Research Article

Performance Prediction of the Family Medicine Applicants of Oman Medical Specialty Board in Multiple Mini-interviews Using the International Foundations of Medicine Exam and Situational Judgment Test

Rahma Said Al Hadhrami [1, *], Najlaa Jaafar [2], Badriya Al Farsi [3] and Asma Al Belushi [4]

[1]Consultant Family Medicine, Department of Family Medicine and Public Health, Sultan Qaboos University Hospital, Muscat, Oman
[2]Senior Consultant Family Medicine, Primary Health Care, Ministry of Health, Muscat, Oman
[3]Senior Consultant Family Medicine, Armed Forced Hospital, Muscat, Oman
[4]Consultant Emergency Medicine, Director of Trainee Affairs Department, Oman Medical Specialty Board, Muscat, Oman

Corresponding author: Department of Family Medicine and Public Health, Sultan Qaboos University Hospital, Muscat, Oman. Email: rahmasaid@squ.edu.om

Abstract

Background: Entry into postgraduate medical training programs should be based on valid selection measures predictive of training performance. A situational judgment test (SJT) is often used during the residency selection process to measure knowledge and behavioral attributes.

Objectives: This study aimed to test the validity of the International Foundations of Medicine (IFOM) examination and an SJT for predicting performance during multiple mini-interviews (MMI) of the applicants for a family medicine residency program in Oman.

Methods: This retrospective cross-sectional study was conducted at Oman Medical Specialty Board (OMSB) during the application process of residency for the academic year 2021. Data were collected from 80 applicants to the OMSB Family Medicine Residency Program. Before participating in the MMI panels, all applicants underwent pre-interview tests (comprising the IFOM examination and SJT). Multiple regression analysis was performed to determine the predictors of the MMI score.

Results: A total of 67 applicants completed the entire selection process cycle and were included in the study. The SJT score was a significant predictor of the MMI score (r = 0.367; P = 0.002), whereas the IFOM score was not a significant predictor (r = 0.105; P = 0.396).

Conclusions: This study supports using SJT as an additional selection measure alongside the MMI panel for entry to OMSB residency training.

Keywords: Graduate Medical Education, Internship and Residency, Selection Process, International Foundations of Medical Examination, Situational Judgment Test, Multiple Mini-interviews, Oman

1. Background

Using valid and reliable methods in selecting candidates for medical specialty training should be of high priority to any academic institute (1). A concise analysis of what is needed for the job and its purpose should determine the selection method and the candidates most likely to perform well during training. Therefore, combined with academic and clinical competencies, an assessment of the applicant's personal attributes is essential to the selection process (2). Moreover, to avoid any selection bias, the chosen selection method should be tailored and sensitive to diversity issues (1, 2).

At Oman Medical Specialty Board (OMSB), the regula-

tory body responsible for overseeing postgraduate medical education in Oman, all residency training applicants are required to pass the International Foundations of Medicine (IFOM) examination. Regardless of their chosen specialty, applicants who have passed the IFOM examination subsequently undergo three panels of multiple mini-interviews (MMIs) over four days from 8:00 am to 5:00 pm, which is highly connected with interviewers' motivation to improve their profession and future patient care (3). Approximately 80 - 100 candidates apply yearly to the OMSB Family Medicine Residency Program as their first or second choice, representing more applicants than other programs. For the 2021 academic year, the OMSB Education

Committee piloted an additional selection measure (a situational judgment test (SJT)) to assess attributes other than knowledge alone.

Tests assessing situational judgment seek to evaluate the applicant's responses to professional dilemmas likely to be encountered in clinical practice (4-6). Worldwide, SJTs are widely used in the selection of high-stakes industry personnel, as well as medical students in the UK, Belgium, Canada, Australia, and Latin America. In the UK, 50% of the total scores of final-year medical students come from the SJT (6-11). In addition, SJTs were useful for teaching empathy, assertiveness, inter-professional communication, and ethics to pharmacy students (12).

