

Ongoing formative assessment in the training of post graduate students

Rôle de l'évaluation formative continue appliquée aux résidents

Nesrine Mejri, Houda El Benna, Soumaya Labidi, Nouha Daoud, Hamouda Boussen

Medical oncology department. Abderrahmen Mami Hospital / Université Tunis El Manar, Faculté de médecine Tunis,

RÉSUMÉ

Introduction: Il existe de plus en plus de preuve validant le rôle de l'évaluation formative comme méthode d'apprentissage. L'intégration de l'évaluation formative chez les résidents peut être difficile.

Objectif : Décrire une activité d'apprentissage basée sur l'évaluation formative continue chez les résidents d'oncologie.

Méthodes: Nous avons effectué une étude descriptive sur une période de 3 ans, faisant participer 25 résidents en oncologie. La première phase consistait en la préparation des tests conformément aux objectifs du cursus du résidanat puis une deuxième phase d'examen par petits groupes suivi d'une séance de correction interactive immédiate. La troisième phase consistait en la distribution d'une enquête tous les six mois évaluant la satisfaction (1 à 5) des résidents.

Résultats: Vingt-deux tests ont été effectués entre 2016-2018. Au premier test, la médiane des scores était de 51/100 [30-72] avec une moyenne de 53/100. Les scores médians de tous les élèves se sont significativement améliorés entre le premier et le dernier test passant à 68/100 ($t(16)=3,172, p<0,02$) avec une moyenne de 65/100 et une diminution de l'écart entre les résidents. Tous les étudiants ont attribué une satisfaction de 5 à la séance de correction. La majorité ont déclaré que le test avait répondu à leurs attentes (73% 4-5) et l'ont considéré comme ayant un impact sur leur pratique quotidienne (77% 4-5).

Conclusion: L'évaluation formative est une activité éducative associée à une amélioration des scores au fil du temps avec un haut niveau de satisfaction des résidents.

Mots-clés

Apprentissage, résidanat, évaluation formative

SUMMARY

Introduction: There is growing evidence that formative assessment is valuable tool in enhancing learning. Integrating formative assessment into post graduate students can be challenging.

Aim: Authors aimed in this study to describe an ongoing formative assessment activity in post graduates. We reported resident's performance and satisfaction.

Methods: Authors performed an exploratory study over a 3-year period. Twenty five oncology residents participated. The first phase was test preparation by senior oncologists, according to residency curricula then taking the test by a small group of residents with an immediate feedback. The third phase was distribution of a survey each 6 months evaluating resident's perception of the testing.

Results: Twenty two tests were taken by 17 medical oncology, 2 surgical oncology and 4 radiation therapy residents. At the first test, median scores was 51/100 [30%-72%] with a mean of 53/100. Individual scores of each resident improved with time, becoming 68/100 ($t(16)=3.172, p<0.02$) with a mean of 64/100 and decrease between student's scores. All the students rated the correction session with a 5. The majority reported that the test reached their expectations (73% rated4-5), and considered it, as having an impact on their daily practice (77% rated4-5). Residents also considered the test as highly difficult (80% rated4-5).

Conclusion: Ongoing formative assessment showed improvement in overall knowledge of residents with high level of satisfaction.

Key-words

Learning, residency, ongoing formative assessment

INTRODUCTION

Traditional medical education and learning tools are based on didactic conferences with several clinical term courses and placements. Multiple studies have shown that at the end there is no difference in the long-term retention of knowledge between students who attend didactic conferences and those who do not. Various modern educational methods are now adopted such as Problem-based-learning and web-based learning mainly among undergraduate students [1]. Assessment of knowledge which is mainly based on multiple-choice questions or short essays or clinical cases isn't sufficient to assess the competences of the students. Besides, these competences are in progress and needs to be assessed continuously. Multiple recent studies have shown the impact of ongoing assessment on the improvement of the students' skills and competences [2, 3]. Repeated practice in retrieving information from memory seems to greatly enhance future recall and helps students learn about their strengths and weaknesses, it also helps teachers improve their teaching methods [4], providing them additional educational resources. Postgraduate education in medical oncology lasts 4 years in our country. The residents are trained by specialist's staff members from the Faculty of Medicine of Tunis. We aimed to describe an ongoing formative assessment activity in small groups of post graduate oncology residents. We reported residents' performance/satisfaction and the challenges reported by the evaluators.

