



## Twelve tips for completing quality in-training evaluation reports

Nancy Dudek & Suzan Dojeiji

To cite this article: Nancy Dudek & Suzan Dojeiji (2014) Twelve tips for completing quality in-training evaluation reports, *Medical Teacher*, 36:12, 1038-1042, DOI: [10.3109/0142159X.2014.932897](https://doi.org/10.3109/0142159X.2014.932897)

To link to this article: <http://dx.doi.org/10.3109/0142159X.2014.932897>



Published online: 02 Jul 2014.



Submit your article to this journal [↗](#)



Article views: 719



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 4 View citing articles [↗](#)

## TWELVE TIPS

# Twelve tips for completing quality in-training evaluation reports

NANCY DUDEK & SUZAN DOJEJI

University of Ottawa, Canada

## Abstract

Assessing learners in the clinical setting is vital to determining their level of professional competence. Clinical performance assessments can be documented using In-training evaluation reports (ITERS). Previous research has suggested a need for faculty development in order to improve the quality of these reports. Previous work identified key features of high-quality completed ITERS which primarily involve the narrative comments. This aligns well with the recent discourse in the assessment literature focusing on the value of qualitative assessments. Evidence exists to demonstrate that faculty can be trained to complete higher quality ITERS. We present 12 key strategies to assist clinical supervisors in improving the quality of their completed ITERS. Higher quality completed ITERS will improve the documentation of the trainee's progress and be more defensible when questioned in an appeal or legal process.

## Introduction

Work-based assessment (WBA) is thought to be the best method of assessing professional competence (Epstein & Hundert 2002). There are several tools for WBA including the mini-clinical evaluation exercise, direct observation of practical skill, multi-source feedback and in-training evaluation (Govaerts & van der Vleuten 2013). In this article, we focus on in-training evaluation which is documented on an In-training Evaluation Report (ITER) (Turnbull & van Barneveld 2002). ITERS are also referred to as, clinical performance reports, performance assessment forms, clinical performance progress reports and end of clinical rotation reports. They usually consist of a list of items on a checklist or rating scale and written comments. ITERS serve both a formative and a summative role. Effective ITERS provide feedback to the trainee that can be used to modify and develop future performance (Turnbull et al. 1998). On the other hand, given that ITERS document performance they can be used as evidence that a trainee has met a set standard, a summative role.

Physicians who supervise medical trainees have indicated that they want specific training to enhance their ability to complete ITERS (Dudek et al. 2005). This perceived need is consistent with observed needs noted in the literature. There is evidence to suggest that the final assessment (i.e. pass versus fail) written on the ITER is not always consistent with the assessor's judgement of the performance, especially for the poor performer (Cohen et al. 1993; Hatala & Norman 1999; Speer et al. 1996). A significant part of the problem in failing to report poor clinical performance is that supervisors often do

not know what to document when completing an ITER (Dudek et al. 2005).

Several authors (Holmboe et al. 2004; Littlefield et al. 2005), including the Advisory Committee on Educational Outcome Assessment (Swing et al. 2009), have proposed that assessor training is a key component in addressing the problem of quality assessments in residency programs, with some suggesting that rater training may be the "missing link" in improving assessment quality (Holmboe et al. 2011).

Finally, with Competency Based Medical Education (CBME) curricula, there will be an increasing requirement of direct observation and assessment methodologies that reflect trainee performance accurately. A quality ITER will become a critical component of medical trainee assessment with CBME (Bullock et al. 2011).

Dudek et al. identified nine key features of high-quality completed ITERS, eight of which deal with the quality of the written comments suggesting that the focus of improving ITER quality should be on the narrative comments which is a shift from past work which focused on improving ITER quality by improving the reliability of the assigned ratings (Dudek et al. 2008). Various faculty development (FD) programs have demonstrated faculty can be trained to complete higher quality ITERS primarily by focusing on improving the narrative comments (Dudek et al. 2012, 2013a, 2013b). This is in line with a recent strong call in the literature for more emphasis on qualitative assessments, with some even suggesting that narrative descriptions replace numerical ratings for clinical performance (Hanson et al. 2013). Rich narrative evaluations of performance enhance the formative function of ITERS but

*Correspondence:* Nancy Dudek, MD, MEd, The Ottawa Hospital Rehabilitation Centre, Room 1105D, 505 Smyth Road, Ottawa, Ontario, Canada K1H 8M2. Tel: 1 613 7377350 ext. 75596; Fax: 1 613 7396974; E-mail: ndudek@toh.on.ca

also are required for defensible decisions in summative assessments (Govaerts & van der Vleuten 2013).

