

Received 2016-10-06
Revised 2016-11-11
Accepted 2017-01-13

Teachers Evaluation Methods in Medical Education: Round Views of Faculty Members and Educational Experts

Peivand Bastani¹, Soudabeh Vatankhah², Ali TaherNejad³, Afsaneh Ghasemi⁴✉

¹Health Human Resources Research Center, School of Management and Medical Informatics, Shiraz University of Medical Sciences, Shiraz, Iran

²Health Services Management Department, School of Health Management and Information Sciences, Iran University of Medical Sciences, Tehran, Iran

³Baghiatallah University of Medical Sciences, Tehran, Iran

⁴Fasa University of Medical Sciences, Fasa, Iran

Abstract

Background: Since there is no agreement on the best approach of teachers' evaluation, this study was conducted to determine medical teachers' evaluation methods and clarify the viewpoints of Iranian faculty members toward them. **Materials and Methods:** A mix method study was conducted in two phases, systematic review and survey, in Tehran University of Medical Sciences on 400 faculty members. In phase one, 24 studies were analyzed among 1520 and based on that, the viewpoints of faculty members about 14 methods were assessed through a validated questionnaire. Independent t-test and one-way ANOVA were used for data analysis. **Results:** The participants' age mean was 48.62±5.23 and most of them were assistant professors (121/36.01%). About 280 participants (83.3%) chose "mixed method rating" as the best way of evaluation; 68.7% of the participants though "student rating" cannot be an appropriate indicative for evaluating teachers' performance. The findings indicated statistical relationships between the average of some evaluation methods (student rating, peer evaluation, self-ratings, teaching scholarship, teaching awards) and the faculty members' gender ($P<0.05$). There was also a significant relationship in average of student rating, peer evaluation, mentor's advice and self-ratings with participants' age ($P<0.05$). **Conclusion:** None of the evaluation methods can be sufficient to show a correct status of teachers' performance. It is obvious that mix method evaluation as a combination of different measures and methods can be considered as a comprehensive approach; it is recommended to be applied in this university, and then compare teachers' satisfaction and performance before and after this transition. [GMJ.2017;6(3):233-39] DOI:10.22086/gmj.v0i0.725

Keywords: Evaluation; Medical Education; Higher Education; Performance

Introduction

In the last two decades, the significance of teaching evaluation has been emphasized

in higher education and medical education as well. For this reason, many medical schools and universities have searched for ways to effectively and constructively evaluate perfor-

GMJ

©2017 Galen Medical Journal
Tel/Fax: +98 71 36474503
PO Box 7193616563
Email: info@gmj.ir



✉ **Correspondence to:**
Afsaneh Ghasemi, Fasa University of Medical Sciences,
Fasa, Iran
Telephone Number: +989177285617
Email Address : naz7899@yahoo.com

mances of their faculty members [1]. Furthermore, as the faculty members are considered as the most important elements of the higher education systems, designing an appropriate and suitable evaluation system for evaluating their performance can be supposed as a significant indicator for the whole education process [2]. In spite of different findings on the topic of teaching effectiveness and different methods of evaluating teachers' performance, there is no agreement on the best approach [3]. Moreover, although student ratings have dominated as the primary and almost only measure of teaching performance for the last 50 years in many countries, evidences show that most of the Iranian Universities of Medical Sciences already use student ratings for their summative decisions [4]. As universities continue to become more student-oriented, students' perceptions of higher educational facilities and services are becoming more important [5]. Different evidences show that scale rating by students cannot be the only source of teacher evaluation [6]. In another word, evidence of teaching effectiveness is used not only to evaluate student experience and outcomes, but also to substantiate applications for promotion. Peer evaluation, learner evaluation and teaching portfolios are considered as the main methods of promotion in medical education [7]. Peer evaluation can be considered not only as a source of formative feedback but also as a reflective process for teachers and a qualitative evidence of student evaluation [8]. Although, evidences indicate that time constraints, busy workloads and fears of scrutiny and criticism are the main barriers for medical teachers participating in the process [9]. Teaching portfolios are mentioned as another reflective practice in medical evaluation which can be used as an effective tool for assuring life-long learning, as it is considered in medical education [10]. Rather than these methods -student rating, teaching portfolios and peer ratings- there are various ways of teacher evaluation as follows: external expert ratings, self-ratings, videos, student interviews, alumni ratings, employer ratings, mentor's advice, administrator ratings, teaching scholarship, teach-

