

Feasibility, relevance and effectiveness of teaching and assessment of ethical status and communication skills as attributes of professionalism

Abid Jameel,¹ Sahibzada Mahmood Noor,² Shahid Ayub,³ Sobia Sabir Ali,⁴ Yoon Soo Park,⁵ Ara Tekian⁶

Abstract

Objective: To examine the feasibility and effectiveness of teaching and assessing professionalism in a developing country.

Methods: The pre-intervention and post-intervention study was conducted from January to August 2012 and comprised 7 workshops of three days each that were held at four teaching hospitals of Khyber Pakhtunkhwa province in Pakistan. Overall, there were 10 Objective Structured Clinical Examination stations and 10 written scenarios. After the pre-test, workshop was held on various aspects of professionalism which was considered 'intervention', and it was followed by a post-test similar to the pre-test at the end of day 3. Stata 12 was used for all statistical analyses.

Results: There were 136 postgraduate residents in the study. The correlation between Objective Structured Clinical Examination stations and written exam for pre-test was 0.42 ($p < 0.001$), while for post-test the correlation was 0.17 ($p = 0.046$). Correlation between pre-test written and post-test Objective Structured Clinical Examination stations was 0.23 ($p = 0.001$), but the correlation between pre-test Objective Structured Clinical Examination stations and post-test written was not significant ($p > 0.05$). The standardised effect size for the adjusted regression was 0.37 for both comparisons ($p < 0.001$). Mean pre-test scores were 38.13 \pm 13.13% vs. 76.50 \pm 14.4% for the post-test score ($p < 0.001$).

Conclusion: Although post-test scores increased significantly both for the Objective Structured Clinical Examination stations and the written scenarios, the former has shown a higher reliability compared to the written test. Furthermore, teaching and assessment of professionalism was found relevant, effective and feasible in resource-constrained countries. Teaching and assessment of professionalism has become globally relevant and is recommended to be included in the curricula of medical institutions.

Keywords: Professionalism, Physician-patient relationship, Professional development, Postgraduate training, Curriculum development, OSCE, Pakistan, Developing countries. (JPMA 65: 721; 2015)

Introduction

According to the unwritten "societal contract" between the medical profession and the society, societal expectations of the medical profession include acquisition and maintenance of professional competence and skills throughout their career as well as professional values such as honesty, integrity, high moral standards, accountability and an altruistic behaviour.¹ However, unethical and unprofessional conduct among physicians has been reported in literature.^{2,3} In one report, 61% final year medical students reported unethical behaviour by their teachers.³ Another study with second year medical residents regarding unethical behaviour showed falsification of patient records and mistreatment of patients during the first year of residency.² Therefore,

there is a need to improve the professional behaviour of doctors in order to regain trust in the society.

As a result, most accreditation bodies as well as medical associations throughout the developed world have come up with their own documents on professionalism. These include "A Physician's Charter" by the American Board of Internal Medicine (ABIM) and the European Federation of Internal Medicine (EFIM) and "Medical Professionalism in a Changing World" by the Royal College of Physicians, London.^{4,5} Teaching and assessment of professionalism have found a place among the essential competencies of the Accreditation Council for Graduate Medical Education (ACGME) and the Canadian Medical Education Directives for Specialists (CanMEDS).^{6,7} Incorporation of professionalism in Global Minimum Essential Requirements (GMER) illustrates the importance of this domain for doctors all over the world.⁸

However, the situation is different in many developing countries where professionalism is neither taught nor assessed at undergraduate or postgraduate levels.⁹

.....
¹⁻⁴Department of Medical Education, Postgraduate Medical Institute, Hayatabad, Peshawar, Pakistan, ^{5,6}Department of Medical Education, University of Illinois at Chicago, USA.