ITERS should not be the sole form of assessment. Ideally, they should be incorporated into a system of assessment that is designed and conducted at a higher level within an institution. The discussion of such a system is beyond the scope of this article. Rather, here we focus on the individual assessor and present 12 key strategies to assist clinical supervisors in improving the quality of their completed ITERS as higher quality ITERS will contribute to an overall enhancement of the assessment system. Notably, many of these strategies can be applied to improve the quality of the narrative comments included on many other forms of WBAs such as multi-source feedback.

### Tip 1

#### Know your institutional process for in-training evaluation

Medical schools and residency programs have policies regarding ITE. They are typically quite detailed regarding the evaluation *process* (i.e. when the evaluation should occur, that the evaluation is reviewed with the trainee, etc.). The rules also remind supervisors that assessments should align with the learning objectives for that particular clinical activity. Thus, it is imperative that the supervisor be familiar with what these are and ideally discuss them with the trainee at the start of the clinical training period. Following the policies is especially important in the case of an appeal. If the *process* was not properly completed (i.e. missing a mid rotation assessment), often even a well done ITER will not stand up to an appeal (Dudek et al. 2013a).

### Tip 2

#### Identify the relative strengths and weaknesses of the trainee's performance in the ratings

Supervisors need to avoid routinely assigning the same rating to all aspects of the trainee's performance (e.g. all ratings are 4 of 5) as the vast majority of people have some variability in their performance. Variability in the ratings suggests that the assessor took the time to reflect on all of the checklist items rather than just making a global judgement (Dudek et al. 2008).

There are a couple of strategies that can be used to help the clinical supervisor avoid the "straight line" marking for ratings. One suggestion is to complete the comments section first, using the checklist items to frame the comments. This may represent a change in practice for many supervisors as typically the comment section is placed after the ratings. After all the written comments are noted, the particular strengths and weaknesses of the trainee being assessed will be more obvious. This process makes it easier to reflect those strengths and weaknesses in the ratings. Another suggestion is to think about the trainee's performance and define his or her top two and bottom two areas of performance. For a strong

trainee these both might be near the top of the rating scale, such as 4/5 and 5/5. Both indicate a high level of performance, but the different ratings would identify relatively stronger areas of performance (Dudek et al. 2013a).

### Tip 3

#### Provide detailed comments by including specific examples of strengths and weaknesses

Comments are the hallmark of a quality ITER (Dudek et al. 2008). High-quality ITERS include comments that are written in enough detail to enable an independent reviewer (e.g. program director, post-graduate dean) to clearly understand the trainee's performance (Dudek et al. 2008). This is especially important when the trainee appeals the ITER (typically, in the case of a failing ITER). The appeals board will need to ensure that the ITER includes enough information to justify the failing grade. Furthermore, detailed comments on performance enhance the legitimacy of the feedback provided in the ITER for all trainees. Specific and detailed written comments align with the recommended techniques for providing feedback to trainees (Hewson & Little 1998; Sargeant & Mann 2010).

Specific examples of strengths and weaknesses provide evidence for your judgements regarding the calibre of the trainee's performance. These examples will increase the quality of the completed ITER (Dudek et al. 2008). Stating that the trainee has "poor communication skills" is not very useful to the trainee or defensible from the perspective of the governing institution. Commenting on a specific aspect of their skills, such as, "Tendency to use too much medical jargon when explaining issues to patients," is much more useful. However, it becomes most defensible when you can provide an example to validate your assessment: "Tendency to use too much medical jargon when explaining issues to patients. Example – In the patient with an abnormal lesion on the chest X-ray you said it could be an infiltrate, a granuloma, a malignancy..." (Dudek et al. 2013a).

### Tip 4

#### Provide behaviour-based comments

The focus of the comments should be on what the trainee did as opposed to what his or her attitude may or may not have been (Dudek et al. 2013a). In other words, it is not useful for the supervisor to state that the resident has "a poor attitude about working with elderly patients" during his or her geriatrics rotation. Rather, the focus should be on the problem behaviours that the supervisor observed. The resident may well not like working with a geriatric population; nevertheless, there are expectations for a trainee's behaviour when on a geriatrics rotation.

As an example, instead of writing "This resident is lazy," reflect on the behaviours that lead to the conclusion that he or she is "lazy". For example, documenting that "the trainee consistently arrived 15–20 minutes late to clinic, resulting in

the patients' waiting unnecessarily," constitutes a more objective comment on the ITER.