METHODS

Medical staff:

Our department is specialized in medical oncology, affiliated to university of medicine, university Tunis El Manar. Senior doctors are graded assistant/associate/professors in medical oncology with experience in medical pedagogy and teaching. A group of residents, (between 4 and 6) are assigned each year for a training course within their residency program with a turn-over of 6 months. Residents in surgical oncology and/radiation oncology were allowed to be assigned in our department as part of their 'optional' training courses. They were between the first and the fourth year of their residency. All residents participated to a periodic test on a monthly basis. They gave voluntary informed consent before participating to the study. At the end of each 6-month period, all participants were given a satisfaction questionnaire to evaluate their perception of the relevance of this activity.

Assessment Test:

The test consisted of 20 questions, based the everyday clinical activity of medical oncology. Different types of question formats were used: essay-type, Objective Structured Clinical Evaluation (OSCE), questions and mini-clinical evaluation exercise (mini-CEX). Questions were mixed between ones requiring recognition of facts and others requiring information production. Three senior doctors participated to the preparation and validation of the test, in accordance to the learning objectives in residency curriculum. Questions were routinely updated to the latest national and international clinical practice guidelines.

Pedagogical scenario:

The tests took place once a month. Residents were asked to choose the theme and were informed about the exact date of the test at least a week before. The test duration was 45 minutes. We asked the residents to take the test individually. After that, a correction session with feedback was held immediately (except for 2 tests, where feedback was delayed by 2 weeks) and animated by at least two seniors who participated to test preparation. During the correction session (2-3 hours), each resident read and discussed his response. Seniors explained several resident's interrogations; gave them references for further reading about the topic. Residents also discussed the questions with their colleagues. They were also asked to take notes about the new learned concepts directly on the distributed test document using a dedicated correction zone. After the end of the correction session, seniors collected the copies and gave a score from 0-100% after identifying correct and false answers and using a predefined scoring system. The final step was to give back the copies to residents; the one with the highest score was named the "Champion of the month" as a reward and for motivation purpose. After each test, seniors eliminated from the questions "bank", all questions with <10% correct-response rate and corrected the questions reported as confusing by residents.

Satisfaction assessment:

After the end of each 6-month period, we distributed a questionnaire to evaluate the resident's perception about the test, according to a 5-point Likert scale responses from strongly disagree (0) to strongly agree (5). We asked questions about: the impact on the daily practice, adherence to their expectations, difficulty of the test, quality

of the correction session, scoring system and desire to continue the activity. The survey was anonymous. Evaluators were asked to report three major challenges they faced when dealing with ongoing assessment activity. Statistical analysis: For each test, the score of each resident was registered. Median scores of the first and last score of each resident were calculated. Scores were presented in percentages, the scores improvements too. Two-tailed paired sample t-tests were conducted to examine whether the differences between the first and the last test were statistically significant, $p < 0,05$ was considered significant.

RESULTS

General characteristics: A total of 25 residents participated to the test as following specialities: 17 were attending medical oncology speciality, 2 surgical oncology speciality and 4 radiation oncology speciality. Sixty-eight were female. Most of them (20) were at their fourth year of residency, 1 at the third year, 2 at the second year, one at the first year. Mean age was 31 years-old. During 3 years, 22 different tests were taken by the different groups of residents (5-7 in each group). A median number of 5 tests/student were performed, ranging between 3 and 9. In 8 tests residents chose a dedicated unique theme and in 16 tests they chose general oncology.

Resident's performance: At the first test we observed a wide range of scores. Scores at the first test ranged between 30% and 72% ($\Delta=42\%$). It reflected different levels of knowledge at start point. We could not evaluate the impact of residency year on scores because of the low number of residents in the first/second year of residency. Individual scores of each resident improved with time for most students (23/25) as they repeated the tests. Median scores of all students improved significantly between the first and the last test ($t = 3.172, p < 0.02$). It went up from 51% to 68% ($\Delta=14\%$), as shown in figure 2. The gap between residents decreased from 42% to 21% at the last test.

In the group of medical oncology residents, median scores improved by 14% (from 54% to 68%). When we consider the group of radiation oncology and surgical oncology residents, the median score improved by 10% (48% to 58%).

Satisfaction questionnaire: 37 satisfaction questionnaires were collected (13 residents participated twice over 12 months). All the students rated the correction session

with a 5. The majority reported that the test reached their expectations (73% rated 4-5), and considered it as having an impact on their daily practice (77% rated 4-5). They were also satisfied with the scoring system (76% rated 4-5). However, satisfaction was less with the distribution into 5 items (53% 4-5). Residents also considered the test as highly difficult (80% rated 4-5).

Evaluator interview: The 3 evaluators reported the following majors challenges:

- the ongoing assessment is time/energy consuming, including it as another routine activity needs team organization and will power.
- The optimal choice of questions type is challenging.
- The impact of this activity on patient's management is difficult to assess.