ing awards and learning outcome measures [11-13]. Each has its potential strengths and restrictions depending on various elements like contingency of situation, nature of the class and students differences [7]. Regarding the variety of teacher evaluation methods and different preferences of applicants for applying each method considering their acceptance, usability, simplicity, costs, etc., this study was conducted to clarify the viewpoints of Iranian faculty members who are affiliated with Tehran University of Medical Sciences as one of the major medical universities in this country toward different teacher evaluation methods to present a practical evidence for Iranian policy makers and those who have a similar context to evaluate medical faculty members more effectively and accurately.

Materials and Methods

This study was conducted in two separate methodological phases as follows: In the first phase, a comprehensive review of literature was carried out using library and internet search to identify and summarize the most important methods for evaluating faculty members' performance in medical education. The following search engines and databases were searched: Google Scholar, PubMed, ISI web of science, Scopus, Embase, ProQuest, and Iranian National Library Of Medicine (INLM), using a group of MeSH terms and keywords pertaining to teacher evaluation, faculty member evaluation, medical education, and higher education. Searches were conducted using Boolean operators OR/AND between main phrases, and the mentioned keywords were extracted from specific themes of the topic under study. The applied search strategy for this phase of the study is shown in Table-1. Beside articles, extracted guides, blueprints, manuals and reports were also included. Moreover, reference lists of all relevant resources were interrogated as a part of the search strategy. The search strategy was limited to English resources and only to the first two pages of search engine results with no time limitation. Paper-based reports on teacher evaluation methods and also gray literature were not included in the search strategy.

The search for teacher evaluation methods in medical education terms took place between Nov 1, 2013 and Aug 20, 2015 and resulted in 1520 resources. The resulting resources were evaluated, based on their relevancy to the study. This step resulted in the exclusion of 1480 resources which were out of scope and inclusion of 40 studies from reference lists of retrieved resources. Finally, 24 studies were identified as relevant. At the end of this phase, 14 items were extracted as the main methods for evaluating teachers' performance in medical education. In the second phase of the study, a questionnaire was designed to assess the range of agreement with each of the 14 items extracted from the first phase of the study, from those teachers affiliated with Tehran University of Medical Sciences points of view. The questionnaire contained two parts: first for teachers' demographic information such as gender, age, educational degree, academic rank and the school they were affiliated to. The second section was related to 14 evaluation methods consisting of student rating, peer evaluation, teaching portfolios, external expert ratings, self-ratings, alumni ratings, employer ratings, mentor's advice, administrator ratings, learning outcome measures, teaching scholarship, teaching awards, non-participant observation (videos) and mixed method rating; the teachers were requested to clarify their opinions in a Likert scale. The scale included five options from completely agree to completely disagree. Reliability of the questionnaire was checked using Cronbach's alpha coefficient for 30 completed questionnaires and α was calculated 0.78 which indicated the acceptable reliability. Confirming face and content validity of the

questionnaire, the preplanned draft was sent to seven experts (two in epidemiology and methodology, three in medical education and two in health education) via their electronic posts and they were requested to present their comments about the questions. Telephone reminders were used a week after sending the questionnaire electronically and finally their opinions were summarized and necessary changes were applied until the questions were finalized. The research population for this phase of the study was all faculty members from 10 schools affiliated with Tehran University of Medical Sciences (Schools of Public Health, Nursing and Midwifery, Medicine, Allied Medical Sciences, Pharmacy, Dentistry, Nutrition, Rehabilitation Sciences, Advanced Technologies in Medicine and Traditional Medicine). Three schools of this university including evening courses, international campus and virtual school were excluded from the study. Stratified sampling was used to select participants from each school and 400 teachers were chosen according to Table-2.

