Training and validation of standardized patients for evaluation of general practitioners' performance in management of obesity and overweight

Tahereh Changiz, Sepideh Jamshidian, Mohamad Hassan Entezari[1], Nazila Kassaian[2]

Department of Medical Education, Medical Education Research Center, [1]Department of Clinical Nutrition, School of Nutrition and Food Sciences, Food Security Research Center, Isfahan University of Medical Sciences, Isfahan, [2]Department of Medical Education, School of Medical Education, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Abstract

Background: Standardized patient (SP) can serve as a valuable tool to measure the physician performance in actual clinical settings, but it has not been validated for obesity/overweight disorders. This study has been conducted to describe the process of creating reliable and valid SPs for evaluation of general-practitioners' management of obesity/overweight in Iran.

Materials and Methods: A total of 6 obese/overweight volunteers (potential SPs) took part in training. Three scenarios, along with corresponding checklists including 102 items representing different aspects of diagnosis and treatment of obesity/overweight, were developed by an expert group. The SPs were trained using role playing method. During this part, one of the SPs failed. The SPs' portrayal of their respective scenario was online watched in another room and the checklist filled independently by the physician, research assistant and other SPs. The reliability of the checklist to be used by the SPs was assessed by Cronbach's alpha. The overall inter-rater agreement was calculated by the intraclass correlation coefficient statistic for total scores.

Results: The 5 eligible SPs were all women between 20 years and 39 years of age. Inter-rater agreement between the SPs' total scores was 0.899, value (95% confidence intervals) were 11.8 (0.68-0.98) and *P* value was <0.001. The Cronbach's alpha for reliability of completed checklists was 0.91.

Conclusion: SP could be a powerful instrument for evaluating medical performance of general practitioners in the field of obesity/overweight management. Further research is needed to find the more aspects of training and validation of unannounced SPs in this field.

Key Words: Obesity, performance, standardized patient

Access this article online	
Quick Response Code:	**Website:** www.advbiores.net
	DOI: 10.4103/2277-9175.125931

INTRODUCTION

Obesity is a global problem, affecting an estimated 300 million people worldwide and is associated with higher mortality and increased risk of co-morbidities.[1] In 2003, the prevalence of obesity and overweight in Iran was reported as more than 50% as judged by a body

mass index (BMI) of 25 or more while it was indicated an increase in the trend of them.[2] It is documented that modest weight loss as little as 5-10% by diet and exercise can improve health outcomes and health related risk factors.[1]

The primary care physician is in a unique position of influence, which is associated with improved diet and physical activity, readiness for life-style change and weight loss. As the gatekeepers to the health-care system, general practitioners (GPs) can play a vital role in addressing obesity in the consultation.[3] Understanding of patients' weight management experiences in the physician-patient setting will be helpful to gather data on healthy and unhealthy weight management practices or behaviors since they are closely linked to obesity, diet and physical activity.[4]

One of the most important challenges for health-care researchers is measuring physicians' performance in the context of clinical care. Standardized patients (SPs) can serve as a measurement tool to access the information that cannot be obtained from other sources such as the performance of the physical examination and counseling activities.[5] A SP is defined as an individual who has been well-trained to imitate as a patient in a realistic way. SP as an excellent tool has been used in different domains of medical education for assessment of undergraduates' and postgraduates' skills.[6]

Using SP for access to information, controls many biases related to observation, such as memory and social desirability biases, which can be never completely avoided with surveys, clinical vignettes and direct observation.[5]

These advantages are realized only if the physicians are unaware that they are being observed and if the SPs be standardized as storytellers and as performance coders.[6]

Unfortunately, knowledge about SPs as a source of data in health care research is limited and few investigators used SPs to provide evidence that these conditions have been met.[7]

In Iran, SP visits have been used in a limited number of studies to assess health professionals' performance in real practice in the field of psychology.[8] However, there is not any study in Iran and probably the world that evaluates the SP technique for assessment of physicians' performance in the field of obesity/overweight management. Furthermore, there is no published study regarding the physicians'

performance for obese and overweight patients in Iran.