During SJTs, candidates are required to respond to questions assessing both behavioral tendency (i.e., how to behave in a given specific situation) and knowledge tendency (i.e., suitable and effective responses to a given situation). According to McDaniel et al., the portion of SJTs testing knowledge correlates with general cognitive ability, while the section testing behavior correlates with positive or desirable personality attributes (13). Furthermore, Lievens et al. found that SJT and cognitive tests completed each other in selecting doctors (14). Ahmed et al. revealed that SJT overcomes clinical problem-solving in predicting selection center performance (R^2 = 0.51 versus 0.35) (15). However, combining the two tests maximized the predictive ability of performance, explaining 57% of the variance in overall mean scores (15). In addition, Husbands et al. found the reliability of one particular SJT to be high during the undergraduate selection process (Cronbach's alpha coefficient = 0.64), with significant associations observed between SJT scores, integrity, other desirable personality traits, and MMI scores (16).

Recently, SJTs have been incorporated into the postgraduate selection process for the UK Foundation Programme (2, 10, 17). However, the evidence supporting the predictive validity of SJTs is lacking in the postgraduate trainees' performance and selection (4). In particular, Prideaux et al. highlighted the difficulties faced by selection process designers due to the lack of research on this topic (1). A systematic review and meta-analysis conducted by Webster et al. identified 26 studies reporting correlation coefficients related to the validity of SJTs, with a pooled estimate of 0.32 (95% confidence interval (CI): 0.26 - 0.39; $P < 0.001$) (18). In a well-designed selection system, the researchers concluded that SJTs could be a useful selection tool, with studies of postgraduate medical applicants reporting larger validity coefficients in contrast to those conducted in undergraduate settings (β = 0.21, 95% CI: 0.11 - 0.31; $P < 0.001$) (16). However, another meta-analysis found that the internal reliability of SJTs varied

widely (Cronbach's alpha coefficients: 0.43 - 0.94) (19). Furthermore, Juster et al. found that including an SJT in the selection process will result in expanding access to medical education in terms of gender, racial, and ethnic diversity, increasing the variation of future physicians and thereby helping in addressing the needs of a changing patient population (20).

2. Objectives

This study aimed to test the validity of both SJT and the IFOM examination in predicting performance during MMIs among applicants to the OMSB Family Medicine Residency Program. In addition, the study evaluated the complementary role of an SJT in shortening the list of applicants undertaking the MMI panel. Implementing valid, cost-effective measures to shortlist applicants for postgraduate residency training programs will save time and effort. It is beneficial both on the part of interviewers, many of whom are required to interview a large number of applicants under the current selection system, as well as on the part of unsuitable applicants who would otherwise have to undergo unnecessary lengthy interviews and stressful waiting periods before finding out they have been rejected. Moreover, a streamlined and more efficient selection process will positively impact the future applicant's performance during residency, as well as ensure that the best possible candidates are admitted into residency training programs.

3. Methods

This retrospective cross-sectional study was conducted during the OMSB residency application process for the academic year 2021 and included 80 applicants who applied for the OMSB Family Medicine Residency Program as either their first or second choice. Various information was collected about the applicants, including their age, gender, graduating institute, work experience, current position, and grade point average (GPA) during their undergraduate medical degree. All data were obtained from the relevant OMSB administration and examination sections.

All candidates included in the study had passed the IFOM examination as a prerequisite for their residency application. Subsequently, they underwent a two-hour, computer-based SJT under invigilation conditions. The test consisted of 50 questions in nine attributes, including a commitment to professionalism, coping with pressure, effective communication, working effectively as part of a team, learning and professional development, organization and planning, patient focus, problem-solving and decision making, and self-awareness and insight. It had two

formats of best answer or multiple responses of the three most appropriate options (Table 1). The internal reliability of the SJT used during the selection process was high (Cronbach's alpha coefficient = 0.84). Afterwards, all applicants participated in three MMI panels, each consisting of two interviewers. Each panel lasted 10 min, during which the applicant had 5 min to answer one question posed by each interviewer. Questions comprised clinical and non-clinical elements, including those related to the applicant's personality, attitude, history of community service, and level of enthusiasm related to their chosen specialty. Each interviewer then rated the applicant's responses and their proficiency in spoken English (scored from 0 to 4) for a cumulative score out of 22.