Also, it is very useful to provide the outcome of the specific behaviour as this makes the comment a more powerful piece of corrective feedback for the trainee. For example, in considering a "rude" trainee, the supervisor may document that the trainee displayed inappropriate verbal (raised voice) and non-verbal (rolled eyes) communication skills when working with a junior medical colleague. This consistent behaviour resulted in the junior colleague's being fearful of and avoiding the trainee, which presented a patient safety concern when the junior colleague required support or advice on a patient problem.

## Tip 5

### Include the trainee's response to feedback

Most clinical supervisors agree that how trainees respond to feedback is an important feature of their performance (Dudek et al. 2008). Trainees who accept the provided feedback and adjust their performance accordingly tend to continue to learn in a very positive manner. Alternatively, trainees who become defensive about their performance and dismissive of the feedback tend to be more difficult to teach.

During the rotation, you will have given the trainee some specific feedback on his or her performance in several domains (e.g. communications skills, physical examination skills). The trainee's response to this feedback should be noted. For example: "Trainee tends to respond positively to feedback. Example – noted that trainee tested sensation only in dermatomes as opposed to peripheral nerve territories. This feedback was given to trainee. On observation at a later point during the rotation she had altered her physical examination appropriately" (Dudek et al. 2013a).

Remembering to follow-up on feedback can be challenging. Try keeping a list of feedback items that you provided to the trainee during the rotation, and review the list regularly. Make a point of following up on one item the next time that you have an opportunity to observe the trainee's work. This will also assist you with data collection (refer to Tip 10).

## Tip 6

### Match the ratings to the comments

This seems like a straightforward concept, but it is not uncommon for a program director to review an ITER and need to ask "Which one is right? Are the ratings the real story or the comments?" It is important that the ITER can stand on its own without the need for a follow-up clarifying conversation (Dudek et al. 2008).

Ensure that you read over your report once it is completed and look for inconsistencies. Areas of strength in the ratings should have supportive comments, and likewise for areas of weakness. If you are unsure, have colleagues review a couple of your completed forms to see whether they note any discrepancies (Dudek et al. 2013a).

Alternatively, if you feel that the ITER form does not facilitate adequately your observations and assessments on the rotation, raise this concern with your institution's representative responsible for trainee assessment. Potentially, other supervisors are having similar challenges and the form itself may need to be modified.

## Tip 7

### Provide recommendations for ongoing development

As the ITE process is formative and summative, specific recommendations as to where trainees should focus their learning are essential for quality ITERs (Dudek et al. 2008). This is extremely beneficial for a trainee who has failed a rotation as it provides a "road map" for developing the remediation program for that trainee. Similarly, recommendations for improvement benefit mid and high performing trainees as constructive comments guide their growth (Boehler et al. 2006).

Improving the quality of comments in this area can be accomplished by giving more specifics. For example, a recurrent comment frequently cited in ITERs is "should read more". This recommendation is easy to improve on by simply stating what specific areas the trainee should be reading about (Dudek et al. 2013a).

## Tip 8

### Collect data from other sources

Trainees work with a large number of other health care professionals (e.g. nurses) and non-medical professionals (e.g. administrative assistants) as part of their clinical practice. It is important to collect information regarding trainees' performance from others who have observed them during their rotations. For example, nurses typically will interact directly with residents. It is imperative that we understand whether residents are communicating and collaborating with the nurse in a productive and professional manner. Therefore, we need to advise these other professionals that their feedback on trainee performance will be requested. Other assessment systems can be implemented to capture this information, such as multi-source feedback. However, it is also very useful to collect this information directly if you are the supervising clinician. Given that specific information is needed, the supervisor has to ask specific questions. If a nurse states, "I really like Dr. X. He's great to work with" ask the nurse why he or she feels that way and whether he or she can provide an example to illustrate that impression (Dudek et al. 2013a).

## Tip 9

### Establish on-going data collection (what and how much)

Often clinical supervisors are presented with an ITER to complete, but they simply cannot remember the specific examples (Dudek et al. 2005); rather, they just have a general

impression of the trainee's performance. As outlined in the previous tips, a quality ITER requires a lot of details regarding the trainee's performance. In order to provide this type of detail, the clinical supervisor needs to develop a strategy to record their assessment of trainee performance in-the-moment (Turnbull et al. 1998). Similar to the concept that clinicians need to make notes about patients in order to remember details of findings and management plans, clinical supervisors need to make similar notes on their trainees. It is unreasonable to expect that busy clinicians will be able to remember their trainees without any notes to reference, especially if the clinician is supervising multiple trainees during a rotation.