Correspondence: Abid Jameel. Email: ajameel99@yahoo.com

Reports of unprofessional behaviours exhibited by physicians have emerged from many developing countries.^{10,11} A study conducted among medical students at one of the top medical institutions in Pakistan concluded that the level and understanding of professional values were far below the acceptable level.¹² However, it is encouraging to note that the response of students to training in professionalism was positive.^{13,14}

In order to address the issue of training residents in professionalism in resource-poor countries, Postgraduate Medical Institute (PGMI), Peshawar, in collaboration with the Foundation for Advancement of International Medical Education and Research (FAIMER), Philadelphia, initiated a project to examine the feasibility, effectiveness and short-term impact of teaching and assessment of professionalism.

The current study was planned to examine the internal structure of a professionalism assessment administered to residents and whether this short-term training led to any significant improvement in professionalism scores measured using a written examination and Objective Structured Clinical Examination (OSCE). PGMI is responsible for training of all postgraduate residents across 27 different specialties in all public-sector teaching hospitals of Khyber Pakhtunkhwa (KP) province in northern Pakistan.

Subjects and Methods

The pre-intervention and post-intervention study was conducted from January to August 2012 and comprised 7 workshops (intervention) of three days each that were held at four teaching hospitals of the KP province in Pakistan. The study was approved by the institutional research board of PGMI. The PGMI-affiliated hospitals were Lady Reading Hospital, Khyber Teaching Hospital and Hayatabad Medical Complex in Peshawar and Ayub Teaching Hospital in Abbottabad. Convenience sampling was used to pick up postgraduate Residents as the study sample from various Medicine and Surgical disciplines (Table-1). A pilot study was conducted on 20 final year medical students of Khyber Girls Medical College, Peshawar, in January 2012. Two workshops each at the four teaching hospitals were held enrolling 20 first and second year Residents. Teaching and assessment strategies adopted in the workshop included pre-test and post-test assessment in the form of objective structured clinical examination (OSCE) with 10 stations with standardised patient (SP) encounters and 10 written scenarios requiring short answers (Table-2). Pre-test for residents on day one was followed by group discussion sessions on each station using moral reasoning for

justification of responses on day two and part of day three. After each case discussion, there was a brief faculty-led presentation on the topic. This interactive teaching session was followed by a post-test at the end of day three using the same stations, SPs and marking system as the pre-test.

OSCE scoring was done by the SPs on a scoring sheet with a checklist (Table-2) while written test was scored using a scoring key for the test questions. SPs were recruited from the hospital staff, including nurses for female SPs. They were trained in portraying their specific role over a period of two weeks. Checklists for OSCE stations were translated into Urdu language so that SPs could read and understand them since some of the SPs were not familiar with English language. Instructions for Residents were in English, but the language used during SP interaction was Urdu. Prompts were provided to SPs to facilitate Residents as required. SPs were also trained in scoring the checklist for the OSCE stations as set out in literature.^{15,16} Written case scenarios were taken mainly from the professionalism vignettes in the document "Project Professionalism" created by the American Board of Internal Medicine (ABIM). Differences in pre- and post-test scores were noted for all stations and all workshops to analyse improvement in Residents' knowledge of core principles of professionalism and their application. Workshop evaluation was carried out at the end of each workshop.

The internal structure of the professionalism assessment was examined using pre-test data collected from the 10 OSCE stations and 10 written scenarios. Item-test correlations were used to measure the association between a station to the total test score; mean inter-item correlation was used to examine associations between scores from each station. Internal consistency reliability was measured using Cronbach's alpha. To examine the reliability of the combined written test and the OSCE, a generalisability study was conducted using the p Xi: d design where examinees (p) were crossed by items nested in the two domains of assessment formats. Written test and OSCE items (i) were specified as random and the assessment format (d) was specified as fixed. Differences between pre- and post-test scores were calculated by paired t-test analysis. Analysis of covariance was performed to examine the change in pre- and post-test scores after controlling for OSCE and the written examination. Stata 12 was used for all statistical analyses.

Results

There were 136 postgraduate residents in the study.