The frequency of categorized variables was obtained using the Statistical Package for the Social Sciences (SPSS) version 23 (IBM Corp., Armonk, NY, USA). Pearson's correlation coefficient was used to evaluate the correlations of MMI scores with SJT and IFOM examination scores. In addition, a multiple hierarchical regression analysis was conducted to assess the predictors of the MMI score. A receiver operating characteristic (ROC) curve analysis was also performed to calculate the area under the curve (AUC) and determine the ability of SJT to categorize applicants as recommended or not recommended for the OMSB Family Medicine Residency Program based on their MMI scores. Overall, the AUC values of $\leq$ 0.5, 0.6 - 0.7, 0.7 - 0.8, 0.8 - 0.9, and 0.9 - 1 indicated no, poor, acceptable, excellent, and outstanding accuracy in selecting candidates, respectively (21). Ethical approval for this study was obtained from the OMSB Institutional Research and Ethics Committee.

4. Results

A total of 67 applicants completed the entire selection process (i.e., underwent the SJT and appeared for the MMI panels) and were included in the study. Of these, five (7.5%) were male, and 62 (92.5%) were female. The majority aged 25 - 29 years (n = 46; 68.7%), with less than five years of experience since graduation (n = 50; 74.6%). All, except one of the applicants, were of Omani nationality (n = 66; 98.5%), and most had graduated from institutions based within Oman (n = 61; 76.1%). The applicants were currently working in various types of job, such as general practitioner (n = 23; 34.3%), enrolled in a general foundation program (n = 27; 40.3%), or undertaking their internship year (n = 17; 25.4%). A total of 47 applicants (70.1%) had chosen family medicine as their first choice of specialty. More than half of the applicants had scored < 60 in the IFOM examination (n = 35; 52.2%), while almost three-quarters (n = 50; 74.6%) had scored > 70 in the SJT. A total of 18 applicants (26.9%) scored

$\geq$ 15 in the MMI panels and were, therefore, recommended for entry into the OMSB Family Medicine Residency Program (Table 2).

Overall, the SJT score was significantly correlated with the MMI score (r = 0.367; P = 0.002), whereas IFOM score (r = 0.105; P = 0.396), age (r = -0.010; P = 0.936), years of experience since graduation (r = 0.056; P = 0.651), and GPA (r = 0.036; P = 0.772) were not significantly correlated (Table 3). According to multiple hierarchical regression analysis, model 1 resulted in an R^2 value of 0.134, indicating that the SJT score accounted for 13.4% of the variance in the MMI score. Model 2 resulted in an R^2 value of 0.138, showing that the two variables together (the SJT score and IFOM examination score) accounted for 13.8% of the variance in the MMI score. The R^2 change between the two models was 0.004, signifying that the addition of the IFOM examination score caused a 0.4% additional variance to the MMI score (P = 0.586). Consequently, the inclusion of an additional variable (the IFOM examination score) in model 2 did not result in a statistically significant increase in the variance of the outcome variable (MMI score) compared to model 1 (P = 0.586 versus 0.002) (Table 4).

The regression model validated the ability of the SJT score to predict the MMI score according to the following formula: MMI score = 1.512 + 0.152 × SJT score (Table 5). However, upon calculating the ROC curve, the AUC was found to be 0.633 (95% CI: 0.482 - 0.785), indicating that the SJT score had a poor accuracy in categorizing individuals as either recommended or not recommended for the residency program (Figure 1). No significant differences in the mean MMI scores were observed between candidates based on the graduating institution (P = 0.260) or current position (P = 0.460) (Table 6).

5. Discussion

To date, most literature related to the application of SJTs has evaluated the predictive potential of this type of test on the performance of applicants in the end-of-training assessments (22), with few studies highlighting the use of SJTs as a selection method for entry into postgraduate training (4, 23). This gap in the existing research, along with introducing an SJT in the selection process for entry to the OMSB Family Medicine Residency Program, was the source of inspiration for the present study. Consequently, the SJT was performed for the first time in Oman, and the OMSB Education Committee of Family Medicine program decided not to use SJT scores as a shortlisting mechanism prior to the MMI panels for 2021 applicants. However, this decision may be amended in the future once

Table 1. Example of Two-Format Questions in Situational Judgment Test to the Oman Medical Specialty Board Family Medicine Residency Program