A useful and feasible strategy is to document, for each clinical encounter, one example of something the trainee did well and one example of something that the trainee did that needs improvement. What is a clinical encounter? A clinical encounter can be defined in any manner that fits with the supervisor's clinical practice. Examples include but are not limited to the following: an OR day, an ER shift, an ambulatory clinic, a night on-call and an inpatient service day. If one assumes five working days in a week, the supervisor should have five specific examples of strengths and five specific areas that need improvement by the end of the week (Dudek et al. 2013a). This amount of detail should enable the assessor to provide some clear examples in the comments section of the ITER. This approach facilitates scenarios where the trainee might work with a number of physicians throughout their week but one physician is responsible for completing the ITER. If each physician simply provides their examples from their clinical encounters to the physician in charge of the ITER all of the data can be compiled.

## Tip 10

### Establish a data storage and retrieval strategy

The next step is to establish a strategy for recording the detailed feedback information. The strategies employed vary with the clinical supervisor's preferences, especially with respect to the use of technology. Each person needs to find a method that is practical, sustainable and useful within the context of his or her practice.

In FD sessions conducted across Canadian universities, faculty members identified the following strategies (Dudek 2013a). (1) Keep a resident paper or electronic file and add to it after each clinical encounter. This often works well for the supervisor in an office setting. After seeing patients the supervisor can simply make notes about the trainee's performance at the end of each day. (2) Use a calendar (either paper or electronic) as it helps to contextualize the notes that you have made. For example, if you always run a particular type of clinic on Tuesdays, you will likely remember more easily the particular incident that prompted that note about the resident. (3) Keep notes on index cards or a note pad (whether electronic or paper) because it is something you carry in your pocket already. (4) Dictate your impressions of the resident's performance. These impressions can be transcribed or simply

played back when you need to complete the ITER (Dudek et al. 2013a).

## Tip 11

### Determine if you think the performance assessed is a pass or fail

Some institutions require the supervising physician to make the final decision regarding the "final grade" whereas others only request a recommendation. In either case the supervising physicians are in an excellent position to judge whether or not a trainee's demonstrated performance on the rotation met the training objectives for that rotation from their perspective. In quantitative assessment there is significant effort made to reduce rater effects given the concern about rater bias (Govaerts & van der Vleuten 2013; Hodges 2013). However, current discussions in the literature suggest that these differences may actually represent alternative but complimentary perspectives on trainee performance and therefore, are valid pieces of information (Govaerts & van der Vleuten 2013). The value of these assessments relies on multiple samples from different assessors supervising the trainee in different contexts (Govaerts & van der Vleuten 2013). Therefore, regardless of whether or not the trainee has performed well or poorly in other clinical contexts, there is great value in an ITER that describes the performance as "passing or failing" from the perspective of that clinical supervisor.

## Tip 12

### Assess and improve your own ITERs

Participating in FD sessions aimed at improving the quality of written comments increases the quantity and quality of the feedback (Dudek et al. 2012; Salerno et al. 2003). FD offices should be encouraged to provide this type of training as better quality narrative comments will improve all types of WBA that include comments (Roberts 2013).

Individually, supervisors can assess the quality of their ITERs by using the Completed Clinical Evaluation Report Rating (CCERR) (Dudek et al. 2008). The CCERR has repeatedly demonstrated strong evidence for validity and provides a reliable rating of the quality of ITERs completed by clinical supervisors (Dudek et al. 2008, 2012, 2013b). A clinical supervisor can simply give a sample of their ITERs to a colleague and have them rate the ITERs using the CCERR. An improvement in future CCERR scores has been demonstrated after clinical supervisors review the results of these ratings (Dudek et al. 2013b).

## Conclusion

ITER completion is an essential component of the clinical supervision process. The clinical supervisor's assessment of trainee performance is of great value to trainees and their respective training programs. Clinical supervisors have the ability to complete this task competently by simply translating their observations and verbal feedback into written comments.

These comments should be specific, behaviour-based and supported by examples. In order to provide these high-quality comments each supervisor needs to find a method to document what they see during the rotation so that this data will be available to them when they are presented with an ITER to complete.

## Notes on contributors

NANCY DUDEK, MD, MEd, is an Associate Professor in the Faculty of Medicine at the University of Ottawa.

SUZAN DOJEJJI, MD, MEd, is an Associate Professor in the Faculty of Medicine at the University of Ottawa.

**Declaration of interest:** The authors report no declarations of interest. The authors alone are responsible for the content and writing of the article.

## References

Boehler ML, Rogers DA, Schwind CJ, Mayforth R, Quin J, Williams RG, Dunnington G. 2006. An investigation of medical student reactions to feedback: A randomised controlled trial. *Med Educ* 40:746–749.