Table 2. Distribution of Participants According to their Schools

Schools	Sample size
Public Health	50
Nursing and Midwifery	50
Medicine	100
Allied Medical Sciences	40
Pharmacy	40
Dentistry	40
Nutrition	25
Rehabilitation Sciences	25
Advanced Technologies in Medicine	20
Traditional Medicine	10

Table 1. Search Strategy

Search Engines and Databases:

Google Scholar - PubMed - ProQuest - ISI web of science - Scopus, Embase - Iranian National Library Of Medicine (2000 to present)

Limitations: Language (only resources with at least an abstract in English)

Date: Up to Aug 2015

Strategy: #1 AND #2

#1	Teacher Evalu*Methods, Faculty Member* Evalu* Methods
#2	Medical Education OR Higher Education OR Undergraduate Education OR Postgraduate Education

The professors' degrees were as follows: 121 (36.01%) were assistant professors, 94 (27.98%) were associate professors and 41 (12.2%) were professors, while the others (70.21%) were faculty coaches. Questionnaires were distributed among selected teachers and after justifying the aim and significance of the study and also assuring the confidentiality of data and finally achieving their voluntary verbal consent, they were requested to answer the questions. In the study, 336 questionnaires were completely filled (response rate = 84%) and the related data were entered into SPSS version 16 and analyzed; descriptive statics and analytical ones like independent t-test and ANOVA were assessed at significance level of 0.05.

Results

Results showed that among 336 participants of the study, 200 were males and 136 were females (59.52% and 40.48%, respectively). The average of age was 48.62 ± 5.23 and most of them were assistant professors (121 participants, 36.01%) and associate professors (94 participants, 27.98%) respectively, while the least participants were 41 full professors (12.2%). Moreover, most of

the participants were affiliated with medicine school (65 teachers, 19.34%) while the least participants were related to Advanced Technologies in Medicine School (2.98%). Other findings indicated that among the 14 teacher evaluation methods, 280 participants (83.3%) chose "mixed method rating" as the best way of evaluating and "external expert ratings" and "peer evaluation" were considered as the second and third options by the research participants, respectively. Other results show a good acceptance of the new evaluation techniques by the participants; for example, about 55% of the present faculty members thought that "non-participant observations" and "mentor's advice" can be considered as good evaluation methods. On the other hand, about 60% of the participants chose "employer rating" and "administrator rating" as the last preferred options for their evaluation. At the same time, 68.7% of the present faculty members thought that student rating cannot be an appropriate indicative for evaluating teachers' performance lonely (Table-3). Findings presented in Table-4 indicate that there was a statistical relationship between the average of some evaluation methods (student rating, peer evaluation, self-ratings,

Table 3. Participants' Viewpoints about Teacher Evaluation Methods

Teacher evaluation methods	Frequency of agreed participants		Frequency of disagreed participants		Frequency of neuter participants	
	No.	%	No.	%	No.	%
Student rating	57	17	231	68.7	48	14.3
Peer evaluation	240	71.7	50	14.9	46	13.7
Teaching portfolios	150	44.6	174	51.8	12	3.6
External expert ratings	243	72.3	80	23.8	13	3.9
Self-ratings	178	53	100	29.8	58	17.2
Alumni ratings	164	48.8	159	47.3	13	3.9
Employer ratings	100	30	200	60	36	10
Mentor's advice	182	54.2	52	15.5	102	30.4
Administrator ratings	90	27	195	58	51	15
Learning outcome measures	170	50.6	100	29.8	66	19.6
Teaching scholarship	159	47.3	159	47.3	18	5.4
Teaching awards	158	47	89	26.5	89	26.5
Non-participant observation (videos)	190	56.5	69	20.5	77	22.9
Mixed method rating	280	83.3	16	4.7	40	12

teaching scholarship, teaching awards) and gender of the faculty members ($P < 0.05$). Other findings showed statistical relationships in student rating and peer evaluation with the participants' academic rankings ($P = 0.002$ and $P < 0.001$, respectively). There was also significant relationships in self rating, teacher scholarship and alumni rating with educational degree ($P < 0.05$) and also in average of student rating, peer evaluation, mentor's advice and self-ratings with participants' age ($P < 0.05$).

Discussion

According to the importance of teacher evaluation in higher education and specially medical education, many medical schools and experts in the scope of medical education have searched for the most effective and constructive methods to evaluate the faculty members' performance [14]. In this regard, the present study tried to summarize different propose methods for teacher evaluation presented in the literature. Findings showed that among these evaluation methods, the present participants preferred mixed method rating, external expert ratings and peer evaluation. Instead, they thought that employer rating and administrator rating cannot be assumed as effective