Format	Question	Options
Best answer	You have witnessed one of your colleagues indulging in inappropriate and socially unacceptable behaviour with a patient.	A. Advice your colleague to not to repeat it a gain.
		B. Discuss the issue with your team members.
		C. Inform your colleague that you saw his behaviour and warn him not to repeat it.
		D. Pretend you saw nothing.
		E. Report the issue to your supervisor.
Multiple responses to the three most appropriate options	You have been approached by a colleague as you are a batch representative, PGY3 resident, who has been on shifts with another PGY2 resident, for the last two weeks. He tells you that PGY2 has gradually become increasingly careless in monitoring and documenting patient records. On three occasions, he tells you that he has found him asleep in the common room whilst on duty. You know PGY2 very well and have never witnessed such behavior when you have worked with him previously.	A. Tell PGY3 resident that you have never witnessed such behavior from PGY2 resident.
		B. Suggest to PGY3 resident that he speaks to PGY2 resident directly about his concerns.
		C. Advise PGY3 resident to document his concerns.
		D. Inform a senior colleague about what the PGY3 resident has told you.
		E. Tell PGY3 resident you will speak to PGY2 resident about his behavior.
		F. Ask other members of the team whether they have witnessed this behavior in PGY2 resident.
		G. Advise PGY3 resident to speak to a senior colleague about his concerns.
		H. Ask PGY3 resident whether he has any evidence that patient safety is being compromised.

the validity of this SJT as a selection criterion is definitively confirmed.

The current study found a non-significant correlation between the IFOM examination and MMI scores, whereas there was a significant correlation between SJT and MMI scores. This can be explained by the fact that the IFOM examination concentrates on applied knowledge, while the MMI panels gauge various non-academic attributes and behaviors, similar to SJTs (23). Therefore, it is reasonable that SJTs would be a better predictor of MMI scores compared to the IFOM examination. Patterson et al. found that both SJT and MMI scores had strong correlations with the objective structured clinical examination (OSCE) scores (r = 0.44 and r = 0.46, respectively; P < 0.001 for both) during an end-of-training assessment. The researchers attributed these findings to the similarities between these methods because both extend beyond applied knowledge (23, 24). For the same reasons, Patterson et al. and Tiffin et al. found SJT scores to be a strong predictor of practical assessment of clinical examination skills scores during an end-of-training assessment, especially with construct-driven SJT

(23, 25). In addition, McDaniel et al., in a meta-analysis, indicated that behavioral tendency questions within SJTs resulted in the largest incremental validity compared to knowledge tendency questions (0.05 versus 0.03) (13).