In this study, we assessed the accuracy (validity) and consistency (reliability) of the SPs for the evaluation of physicians' performance in the field of weight management. We report here the procedures and results of training and validation of SPs' portrayal of obesity or overweight disorders that can be used by other researchers.

MATERIALS AND METHODS

The study was approved by the Ethics Committee at the Research Council of Isfahan University of Medical Sciences.

Selection of SPs
A total of 10 obese or overweight individuals (nine women and one man) with BMI over 25 were recruited from volunteer associations and Clinical Skills Learning Center of Isfahan University of Medical Sciences. They were interviewed to confirm their suitability to participate as SP and assessed with certain criteria related to well-being, availability and understanding of the study objectives. Six individuals were found eligible to participate in the training after which they signed a written informed consent form.

They were between 20 years and 43 years of age; one of them was a man and all had some knowledge of overweight disorders.

Scenario development
Three different scenarios; a scenario for man, a scenario for married woman and a scenario for single woman were compiled by consensus among the panel of experts. In all three scenarios, SP's chief complaint was obesity, seeking GP's help for weight loss. The key points of standardized information were life-style behaviors, the prior treatment, accompanied diseases and family support.

The scenarios were based on issues relating to the diagnosis and treatment of obesity/overweight disorders, which had been emphasized in a guidebook for GPs.[5] The SPs kept their own names and birth date records, but marital status and employment was standardized. Each scenario included 15 key points of standardized information to be addressed in the visit.

Checklists for evaluation of GPs' performance
To evaluate the performance of the GPs regarding to overweight or obesity management, aspects and guidelines[9-11] were identified and three checklists were compiled by experts in a focus group; consisted

of two expert GPs, medical education expert, research assistant and two nutrition experts. The checklists were related to important issues to be considered by the GPs when encountering a case of obese/overweight. Each checklist included 6 parts with binary questions (yes/no): History taking (57 items), physical exams (9 items), explanation (15 items), intervention (9 items), prescriptions (3 items) and referral to a specialist (5 items). The checklist also had 4 open-ended questions including the time that the GP spent the time s/he suggested for the next visit, the referral address for more information, and the SP's satisfaction rate.

To score the checklist, the weight of each item (from 1 to 3) was defined by the expert panel, according to the importance of each item in the management of obese/overweight client. Minimum and maximum attainable total score for checklist were 65 and 237, respectively. Content validity of the checklists was assessed by the panel of experts including two expert GPs, medical education expert, research assistant and two nutrition experts.

SPs' training
The SPs were given a detailed description of the scenario, consisting of information that should be shared with the physician and information that has to be provided only if the physician asked questions. During training, the SPs engaged in role-playing based on their scenarios with a physician on the research team. The tape recording was also obtained while they were practicing role-plays. This part of the training was spread over two meetings and lasted 8 h.

Next, the SPs were instructed how to fill in the checklists and record the practice of the physicians. During this process, the ambiguities were clarified. In this part which lasted 8 h, the male SP was failed and excluded.

Validation of SPs
Each of the 5 SPs was portrayed her scenario with a GP on the research team. All interactions were audio recorded. The other SPs and the research assistant watched them through live video in the other room and afterward all the trainees and trainers completed the observational rating scale individually.

The research assistant and other SPs were used a grid identical to the one used by the SP. The audio recordings and checklists of each SP performance were analyzed by the research assistant who rated the SP's accuracy of portrayal of the scenario and coding the physician's performance. Finally, each SP made a blind visit to another physician who had agreed to participate in the study in the real setting

and then the notes about SPs performance made by the physician were analyzed by the research assistant and the eligible SPs who had reached 90% accuracy in portrayal and coding were selected.

Analysis
Accuracy of completion of the checklist for physician performance was evaluated that could be validated by audio tape recordings. The overall inter-rater agreement was calculated by the intraclass correlation coefficient statistic for the total score. Cronbach's alpha was measured for checklist reliability.

Statistical analysis was performed with SPSS-16 for Windows (Inc., Chicago, IL) and $P \leq 0.05$ was considered statistically significant.