evaluation methods. At the same time, they believed that although students' viewpoints as the main stakeholders of the teaching process can be helpful and significant, student rating cannot be effective alone and it should be applied along with other methods. In this regard, Bastani *et al.* (2013) showed that mixed method evaluation is the only way ending in comprehensive feedback of teaching quality and matches 360 degree evaluation, and student rating is not enough for teacher evaluation [5]. Despite these findings, Safavi *et al.* (2013) demonstrated that student rating can have the efficacy for evaluating theoretical teaching in the medical sciences faculties and defined the influenced aspects of teaching and administrative practices in such faculties [4] though the efficiency of this method in the field, laboratory, etc., which should be investigated by furtherer studies. Other findings such as Schiekirka *et al.* (2012) emphasized on the importance of applying students' perception about their teachers' quality of work and in this regard, they claimed that paying attention to asking about teachers' outcomes along with their characteristics can be helpful [15]. Berk (2009) presented the 360 multisource feedback model to evaluate teaching and professionalism and claimed that this model

Table 4. Participants' Viewpoints about Teacher Evaluation Methods According to their Genders

Teachers evaluation methods	Mean \pm SD (Male)	Mean \pm SD (Female)	P Value
Student rating	40.1 \pm 0.59	3.92 \pm 0.85	< 0.001
Peer evaluation	4.18 \pm 0.79	4.25 \pm 1.1	0.002
Teaching portfolios	3.33 \pm 0.65	3.5 \pm 0.87	0.3
External expert ratings	4.36 \pm 0.1.2	4.39 \pm 0.92	0.06
Self-ratings	3.4 \pm 0.92	3.15 \pm 1.1	< 0.001
Alumni ratings	3.25 \pm 0.95	3.21 \pm 1.06	0.06
Employer ratings	2.95 \pm 0.89	3.1 \pm 0.75	0.3
Mentor's advice	3.89 \pm 0.85	3.95 \pm 1.1	0.07
Administrator ratings	2.9 \pm 0.59	2.75 \pm 0.98	0.36
Learning outcome measures	3.75 \pm 0.5	3.52 \pm 0.59	0.4
Teaching scholarship	3.65 \pm 0.68	3.78 \pm 0.79	0.003
Teaching awards	3.6 \pm 1.2	3.23 \pm 0.95	0.002
Non-participant observation (videos)	3.99 \pm 1.1	4.1 \pm 0.98	0.25
Mixed method rating	4.65 \pm 0.95	4.4 \pm 0.37	0.22

*min = 0, max = 5

can be considered as a useful framework for evaluation of faculty teaching performance along with their professionalism [12]. In this 360 degree evaluation model, Berk supposed that a faculty member in a medical context must be evaluated through different clients, colleagues, students, patients, etc., but the present participants emphasized on applying a mixed method of evaluation methods to have a better demonstration of the teacher. Similar to the present results, Aburawi *et al.* (2014) concluded that applying students' perceptions needs some pre-requisite such as establishing a culture of trust among all the stakeholders. In another word, their results emphasized that teacher evaluation, especially from the students' points of view, may have different results for students and their teachers [16]. In addition to what was discussed, Gimbel *et al.* (2011) believed that providing correct feedback about faculty members' evaluation scores can help them to improve their teaching skills and solve the probable problems through their teaching process or classroom environment [17]. This can be very important for the present setting and other Iranian medical universities that only use an electronic system with some restricted questions, which the students must answer before the end of their semester and the teacher can see his/her evaluation score through the stated electronic system just after finalizing the students grades. In conclusion, it seems that the evaluation of faculty teaching performance is complex and it cannot be done applying a unique method. In this regard, most academic medical centers prefer to use the open evaluation format as a better determi-

nant for judging teachers' performance [18]. Furthermore, using quantitative measures along with the qualitative ones may be acceptable as a model to evaluate the effectiveness of teachers' performance [19]. This study had some limitations; first, it was a quantitative study applying a self-answering questionnaire; integrating this design with a qualitative method through semi-structured interviews may help with achieving more reliable and in-depth responses. Restricted study population was another limitation. It is recommended to design national studies in this regard.

Conclusion

According to the present results, none of the evaluation methods can be sufficient to show a correct status of teachers' performance. It is obvious that mix method evaluation as a combination of different measures and various methods can be considered as a comprehensive method and it is recommended to be applied in this university and those Iranian universities with the same setting and compare teachers' satisfaction and performance before and after this transition.

Acknowledgements

This manuscript is supported by Tehran University of Medical Sciences as a research proposal titled "designing a contingency model of faculty members' evaluation system" with the approved number of 91-130-2609.