In the present study, the SJT scores were positively correlated with the MMI scores and were significantly predictive of performance in the MMI panels (r = 0.367; P = 0.002). Therefore, it can be concluded that the SJT used in the present study may be utilized as a prerequisite to shortlisting applicants to the OMSB Family Medicine Residency Program prior to the interview phase. This would be especially beneficial considering the high number of applicants to this specialty compared to other programs, which results in a considerable administrative burden. Patterson et al. similarly concluded that SJTs might offer an additional, standardized benchmark by which program directors could shortlist applicants before the interview phase (26). Husbands et al. also reported statistically significant correlations between the SJT and MMI scores and multiple aspects of the HEXACO Personality Inventory (16). These findings further support the reliability and validity of SJTs
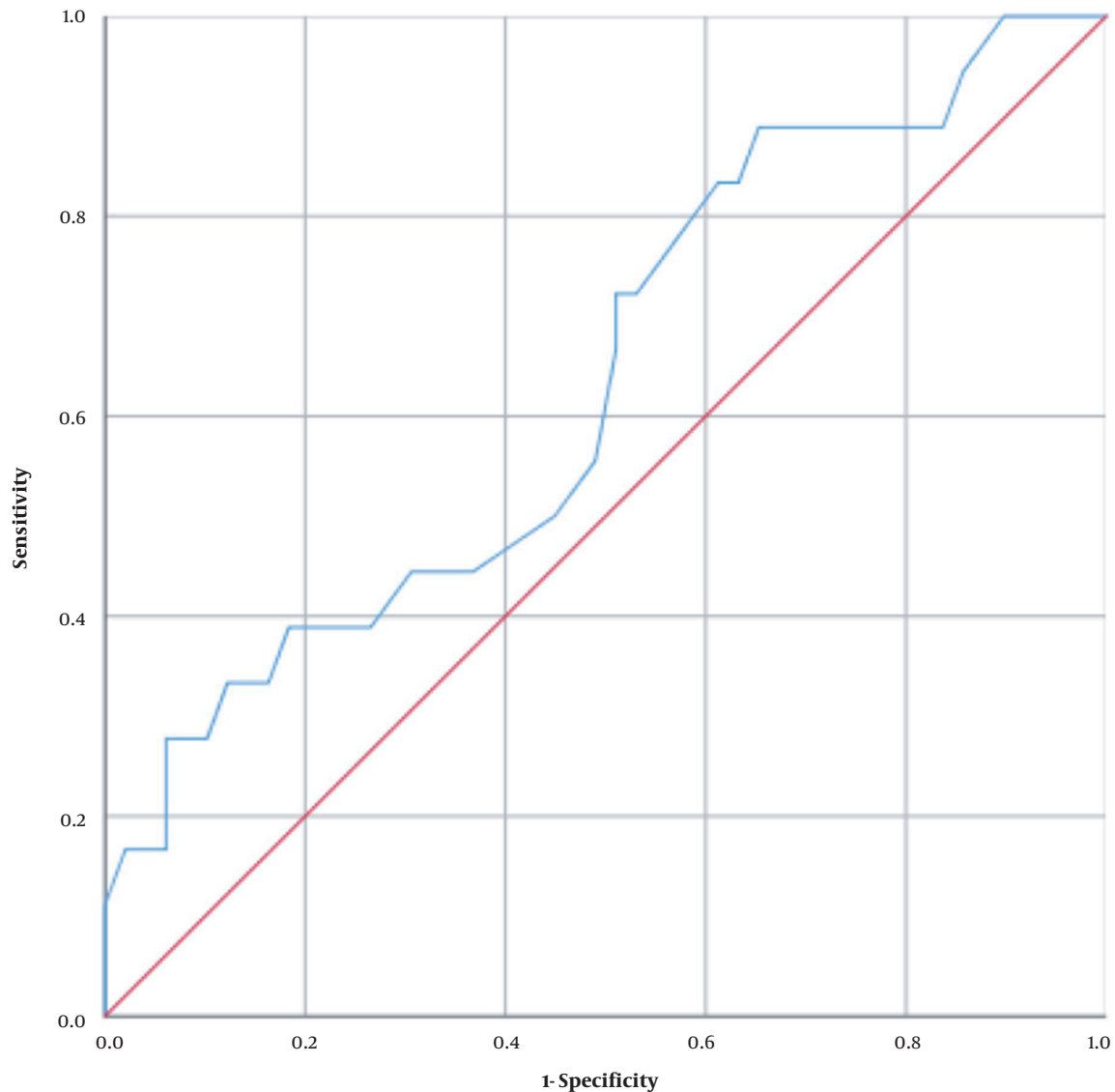
Figure 1. Receiver-operating characteristic curve of situational judgment test scores as a selection method to predict multiple mini-interviews scores among applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

and strengthen their utilization as a selection method, either as a pre-interview screening tool or as a partial replacement for MMI panels altogether (16). In another study, Patterson et al. confirmed the reliability of SJTs in medical selection, predicting performance in different medical education settings (undergraduate and postgraduate) and their incremental validity over other selection methods (9). Furthermore, the variations in SJT validity, reliability, and diversity benefits are attributed to the cultural differences between Western and Asian contexts concerning professionalism, values, and communication style (27).

Overall, the SJT scores were found to explain 13.4% of the variability in the MMI scores in the current study, which is considerably lower than the previous research, showing that this selection method explained 22% of the variation in selection center performance when used as a shortlisting tool during selection for specialty training (15). Patterson et al. reported that machine-marked SJT scores added incremental validity (R^2 = 0.143) when evaluating non-clinical attributes for the postgraduate medical selection compared to a clinical problem-solving test alone (26). Moreover, the researchers found that SJT was the most in-

Table 2. Demographic Characteristics of Applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