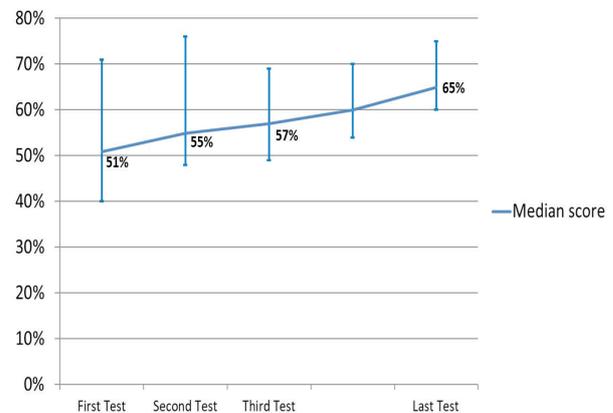


Figure 1: Evolution of the median scores of all residents from the first to the last test

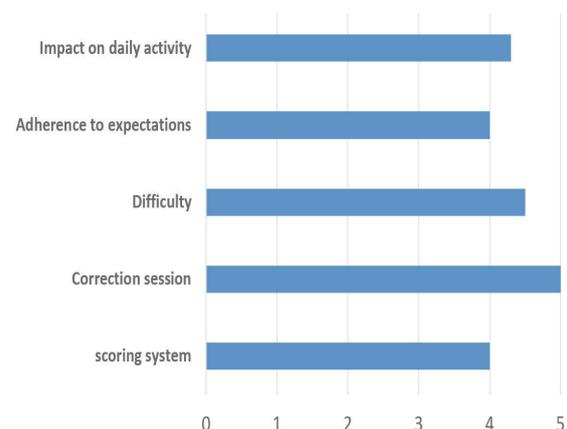


Figure 2: Resident's perception of the test; from 0 for totally disagree to 5 totally agree

DISCUSSION

The authors explored the educational role of ongoing formative assessment of clinical skills in the context of oncology residency. The authors found that ongoing formative assessment is a feasible pedagogical strategy with some challenges. Ongoing formative assessment, lead to an increase of the overall knowledge level of residents and showed a decrease in the gap between learners with time. Residents valued this activity as having an impact on their practice and rated highly the feedback. Those results provide encouraging evidence on implementing routine testing as an educational tool in post graduate residents and were also consistent with prior published research.

The ongoing formative assessment effect is a cognitive term referring to the finding that taking practice tests on studied material promotes greater subsequent learning and retention on a final test compared to more common study strategies [5]. Cognitive psychology laboratory studies by Roediger et al. and Karpicke et al. consistently demonstrated that recalling previously learned information (retrieval practice) enhances the ability to recall the information in the future (retrieval effect) [6, 7]. Expanding retrieval practice promotes short-term retention, but equally spaced retrieval enhances long-term retention [8]. In the systemic review of Green et al, it was highlighted in many clinical trials that learners who are asked to recall information and make an effort in problem solving show better learning, retention, and transfer than students who spend the same time repeatedly studying the same material [9, 10]. More recently, investigators have demonstrated the retrieval effect in health professions education. It was reported in several health domains where trainees exposed to testing (versus studying) demonstrated superior medical knowledge in undergraduate nurses [11, 12] and better skills in cardiopulmonary resuscitation practice [13] and in radiograph interpretation [14], with effects lasting up to 6 months. Karpicke suggested that classroom quizzing may represent a pedagogical strategy to retrieval practice. In fact McDaniel study showed that low-stakes quizzing is an effective way to get to practice active retrieval in the classroom [15]. Implementation of ongoing assessment in workplace for post graduate should be competency based [16] and must also reflect realities of the practice of the specialty.

The extent to which workplace assessment tools can be used to detect and manage underperformance in postgraduate trainees is unclear [17].

Ideally, “tests” should be repeated, spaced over time, utilize items that require production of information, and include feedback with the correct responses. It has been demonstrated that repeated studying after learning had no effect on delayed recall, but repeated testing produced a large positive effect [18]. Lyle and Crawford found in their study that quizzing enhanced student performance in class, but they also reported that students liked the frequent quizzing and viewed it as a valuable learning tool [19].

The best testing interval is still unknown, it ranged from few days to few weeks among several studies [9, 20]. In our exploratory study, the tests were performed in small groups of residents (maximum 6). Indeed, previous studies showed that small classes induce a better student’s achievement, but it does not seem to have an impact on reducing the gap between students [21, 22]. The magnitude of testing effect depends on several factors such as test format, differences among participants, context of the study (laboratory, classrooms)... More convincing data would come from comparing practice tests to other non-testing activity matched for time and content. In our study tests were repeated monthly, it was the time balance that we had between a busy clinical activity schedule and necessary time for a new test preparation.