RESULTS

After training procedures, 5 SPs were selected to participate in the validation phase of the study. They were all women between 20 years and 39 years of age. The characteristics of the eligible SPs have been shown in Table 1.

Overall Inter-rater agreement was 0.899, F value (95% confidence intervals) were 11.8 (0.68-0.98) and P value was < 0.001 for the total score of the checklist.

The Cronbach's alpha for reliability of completed checklists was 0.91.

The individual test results for the 5 participants, the research assistant and the physician are shown in Table 2.

During the blind visit to the physician in the real setting, one of the SPs was excluded because she couldn't get the eligible score.

DISCUSSION

In medical education, there is a gap between what physicians essential performance based on the practice guidelines and what they actually do. Bridging this

Table 1: Demographic and anthropometric characteristics of the eligible standard patients

SP code	Age (years)	Waist circumference (cm)	Height (cm)	Weight (kg)	BMI (kg/m²)
1	20	87	176	83	26.8
2	28	86	162.5	75.5	28.6
3	39	91	151	62	27.2
4	21	88	164.5	72	26.6
5	31	100	159	92	36.4

SP = Standardized patient; BMI = Body mass index

Table 2: Scores given to the GP's performance in 5 SP-physician encounters by trained SPs as well as a research assistant and physician herself

Raters	Encounter 1	Encounter 2	Encounter 3	Encounter 4	Encounter 5
Physician	30.4	36.9	42.1	21	33.6
Research assistant	34.1	36.4	47.2	37.4	29.9
SP 1	37.9	34.1	56.1	39.7	24.8
SP 2	32.7	37.9	58.4	40.7	36.4
SP 3	26.2	29	51.9	36	22
SP 4	29.9	29.4	40.2	37.4	28
SP 5	34.6	45.3	41.6	37.4	36.4

SP = Standardized patient; GP's = General practitioners

gap, different methods help to assess how learning is transferred to the professional practice.

Use of the SP is a valuable tool for performance assessments in medical education and research. Two main advantages of SP using are: (1) The SP as a standardized stimulus can control case-mix variability, and (2) SPs can themselves provide access to the main information on the subjects' performance such as the case histories, physical examination and counseling data. Moreover, using SP controls for biases related to the situation of being observed.[12]

The high expense of using SPs in comparison with other assessment methods is the drawback of unannounced SPs. However, SPs are considered as a valid tool for evaluating the real clinical practice.[13]

In this study, we evaluated the aspects of SP-based research to estimate physician's performance in the field of weight management of obese/overweight and reported our experience, the reliability of coding and portrayal aspects.

There are few medical investigators who have used the SP technique and this study may be a good experience that could be used by other researchers planning to use the SP method. To the best of our knowledge, this is the 1st study in Iran and probably the world that evaluates the SP technique for assessment of GPs' performance in the field of obesity management.

In available studies, to document and assess GPs practices regarding the prevention and management of overweight and obesity, questionnaires were completed by GPs or patients in the waiting room or through the telephone survey interview or postal surveys.[14-17] In respect of the method, the results of our study were better than those previously reported in this field.

It is known that the use of SPs is certainly the best way to evaluate counseling aspect of medical performance.

However, evaluation with SPs is more complicated in practice-based settings and must therefore be applied with the greatest possible rigor.[8]

It have been surmised that inaccuracies of recording by unannounced SPs might be a bias in studies on provider performance.[18]

The high correlation coefficient between raters in our study confirms that careful and systematic training of SPs can be protective against such bias, which is in concordance with studies in other areas of medicine.[8]

Our experience in this study showed that SPs can be trained to provide a reliable and valid assessment of GPs' performance with regarding diagnosis and treatment of obesity/overweight disorders.