Conflict of Interest

There was no conflict of interest in the study.

References

1. Dush SK, Patro S, Behera BK. Teaching Methods and Its Efficacy an Evaluation by the Students. *J of Indian Acad of Forensic Med.* 2013; 35(4):321-4.
2. Bland CJ, Wersal L, VanLoy W, Jacott W. Evaluating faculty performance: A systematically designed and assessed approach. *Acad Med.* 2002; 77(1):15-30.
3. Bastani P, Amini M, TaherNejad K, Shaarbafchizadeh N. Faculty members' viewpoints about the present and the ideal teacher evaluation system in Tehran University of Medical Sciences. *J Adv Med Educ Prof.* 2013;1(4):140-7.
4. Safavi SA, Kamariah AB. The Utility of Student Evaluations for Medical Sciences Teachers and Administrators. *Int J of Bus, Soc and Sci Res.* 2013; 4(13):266-72.
5. Bastani P, Amini M, TaherNejad A, Rouhollahi N. The Tehran University

- of Medical Sciences Faculty Members' Viewpoints about the Teachers' Evaluation System: A Qualitative Study. *J of Torbat Heiydariyeh University of Med Sci.* 2014; 2(1):7-16. [Persian]
6. Azizi K, Aghsmolaei T, Parsa N, Dabbaghamanesh T. Comparison of differences in performance evaluation of faculty by students with faculty's self-assessment. *Med Educ Prof.* 2014; 2(3): 108-13.
 7. Bastani P, Rouhollahi N, Tahernejad A. Validity and Reliability of Teachers' Evaluation Questionnaires from Students Point of View in Tehran University of Medical Sciences. *Biannual Medical Education. J of Babol University of Med Sci.* 2014;3(1):7-14. [Persian]
 8. Siddiqui ZS, Jonas-Dwyer D, Carr SE. Twelve tips for peer observation of teaching. *Medical Teacher.* 2007; 29:297-300.
 9. Hammersley-Fletcher L, Orsmond P. Evaluating our peers: is peer observation a meaningful process?. *Stud in High Educ.* 2004; 29:489-503.
 10. Adshead L, White P, Stephenson A. Introducing peer observation of teaching to GP teachers: a questionnaire study. *Med Teach.* 2006; 28(2):68-73.
 11. Belcher R, Jones A, Smith LJ, Vincent T, Naidu SB, Montgomery J, et al. Qualitative study of the impact of an authentic electronic portfolio in undergraduate medical education. *BMC Med Educ.* 2014, 14:265-72.
 12. Berk RA. Using the 360° multisource feedback model to evaluate teaching and professionalism. *Med Teach.* 2009; 31: 1073-80.
 13. Berk RA. Survey of 12 strategies to measure teaching effectiveness. *Int J of Teach and Learn in Higher Educ.* 2005; 17(1):48-62.
 14. Wen Sh H, Xu JS, Carline JD, Zhong F, Zhong YJ, Shen Sh J. Effects of a teaching evaluation system: a case study. *Int J of Med Educ.* 2011; 2:18-23.
 15. Schiekirka S, Reinhardt D, Heim S, Fabry G, Pukrop T, Anders S, et al. Student perceptions of evaluation in undergraduate medical education: A qualitative study from one medical school. *Med Educ.* 2012, 12:45-52.
 16. Aburawi E, McLean M, Shaban S. Evaluation of Faculty. Are medical students and faculty on the same page? *Sultan Qaboos Univ of Med J.* 2014, 14(3): 361-68.
 17. Gimbel RW, Cruess DF, Schor K, Hooper TI, Barbour GL. Faculty performance evaluation in accredited U.S public health graduate schools and programs: a national study. *Int J of Med Educ.* 2011; 2:18-23.
 18. Afonso NM, Cardozo LJ, Mascarenhas OAJ, Aranha ANF, Shah Ch. Are Anonymous Evaluations a Better Assessment of Faculty Teaching Performance? A Comparative Analysis of Open and Anonymous Evaluation Processes. *Fam Med.* 2005;37(1):43-7.
 19. Roos M, Kadmon M, Kirschfink M, Koch E, Jnger J, Strittmatter-Haubold V, et al. Developing medical educators _ a mixed method evaluation of a teaching education program. *Med Educ online.* 2014; 19: e23868.