The best test format is still not clear. Test that required use of knowledge to solve problems using multiple choice questions, short answer questions and easy and/or key features questions were shown to be more effective than test requiring just recognition of facts [23]. However, Multiple choice questions are used to assess factual knowledge not skills which remains the most important objective of educational tools. On the other hand, mini-CEX is a well accepted tool that can be used in formative assessment within a clinical training environment [24]. Lorwald et al. reported that it was feasible in a surgical postgraduate students [25]. Items that require production of information (short answer, essay) perform better as well [26, 27]. In our study students valued this constructive approach that required them to reflect on their current understandings of knowledge and practice through active engagement. This testing effect applies better when results are without direct consequences for the learner [5]. In our study the only consequence was naming the best

score as “the champion of the month”, which had created a dynamic and enthusiasm over the test.

They also gave the best rating to the correction session, which represents feedback. Feedback should include the correct answer since learners were exposed to true and false information in the test [28, 29]. Feedback also enhances the retention of correct but low confidence responses [30]. Several studies suggest that immediate feedback may be more effective than delayed feedback [31, 32]. However, the exact effect of feedback on learning and retention is rarely investigated. We conclude that repetitive testing of residents during their residency program could be a useful learning tool. Its exact effect and the best test format should be investigated on larger studies with specific methodology.

One of the challenges reported by the evaluators concerned active engagement to this activity. Joshi et al. evaluated the factors that determines active engagement to formative assessment and reported that engaging in this activity is complex and quite a challenge to both students and teachers. Increased acceptability along with the effective implementation of the assessment structure, individual perspectives on feedback, a supportive learning environment and credibility of feedback are all important in this process [33]

Some limitations of our study should be highlighted. The sample size is small which weakens to power of the conclusions, the oncology residency curricula in our country is limited to a small number of residents each year which makes it difficult to include more participants. The lack of a comparator with a non-testing makes it difficult to relate the results mainly to the testing procedure.

Conflict of interest: authors have no conflict of interest to declare

Acknowledgement to all who helped to this work: Manel Dridi, Manel Haj Mansour, Imen Abdelli, Yassine Bel Haj, Haifa Rachdi, Meher Nasri, Salma Gatria, Azza Guebsi, Mehdi Benna, Fahmi M'ghirbi, Yosra Berrazegua, Essia Mezni, Khalil Béhi, Ghita Omelgheith, Khadija Ben Zid, Wafa Sbika, Marouen Benna, Fatma Guemazi, Meriem Saadi, Chaima Bousrih, Imen Harhira, Kaouther Herbegue, Aicha Keskes, Selim kherouf, Hamza Ouaz.

REFERENCES

1. Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: a systematic review. *Med. Educ.* 2014; 48(2):124–135.
2. Pinsk M, Karpinski J, Carlisle. Introduction of competence by design to Canadian nephrology postgraduate training. *Canadian Journal of Kidney Health and Disease.* 2018; 5:2054-59
3. Wagner N, Acai A, McQueen SA et al. Enhancing Formative Feedback in Orthopaedic Training: Development and Implementation of a Competency-Based Assessment Framework. *J. Surg. Educ.* 2019; 76(5):1376–1401.
4. Ong AM-L, Fong WW-S, Chan AK-W et al. Using the Postgraduate Hospital Educational Environment Measure to Identify Areas for Improvement in a Singaporean Residency Program. *J. Grad. Med. Educ.* 2019; 11(4s):73–78.
5. Henry L, Roediger, Jeffrey D, Karpicke. The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science.* 2006;1(3), 181-210
6. Karpicke JD, Grimaldi PJ. Retrieval-Based Learning: A Perspective for Enhancing Meaningful Learning. *Educ. Psychol. Rev.* 2012; 24(3):401–418.
7. Roediger III, HL, Butler, AC. The critical role of retrieval practice in long-term retention - ScienceDirect. *Trends in cognitive sciences.* 2011;15(1), 20-27
8. Wormald BW, Schoeman S, Somasunderam A, Penn M. Assessment drives learning: An unavoidable truth? *Anat. Sci. Educ.* 2009; 2(5):199–204.
9. Green ML, Moeller JJ, Spak JM. Test-enhanced learning in health professions education: A systematic review: BEME Guide No. 48. *Med. Teach.* 2018; 40(4):337–350.
10. Adesope OO, Trevisan DA, Sundararajan N. Rethinking the Use of Tests: A Meta-Analysis of Practice Testing. *Rev. Educ. Res.* 2017; 87(3):659–701.
11. Larsen DP, Butler AC, Iii HLR. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med. Educ.* 2009; 43(12):1174–1181.
12. Messineo L, Gentile M, Allegra M. Test-enhanced learning: analysis of an experience with undergraduate nursing students. *BMC Med. Educ.* 2015; 15(1):182.
13. Kromann CB, Bohnstedt C, Jensen ML, Ringsted C. The testing effect on skills learning might last 6 months. *Adv. Health Sci. Educ.* 2010; 15(3):395–401.
14. Baghdady M, Carnahan H, Lam EWN, Woods NN. Test-enhanced learning and its effect on comprehension and diagnostic accuracy. *Med. Educ.* 2014; 48(2):181–188.
15. McDaniel MA, Agarwal PK, Huelser BJ et al. Test-enhanced learning in a middle school science classroom: The effects of quiz frequency and placement. *J. Educ. Psychol.* 2011; 103(2):399–414.
16. Implementing competency-based medical education: Moving forward: *Medical Teacher.* Vol 39, No 6. [https://www.tandfonline.com/doi/abs/10.1080/0142159X.2017.1315069].
17. Barrett A, Galvin R, Steinert Y et al. A BEME (Best Evidence in Medical Education) systematic review of the use of workplace-based assessment in identifying and remediating poor performance among postgraduate medical trainees. *Systematic reviews.* 2015; 4(1):65.
18. Karpicke JD, Roediger HL. The Critical Importance of Retrieval for Learning. *Science.* 2008; 319(5865), 966-968