In the investigations that measure various characteristics by having rater assigned scores to observed people, or events, it is desirable to measure the extent of agreement when rating the same set of things. This can be treated as a sort of reliability statistic for the measurement procedure. In this study, we used intraclass correlation coefficient for reliability analysis to evaluate the agreement between the raters. This technique can be computed when there are more than 2 judges for purposes of estimating inter-rater reliability. Though intraclass correlation coefficient has applications in multiple contexts, their implementation in reliability is oriented toward the estimation of interrater reliability.[19]

Limitation

Our findings regarding the SPs cannot be generalized. However, what can be transferred to other studies, with similar objectives of evaluating performance, is how we developed the processes of validation.

CONCLUSION

In conclusion, the SP could be a powerful instrument for evaluating medical performance of GPs in the field of obesity/overweight management. Further research is needed to find the more aspects of training and validation of unannounced SPs in this field.

REFERENCES

1. Mokdad AH, Ford ES, Bowman BA, Dietz WH, Vinicor F, Bales VS, *et al*. Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001. JAMA 2003;289:76-9.
2. Akhavan A, Kelishadi R, Sadri GH, Sabet B, Toloui R, Baghai AH. Prevalence of obesity in central part of Iran: Healthy heart project. J Gazvin Univ Med Sci 2003;26:27-35.
3. Martin PD, Dutton GR, Rhode PC, Horswell RL, Ryan DH, Brantley PJ. Weight loss maintenance following a primary care intervention for low-income minority women. Obesity (Silver Spring) 2008;16:2462-7.

4. Potter MB, Vu JD, Croughan-Minihane M. Weight management: What patients want from their primary care physicians. J Fam Pract 2001;50:513-8.
5. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: A prospective validation study of 3 methods for measuring quality. JAMA 2000;283: 1715-22.
6. Fihn SD. The quest to quantify quality. JAMA 2000;283: 1740-2.
7. Tamblyn RM. Use of standardized patients in the assessment of medical practice. Can Med Assoc J 1998; 158:205-7.
8. Shirazi M, Sadeghi M, Emami A, Kashani AS, Parikh S, Alaeddini F, *et al*. Training and validation of standardized patients for unannounced assessment of physicians' management of depression. Acad Psychiatry 2011;35:382-7.
9. National Institutes of Health. Clinical Guidelines on the Identification Evaluation and Treatment of Overweight and Obesity in Adults. New York (USA): NIH Publication; 1998.
10. Scottish Intercollegiate Guidelines Network. Management of Obesity: A National Clinical Guideline, 2010. Available from: http://www.sign. ac.uk./guidelines/index.html [Last accessed on 2012 Feb 25].
11. Health Team Works. Adults Obesity Guideline: Weight Assessment and Management Algorithm, 2011. Available from: http://www. healthteamworks.org/guidelines/obesity.html [Last accessed on 2012 Feb 24].
12. Rethans JJ, Gorter S, Bokken L, Morrison L. Unannounced standardised patients in real practice: A systematic literature review. Med Educ 2007;41:537-49.
13. Glassman PA, Luck J, O'Gara EM, Peabody JW. Using standardized patients to measure quality: Evidence from the literature and a prospective study. Jt Comm J Qual Improv 2000;26:644-53.
14. Campbell K, Engel H, Timperio A, Cooper C, Crawford D. Obesity management: Australian general practitioners' attitudes and practices. Obes Res 2000;8:459-66.
15. Bocquier A, Verger P, Basdevant A, Andreotti G, Baretge J, Villani P, *et al*. Overweight and obesity: Knowledge, attitudes, and practices of general practitioners in France. Obes Res 2005; 13:787-95.
16. Klumbiene J, Petkeviciene J, Vaisvalavicius V, Miseviciene I. Advising overweight persons about diet and physical activity in primary health care: Lithuanian health behaviour monitoring study. BMC Public Health 2006;6:30.
17. Thuan JF, Avignon A. Obesity management: Attitudes and practices of French general practitioners in a region of France. Int J Obes (Lond) 2005;29: 1100-6.
18. Tamblyn RM, Abrahamowicz M, Berkson L, Dauphinee WD, Gayton DC, Grad RM, *et al*. First-visit bias in the measurement of clinical competence with standardized patients. Acad Med 1992;67 (Suppl 10):S22-4.
19. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. Psychol Methods 1996;1:30-46.