Bullock G, Sherbino J, Hodder R, Watling C. 2011. Assessment. In: A white paper prepared for the Royal College of Physicians and Surgeons of Canada, Future of medical education in Canada. Ottawa: Royal College of Physicians and Surgeons of Canada. Available from [www.royalcollege.ca/portal/page/portal/rc/common/documents/educational\\_initiatives/assessment.pdf](http://www.royalcollege.ca/portal/page/portal/rc/common/documents/educational_initiatives/assessment.pdf).

Cohen GS, Blumberg P, Ryan NC, Sullivan PL. 1993. Do final grades reflect written qualitative evaluations of student performance? *Teach Learn Med* 5(1):10–15.

Dudek N, Marks M, Dojeiji S. 2013a. Completing a quality evaluation report – What clinical supervisors need to know – a faculty development workshop. *MedEdPORTAL*. Available from [www.mededportal.org/publication/9320](http://www.mededportal.org/publication/9320).

Dudek NL, Marks MB, Bandiera G, White J, Wood TJ. 2013b. Quality in-training evaluation reports – Does feedback drive faculty performance? *Acad Med* 88(8):1129–1134.

Dudek NL, Marks MB, Regehr G. 2005. Failure to fail – The perspectives of clinical supervisors. *Acad Med* 80(10):S84–S87.

Dudek NL, Marks MB, Wood TJ, Dojeiji S, Bandiera G, Hatala R, Cooke L, Sadownik L. 2012. Quality evaluation reports: can a faculty

development program make a difference? *Med Teach* 34(11):e725–e731.

Dudek NL, Marks MB, Wood TJ, Lee AC. 2008. Assessing the quality of supervisors' completed clinical evaluation reports. *Med Educ* 42: 816–822.

Epstein RM, Hundert EM. 2002. Defining and assessing professional competence. *JAMA* 287(2):226–235.

Govaerts M, van der Vleuten CPM. 2013. Validity in work-based assessment: Expanding our horizons. *Med Educ* 47:1164–1174.

Hanson JL, Rosenberg AA, Lane JL. 2013. Narrative descriptions should replace grades and numerical ratings for clinical performance in medical education in the United States. *Front Psychol* 4:1–10.

Hatala R, Norman GR. 1999. In-training evaluation during an internal medicine clerkship. *Acad Med* 74(10 Suppl):S118–S120.

Hewson MG, Little ML. 1998. Giving feedback in medical education – Verification of recommended techniques. *J Gen Intern Med* 13:111–116.

Hodges B. 2013. Assessment in the post-psychometric era: Learning to love the subjective and collective. *Med Teach* 35:564–568.

Holmboe ES, Hawkins RE, Huot SJ. 2004. Effects of training on direct observation of medical residents' clinical competence – A randomized trial. *Ann Intern Med* 140(11):874–881.

Holmboe ES, Ward DS, Reznick RK, Katsufraakis PJ, Leslie KM, Patel VL, Ray DD, Nelson EA. 2011. Faculty development in assessment: The missing link in competency-based medical education. *Acad Med* 86: 460–467.

Littlefield JH, DaRosa DA, Paukert J, Williams RG, Klamen DL, Schoofield JD. 2005. Improving resident performance assessment data: Numeric precision and narrative specificity. *Acad Med* 80(5):489–495.

Roberts R. 2013. Free-text comments for multisource feedback. *Clin Teach* 10:417–418.

Salerno SM, Jackson JL, O'Malley PG. 2003. Interactive faculty development seminars improve the quality of written feedback in ambulatory teaching. *JGIM* 18:831–834.

Sargeant J, Mann K. 2010. Feedback in medical education: Skills for improving learner performance. In: Cantillon P, Wood D, editors. *ABC of learning and teaching in medicine*, 2nd edn. Oxford: Blackwell Publishing. pp 29–32.

Speer AJ, Solomon DJ, Ainsworth MA. 1996. An innovative evaluation method in an internal medicine clerkship. *Acad Med* 71:S76–S78.

Swing SR, Clyman SG, Holmboe ES, Williams RG. 2009. Advancing resident assessment in graduate medical education. *J Grad Med Educ* 1:278–286.

Turnbull J, Gray J, MacFadyen J. 1998. Improving in-training evaluation programs. *J Gen Intern Med* 13:317–323.

Turnbull J, van Barneveld C. 2002. Assessment of clinical performance: In-training evaluation. In: Norman GR, van der Vleuten CPM, Newble DI, editors. *International handbook of research in medical education*. London: Kluwer Academic Publishers. pp 793–810.