Internal consistency of test scores across the 10 OSCE

Table-1: Specialties of training of participating Residents.

Medical allied specialties: General medicine, pulmonology, gastroenterology, cardiology, neurology, dermatology, haematology
Surgical allied specialties: Neurosurgery, orthopaedic surgery, plastic surgery, paediatric surgery, ophthalmology Gynaecology and obstetrics Radiology

stations was 0.61 and across 10 written scenarios was 0.32 (using Cronbach's alpha). The generalisability coefficient (G-Coefficient) of the combined OSCE and the written examination was 0.65 (Phi-coefficient: 0.56). Decision study (D-study) results indicated a total of 13 stations to

be required to reach a G-Coefficient of 0.71.

Item-test correlation measuring the relative discrimination strength of each OSCE station ranged between 0.20 and 0.67. Average inter-item correlation, correlating one station score to other station scores, ranged between 0.11 and 0.15 for the OSCE stations. Item-test correlation measuring the relative discrimination strength of each written scenario ranged between 0.25 and 0.54. Average inter-item correlation, correlating one station score to other station scores, ranged between 0.03 and 0.06 for the written scenarios.

The test-retest reliability for OSCE stations was 0.40 ($p < 0.001$) while for written examination it was 0.36

Table-2: Professionalism related scenarios used in workshops.

OSCE Station Scenarios	Check-list items	Written Case Scenarios
Dealing with an angry patient	8	Religious beliefs and patient care
Patient autonomy vs. non-maleficence	9	Altruism
Patient autonomy and shared decision making	10	Relationship with pharmaceutical company
Do not resuscitate orders	9	Dealing with depressed patient (autonomy vs. beneficence)
Confidentiality	8	Covering up for negligence
HIV related confidentiality	10	Sexual harassment
Breaking bad news	10	Patient data falsification/ fabrication
Request for a fake certificate by a family friend	9	Misuse of authority/position
Battered woman	11	Research trial payment (conflict of interest)
Sexual impropriety	10	Conflict of interest

HIV: Human immunodeficiency virus.

Table-3: Descriptive statistics of results for individual workshops (pre- and post-test).

Workshop	Format	n	Difference*	Pre-test				Post-test			
				Mean	SD	Min	Max	Mean	SD	Min	Max
1	OSCE	18	42.45	36.11	9.37	23	56	78.56	6.94	56	87
	Written	18	34.06	35.72	11.32	19	62	69.78	12.52	33	86
2	OSCE	23	36.52	34.13	6.13	22	51	70.65	5.51	60	80
	Written	23	44.39	39.52	13.35	9	62	83.91	11.77	62	100
3	OSCE	19	31.05	40.63	7.92	26	57	71.68	8.31	53	85
	Written	19	40.57	35.11	12.78	10	52	75.68	12.89	43	95
4	OSCE	19	32.89	43.11	9.50	31	64	76.00	5.57	61	85
	Written	19	34.16	39.63	15.54	9	71	73.79	12.96	48	95
5	OSCE	21	33.48	42.95	10.23	30	65	76.43	5.76	66	84
	Written	21	27.29	39.57	14.21	14	62	66.86	16.25	28	95
6	OSCE	20	39.10	39.95	10.68	28	74	79.05	5.74	69	88
	Written	20	48.40	34.90	14.21	10	71	83.30	10.98	67	100
7	OSCE	16	38.00	41.25	8.67	28	57	79.25	5.05	71	86
	Written	16	38.94	42.81	8.38	33	57	81.75	14.50	47	95
Combined	OSCE	136	36.12	39.61	9.39	22	74	75.73	6.91	53	88
	Written	136	38.37	38.13	13.13	9	71	76.50	14.40	28	100

Note: *All scores improved significantly, $p < 0.001$, based on paired t-test.

OSCE: Objective Structured Clinical Examination

SD: Standard Deviation.