Variable and Category	No. (%)
Gender	
Male	5 (7.5)
Female	62 (92.5)
Nationality	
Omani	66 (98.5)
Non-Omani	1 (1.5)
Age (y)	
25 - 29	46 (68.7)
≥ 30	21 (31.3)
Years of experience since graduation	
< 5	50 (74.6)
5 - 10	12 (17.9)
> 10	5 (7.5)
Graduating institution	
Sultan Qaboos University	38 (56.7)
Oman Medical College	23 (34.3)
Other [a]	6 (9)
Graduating GPA	
< 2.5	19 (28.4)
2.5 - 2.9	32 (47.8)
≥ 3.0	16 (23.9)
Current position	
General practitioner	23 (34.3)
General foundation program	27 (40.3)
Intern	17 (25.4)
Choice of family medicine as a specialty	
First choice	47 (70.1)
Second choice	20 (29.9)
IFOM examination score	
50 - 59	35 (52.2)
60 - 69	22 (32.8)
70 - 79	9 (13.4)
≥ 80	1 (1.5)
IFOM examination year	
2020	40 (59.7)
2019	25 (37.3)
2018	2 (3)
SJT score	
< 70	17 (25.4)
70 - 79	27 (40.3)
≥ 80	23 (34.3)
MMI score	
< 15	49 (73.1)
≥ 15	18 (26.9)
Final decision	
Recommended [b]	18 (26.9)
Waitlisted	4 (6)
Not recommended	45 (67.2)

Abbreviations: GPA, grade point average; IFOM, International Foundations of Medicine; SJT, situational judgment test; MMI, multiple mini-interviews.
[a] Located abroad
[b] Based on MMI score

fluential independent predictor of selection center scores (adjusted R^2 = 0.213; P < 0.001) compared to two other shortlisting methods (a clinical problem-solving test and structured application form questions) (28). Lievens et al. reported that SJT showed criterion-related validity and incremental validity above cognitive ability and personality tests in predicting job performance (29). In contrast, Lievens and Patterson revealed that SJT resulted in an additional variance of 5.9% over a clinical problem-solving test regarding job performance (30).

In Oman, there is yet no previous data regarding the use of SJTs as a part of the selection process in the postgraduate medical education settings. A recent study by Al Hashmi and Klassen focused on developing an SJT to determine admission into initial teacher education programs (ITEPs) in Oman (31). The results demonstrated that the test had good internal consistency (α = 0.75). However, while SJT scores were significantly correlated with GPA as well as conscientiousness and agreeableness, out of the five personality measures, no correlation was found with interview scores. The researchers recommended using SJT as a promising additional selection method for ITEPs in Oman (31). Ongoing quality assurance of the selection process as a whole is imperative in postgraduate medical education to ensure improved competency in healthcare delivery, thereby enhancing future patient care. Moreover, adding an SJT in the residency selection process would reduce the resources needed for assessment, as SJTs are relatively cheap to administer to candidates compared to other methods (16).

5.1. Conclusions

The findings of this study indicated that an SJT is a valid selection method for entry into residency training alongside the traditional MMI panel. To the best of the authors' knowledge, this study represents the first analysis of SJTs in the medical setting in Oman and adds to existing international research regarding the predictive validity of this type of test in postgraduate medical settings. Furthermore, it opens the door for further longitudinal research evaluating the predictive validity of such selection methods during the end-of-training assessment in Oman, especially in comparison with practical assessments, such as the OSCE or simulated surgery.

5.2. Limitations

This study is limited by its small sample size compared to previous investigations conducted in the same field. Furthermore, the scope of this study was limited to the ability of an SJT to predict MMI score as a selection method for entry into the residency program. Further research is

Table 3. Correlations Between Selection Method and Multiple Mini-interview Scores Among Applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

Variable	MMI Score	Age	Years of Experience Since Graduation	IFOM Examination Score
Age				
r	-0.010	-	-	-
P-value	0.936	-	-	
Years of experience since graduation				
r	0.056	0.956 [a]	-	-
P-value	0.651	<0.001	-	-
IFOM examination score				
r	0.105	-0.233	-0.191	-
P-value	0.396	0.057	0.122	-
SJT score				
r	0.367 [a]	0.058	0.103	0.115
P-value	0.002	0.643	0.407	0.352

Abbreviations: MMI, multiple mini-interviews; IFOM, International Foundations of Medicine; SJT, situational Judgment test.
[a] Correlation is significant using Pearson's correlation coefficient at the 0.01 level (two-tailed).