19. Lyle KB, Crawford NA. Retrieving essential material at the end of lectures improves performance on statistics exams. *Teach. Psychol.* 2011; 38(2):94–97.
20. Kreiter CD, Green J, Lenoach S, Saiki T. The overall impact of testing on medical student learning: quantitative estimation of consequential validity. *Adv. Health Sci. Educ.* 2013; 18(4):835–844.
21. Akabayashi H, Nakamura R. Can Small Class Policy Close the Gap? An Empirical Analysis of Class Size Effects in Japan. *Jpn. Econ. Rev.* 2014; 65(3):253–281.
22. Konstantopoulos S. Do Small Classes Reduce the Achievement Gap between Low and High Achievers? Evidence from Project STAR. *Elem. Sch. J.* 2008; 108(4):275–291.
23. McConnell MM, St-Onge C, Young ME. The benefits of testing for learning on later performance. *Adv. Health Sci. Educ.* 2015; 20(2):305–320.
24. Sharma Y, Patel J. Feasibility of Mini-CEX as Formative Assessment Tool: A Pilot Study. *Indian Journal of Health Sciences and Care.* 2018; 5(2): 55-59.
25. Lörwald A, Lahner F, Greif R et al. Factors influencing the educational impact of Mini-CEX and DOPS: a qualitative synthesis. *Medical teacher.* 2018; 40(4):414-420.
26. Pyc MA, Rawson KA. Testing the retrieval effort hypothesis: Does greater difficulty correctly recalling information lead to higher levels of memory? *J. Mem. Lang.* 2009; 60(4):437–447.
27. Karpicke JD, Aue WR. The Testing Effect Is Alive and Well with Complex Materials. *Educ. Psychol. Rev.* 2015; 27(2):317–326.
28. Butler AC, Roediger HL. Feedback enhances the positive effects and reduces the negative effects of multiple-choice testing. *Mem. Cognit.* 2008; 36(3):604–616.
29. Butler AC, Karpicke JD, Roediger III HL. The effect of type and timing of feedback on learning from multiple-choice tests. *J. Exp. Psychol. Appl.* 2007; 13(4):273–281.
30. Butler AC, III HLR. Testing improves long-term retention in a simulated classroom setting. *Eur. J. Cogn. Psychol.* 2007; 19(4–5):514–527.
31. Dihoff RE, Brosvic GM, Epstein ML, Cook MJ. Provision of Feedback During Preparation For Academic Testing: Learning Is Enhanced by Immediate But Not Delayed Feedback. *Psychol. Rec.* 2004; 54(2):207–231.
32. Epstein ML, Lazarus AD, Calvano TB et al. Immediate Feedback Assessment Technique Promotes Learning and Corrects Inaccurate first Responses. *Psychol. Rec.* 2002; 52(2):187–201.
33. Sharma S, Sharma V, Sharma M et al. Formative assessment in postgraduate medical education-Perceptions of students and teachers. *International Journal of Applied and Basic Medical Research.* 2015; 5(Suppl 1):66.