($p < 0.001$). Correlation between OSCE and written exam for pre-test was 0.42 ($p < 0.001$) while for post-test the correlation was 0.17 ($p = 0.046$). Correlation between pre-test written and post-test OSCE was 0.23 ($p = 0.001$), however, the correlation between pre-test OSCE and post-test written was not significant.

Mean pre-test score for the written component of all seven workshops combined was $38.13 \pm 13.13\%$ vs. $76.50 \pm 14.4\%$ for the post-test score ($p < 0.001$). Mean pre-test score for the active OSCE stations was $39.61 \pm 9.39\%$ vs. $75.73 \pm 6.91\%$ for the post-test ($p < 0.001$). Mean pre-test scores for OSCE stations and the written scenarios combined in seven workshops with 136 residents was $38.9 \pm 6.91\%$ while mean post-test scores was $76.3 \pm 14.4\%$ ($p < 0.001$). Variability in scores was smaller for OSCE than for written test for the post-test results (Table-3). There was a significant improvement in OSCE performance even after controlling for the pre-test effect of the written exam (R-squared: 0.17; $F(2, 133) = 13.39$; $p < 0.001$). Similarly, there was also significant improvement in written exam performance, even after controlling for the pre-test effect of OSCE exam (R-squared: 0.13; $F(2, 133) = 10.03$; $p < 0.001$). The standardised effect size for the adjusted regression was 0.37 for both comparisons ($p < 0.001$). Workshop evaluation by Residents using Likert scale showed a mean rating of 9 on a 10-point scale.

Discussion

The issue of suboptimal levels of professional values due to lack of teaching of professionalism to medical graduates in Pakistan has recently been highlighted in a study.¹² It has rightly stressed upon designing strategies to teach and assess professionalism to medical students in order to improve the professional standing and professional values of our medical graduates. Formal structured training programmes in teaching hospitals in Pakistan have been shown to improve residents' abilities in domains such as communication skills.¹⁷ Various methodologies used to teach and assess professionalism have been reported in literature.^{18,19} Commonly used methods for assessing professionalism include critical incident report, portfolio assessment, professional mini-evaluation exercise (P-MEX), OSCE and multi-source feedback.^{20,21} A survey on students regarding their perceptions about the various aspects of teaching of professionalism concluded that role modelling and case scenario sessions on professionalism were recommended strongly by the students.²² OSCE has also been shown to be a good teaching tool for multiple and complex behaviours where students are exposed to close to real life situations and also receive immediate feedback from the SPs or faculty members after the OSCE.^{15,23} OSCEs

have been used mainly to test the ethical domain of medical students and Residents using standardised patients and multiple examiners. A study reported an inter-rater reliability between 0.67-0.79 and an internal consistency reliability of 0.46 for a six-station OSCE.^{15,16} A different study reported inter-rater reliability ranging from 0.80 to 0.99 and a low correlation of 0.20 across stations.²⁴ Using OSCE stations alone are not sufficiently reliable to measure complex constructs such as professionalism. It has been proposed that using multiple methods of assessment (triangulation) and multiple evaluators will improve the reliability as well as validity of the assessment.²⁵

We therefore used multiple methods for teaching and assessing professionalism to our residents starting with making them go through written case scenarios and multiple OSCE stations with SPs portraying different professionalism-related scenarios. Once they were exposed to real life scenarios they were more inquisitive about the problem faced by them. We followed this by group discussion regarding each case scenario using moral reasoning. This was followed by a facilitator-led interactive lecture on the topic. This strategy resulted in an excellent outcome in the form of significant increase in knowledge and behaviour in post-test and also manifested in high workshop evaluation for its methodology. During discussion on various topics of professionalism, the specific keys or checklists for the stations were not discussed in order to minimise memorisation effect that could confound the results on the post-test.