Table 4. Multiple Hierarchical Regression Analysis of the Predictors of Multiple Mini-interview Scores Among Applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

Model	R	R^2	Adjusted R^2	SE	Change Statistics				
					R^2 Change	F Change	df_1	df_2	Sig. F Change
1 [a]	0.367	0.134	0.121	2.950	0.134	10.090	1	65	0.002
2 [b]	0.372	0.138	0.111	2.966	0.004	0.300	1	64	0.586

Abbreviations: SE, standard error; df, degrees of freedom; Sig, significance.
[a] Predictors: Constant situational judgment test (SJT) score
[b] Predictors: Constant, situational judgment test score, International Foundations of Medicine examination score

Table 5. Coefficients of the Predictors of Multiple Mini-interview Scores Among Applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

Model	Unstandardized Coefficients		t	P-Value	95% CI	
	B	SE			Lower	Upper
1						
Constant	1.512	3.676	0.411	0.682	-5.828	8.853
SJT score	0.152	0.048	3.176	0.002	0.056	0.247
2						
Constant	0.073	4.534	0.016	0.987	-8.984	9.131
SJT score	0.149	0.048	3.075	0.003	0.052	0.245
IFOM examination score	0.028	0.050	0.548	0.586	-0.073	0.128

Abbreviations: CI, confidence interval; SE, standard error; SJT, situational judgment test; IFOM, International Foundations of Medicine.

required to determine whether SJTs are suitable for evaluating long-term clinical training or work performance. Future studies assessing this aspect will provide additional evidence demonstrating the utility of SJTs as a shortlisting selection method prior to the MMI panel.

Table 6. One-Way Analysis of Variance to Compare Mean Multiple Mini-interviews Scores by the Institution and Current Position Among Applicants to the Oman Medical Specialty Board Family Medicine Residency Program (n = 67)

Variable and Category	Mean MMI Score ± SD	F (2,64)	P-Value
Graduating institution		1.377	0.260
Sultan Qaboos University	13.44 ± 3.29		
Oman Medical College	12.32 ± 3.06		
Other	14.31 ± 2.01		
Current position		0.786	0.460
General practitioner	13.79 ± 2.95		
General foundation program	12.70 ± 3.26		
Intern	12.92 ± 3.26		

Abbreviations: MMI, multiple mini-interviews; SD, standard deviation

Footnotes

References

1. Prideaux D, Roberts C, Eva K, Centeno A, McCrorie P, McManus C, et al. Assessment for selection for the health care professions and specialty training: consensus statement and recommendations from the Ottawa 2010 Conference. *Med Teach*. 2011;**33**(3):215–23. [PubMed ID: 21345061]. https://doi.org/10.3109/0142159X.2011.551560.

2. Patterson F, Ferguson E, Lane P, Farrell K, Martlew J, Wells A. A competency model for general practice: implications for selection, training, and development. *Br J Gen Pract*. 2000;**50**(452):188–93. [PubMed ID: 10750226]. [PubMed Central ID: PMC1313648].

3. Burgess A, Roberts C, Sureshkumar P, Mossman K. Multiple mini interview (MMI) for general practice training selection in Australia: interviewers' motivation. *BMC Med Educ*. 2018;**18**(1):21. [PubMed ID: 29370852]. [PubMed Central ID: PMC5785877]. https://doi.org/10.1186/s12909-018-1128-z.

4. Cousans F, Patterson F, Edwards H, Walker K, McLachlan JC, Good D. Evaluating the complementary roles of an SJT and academic assessment for entry into clinical practice. *Adv Health Sci Educ Theory Pract*. 2017;**22**(2):401–13. [PubMed ID: 28181019]. https://doi.org/10.1007/s10459-017-9755-4.

5. Whetzel DL, McDaniel MA. Situational judgment tests: An overview of current research. *Hum Resour Manag Rev*. 2009;**19**(3):188–202. https://doi.org/10.1016/j.hrmr.2009.03.007.

6. Cabrera MA, Nguyen NT. Situational Judgment Tests: A Review of Practice and Constructs Assessed. *Int J Sel Assess*. 2001;**9**(1&2):103–13. https://doi.org/10.1111/1468-2389.00167.

7. Lievens F. Situational judgement tests: introduction and research review. *Rev Psicol Trab Organ*. 2007;**23**(1):93–110.

8. Lievens F, Coetsier P. Situational Tests in Student Selection: An Examination of Predictive Validity, Adverse Impact, and Construct Validity. *Int J Sel Assess*. 2003;**10**(4):245–57. https://doi.org/10.1111/1468-2389.00215.

9. Patterson F, Zibarras L, Ashworth V. Situational judgement tests in medical education and training: Research, theory and practice: AMEE Guide No. 100. *Med Teach*. 2016;**38**(1):3–17. [PubMed ID: 26313700]. https://doi.org/10.3109/0142159X.2015.1072619.