We used a written test based on case scenarios to measure the knowledge domain of our Residents, and OSCE stations to measure the behavioural aspect of professionalism, thereby covering both knowledge and skills of residents in professionalism. Internal consistency reliability of our OSCE stations was 0.61 while the internal consistency reliability of the written test was 0.32 which was very low. The G-coefficient combining OSCE and written stations was 0.65. Thirteen stations each were required to reach a G-coefficient of 0.71. The low internal consistency reliability of the written exam compared to OSCE may be due to the complexity involved in assessing professionalism and indicates that OSCE has a much higher reliability than written exam.

An important aspect of this study is that the psychometric properties of the OSCE stations were not affected even after translation to another language. This can have positive implications for translating these scenarios to other languages and making subtle changes to reflect the

local culture. Another aspect that came out of this study was that the effect of workshop on performance of OSCE and the written test were similar. This may indicate that the focus of workshop was similarly distributed across knowledge and performance.

During the pre-test and the interaction with trainees in discussion sessions, deficiencies in the knowledge and skills of Residents in professionalism became evident. Residents could not differentiate between "do not resuscitate" orders and "euthanasia". Conflict of interest was not understood by most of the residents. Matters related to gender bias, harassment, autonomy of patients and informed consent were also not very clear. Some of these areas of deficiency identified in our study were also reported in an earlier study at Karachi.¹² After clarification of concepts in relationship to our cultural background, the performance of Residents increased sharply in the post-test. Comments of Residents were very encouraging and almost all Residents called for making the attendance at such workshops as mandatory. The deficiencies in knowledge and skills of Residents noted in this project will help us design future workshops and its content.

Regarding workshop logistics and SP training, recruiting female SPs was a particular problem in our part of the world where females avoid contact with males. We overcame this problem by recruiting our nurses as SPs on voluntary basis. Male staff members were recruited and trained as male SPs. Their problem of reading or understanding English was overcome by translating the scenarios, prompts and checklists into Urdu language which facilitated comprehension of scenarios and accurate marking of checklists. This did not affect the psychometric properties of the OSCE, as discussed above. Arrangements for OSCE stations were made by erecting patient screens in a big hall that was converted into OSCE hall with 10 OSCE stations. These simple measures made holding of these workshops for teaching and assessment of professionalism practical and feasible. Based on comments of Residents and request of faculty members, we have currently started holding two-day workshops on professionalism for our faculty members in a format that is mainly based on video triggers and small group discussions. After the completion of these workshops for faculty we will have enough faculty members to train Residents in their own clinical units. This will ensure the long-term sustainability of the project.

Our study has a number of limitations that include a limited number of Residents, assessing Residents from only one province and measuring only the short-term impact of professionalism training in Residents. However,

considering the fact that training structure is almost similar in all public-sector hospitals of the country, it may be reasonable to generalise the results to all public-sector hospitals of the country. It will be interesting to measure the long-term impact of such training in Residents in future. Since the concept of professionalism varies across different cultures, therefore, the scenarios used for teaching and assessing professionalism may have to be changed according to individual, institutional and cultural requirements.

Conclusion

Professionalism in Medicine is a key competence in postgraduate medical education that has to be mastered and practised by doctors throughout their lives. It is therefore imperative that health professional institutions around the world take steps for longitudinal training and continuous assessment of their students and Residents in professionalism. The study has shown that professionalism can be taught as well as assessed in resource-poor settings and can have a very positive impact on doctors.

Acknowledgements

We are grateful to PGMI, Peshawar, and FAIMER, Philadelphia, for the projector support, to Mia Solomon for editorial comments, and the participating Residents and SPs.

References

1. Cruess SR. Professionalism and medicine's social contract with the society. *Clin Orthop Relat Res* 2006; 449: 170-7.
2. Baldwin DC Jr, Daugherty SR, Rowley BD. Unethical and unprofessional conduct observed by residents during their first year of training. *Acad Med* 1998; 73: 1195-200.
3. Hicks LK, Lin Y, Robertson DW, Robinson DL, Woodrow SI. Understanding the clinical dilemmas that shape medical students' ethical development: questionnaire survey and focus group study. *Br Med J* 2001; 322: 709-10.
4. ABIM Foundation. Medical Professionalism in the New Millennium: A Physician Charter. *Ann Intern Med* 2002; 136: 243-6.
5. Royal College of Physicians. Doctors in Society: Medical Professionalism in a changing world. Report of a working party of the Royal College of Physicians of London. London RCP, 2005.
6. Swing SR. The ACGME outcome project: retrospective and prospective. *Med Teach* 2007; 29: 648-54.
7. Frank JR, Danoff D. The CanMEDS initiative: Implementing an outcome-based framework of physician competencies. *Med Teach* 2007; 29: 642-7.
8. Schwarz MR, Wojtczak A. Global minimum essential requirements: A road towards competence-oriented medical education. *Med Teach* 2002; 24: 125-9.
9. Minai FN. Professionalism in residency training. *J Coll Physicians Surg Pakistan* 2009; 19: 675-7.
10. Gadit AAM. Corruption in medical practice: How far have we gone? *J Pak Med Assoc* 2011; 61: 93-4.
11. Hui EC. The contemporary health care crisis in China and the role of medical professionalism. *J Med Philos* 2010; 35: 477-2.
12. Sobani ZA, Mohyuddin MM, Farooq F, Qaiser KN, Gani F, Bham NS

- et al. Professionalism in medical students at a private medical college in Karachi, Pakistan. *J Pak Med Assoc* 2013; 63: 935-9.
13. Adkoli BV, Al-Umran KU, Al-Shiekh M, Deepak KK, Al-Rubaish AM. Medical students' perception of professionalism: A qualitative study from Saudi Arabia. *Med Teach* 2011; 33: 840-5.
 14. Waheed G, Mengal MA, Shah SSA, Shiekh AW. Faculty and students' perception about aptitude of professionalism in admissions process of medical college. *J Ayub Med Coll Abbottabad* 2011; 23: 166-9.
 15. Singer PA, Cohen R, Robb A, Rothman A. The ethics objective structured clinical examination. *J Gen Intern Med* 1993; 8: 23-8.
 16. Singer PA, Robb A, Cohen R, Norman G, Turnbull J. Evaluation of a multi center ethics Objective Structured Clinical Examination. *J Gen Intern Med* 1994; 9: 690-2.
 17. Jameel A, Noor SM, Ayub SA. Survey on perceptions and skills among postgraduate residents regarding breaking bad news at teaching hospitals in Peshawar, Pakistan. *J Pak Med Assoc* 2012; 62: 585-9.
 18. Stern DT, Papadakis M. The Developing Physician - Becoming a Professional. *N Engl J Med* 2006; 355: 1794-9.
 19. Kirk LM. Professionalism in medicine: definitions and considerations for teaching. *Proc (Baylor Univ Med Cent)* 2007; 20: 13-6.
 20. Arnold L. Assessing Professional Behavior: Yesterday, Today and Tomorrow. *Acad Med* 2002; 77: 502-15.
 21. Cruess R, McIlroy JH, Cruess S, Ginsburg S, Steinert Y. The Professionalism Mini-evaluation Exercise: a preliminary investigation. *Acad Med* 2006; 81: 574-8.
 22. Byszewski A, Hendelman W, McGuinty C, Moineau G. Wanted: role models - medical students' perceptions of professionalism. *BMC Med Educ* 2012; 12: 115.
 23. Altshuler L, Kachur E. A culture OSCE: Teaching residents to bridge different worlds. *Acad Med* 2001; 76: 514.
 24. Roberts J, Norman G. Reliability and learning from the objective structured clinical examination. *Med Educ* 1990; 24: 219-23.
 25. Van derVleuten CP, Schuwirth LW. Assessing professional competence: from methods to programs. *Med Educ* 2005; 39: 309-17.
-