10. Singaigireson S, Ramjeeawon N, Ravindra S, Shah N, Singh B. Is it fair for a junior doctor's deanery to be largely based on one test: a student's perspective. *Adv Med Educ Pract*. 2015;**6**:499–500. [PubMed ID: 26257529]. [PubMed Central ID: PMC4525794]. https://doi.org/10.2147/AMEP.S88360.

11. Herde CN, Lievens F, Solberg EG, Harbaugh JL, Strong MH, J. Burkholder G. Situational Judgment Tests as Measures of 21st Century Skills: Evidence across Europe and Latin America. *Rev Psicol Trab Organ*. 2019;**35**(2):65–74. https://doi.org/10.5093/jwop2019a8.

12. Smith KJ, Neely S, Dennis VC, Miller MM, Medina MS. Use of Situational Judgment Tests to Teach Empathy, Assertiveness, Communication, and Ethics. *Am J Pharm Educ*. 2021:8761. [PubMed ID: 34785498]. https://doi.org/10.5688/ajpe8761.

13. McDaniel MA, Hartman NS, Whetzel DL, Grubb W. Situational Judgment Tests, Response Instructions, and Validity: A Meta-Analysis. *Pers Psychol*. 2007;**60**(1):63–91. https://doi.org/10.1111/j.1744-6570.2007.00065.x.

14. Lievens F, Patterson F, Corstjens J, Martin S, Nicholson S. Widening access in selection using situational judgement tests: evidence from the UKCAT. *Med Educ*. 2016;**50**(6):624–36. [PubMed ID: 27170081]. https://doi.org/10.1111/medu.13060.

15. Ahmed H, Rhydderch M, Matthews P. Can knowledge tests and situational judgement tests predict selection centre performance? *Med Educ*. 2012;**46**(8):777–84. [PubMed ID: 22803755]. https://doi.org/10.1111/j.1365-2923.2012.04303.x.

16. Husbands A, Rodgerson MJ, Dowell J, Patterson F. Evaluating the validity of an integrity-based situational judgement test for medical school admissions. *BMC Med Educ*. 2015;**15**:144. [PubMed ID: 26330210]. [PubMed Central ID: PMC4557748]. https://doi.org/10.1186/s12909-015-0424-0.

17. Singh Sidhu H, Mahmood A, Sanghera R, Mandan J. Parallel to the Situational Judgement Test: is the Educational Performance Measure fair in ranking medical students on the UK Foundation Programme? *Adv Med Educ Pract*. 2016;**7**:603–4. [PubMed ID: 27822130]. [PubMed Central ID: PMC5087587]. https://doi.org/10.2147/AMEP.S121242.

18. Webster ES, Paton LW, Crampton PES, Tiffin PA. Situational judgement test validity for selection: A systematic review and meta-analysis. *Med Educ*. 2020;**54**(10):888–902. [PubMed ID: 32353895]. https://doi.org/10.1111/medu.14201.

19. McDaniel MA, Morgeson FP, Finnegan EB, Campion MA, Braverman EP. Use of situational judgment tests to predict job performance: a clarification of the literature. *J Appl Psychol*. 2001;**86**(4):730–40. [PubMed ID: 11519656]. https://doi.org/10.1037/0021-9010.86.4.730.

20. Juster FR, Baum RC, Zou C, Risucci D, Ly A, Reiter H, et al. Addressing the Diversity-Validity Dilemma Using Situational Judgment Tests. *Acad Med*. 2019;**94**(8):1197–203. [PubMed ID: 31033603]. https://doi.org/10.1097/ACM.0000000000002769.

21. Wang XM, Leung VA. The hidden curriculum and limitations of situational judgement tests for selection. *Med Educ*. 2020;**54**(10):872–5. [PubMed ID: 32515515]. https://doi.org/10.1111/medu.14264.

22. Li F, He H. Assessing the Accuracy of Diagnostic Tests. *Shanghai Arch Psychiatry*. 2018;**30**(3):207–12. [PubMed ID: 30858674]. [PubMed Central ID: PMC6410404]. https://doi.org/10.11919/j.issn.1002-0829.218052.

23. Patterson F, Rowett E, Hale R, Grant M, Roberts C, Cousans F, et al. The predictive validity of a situational judgement test and multiple-mini interview for entry into postgraduate training in Australia. *BMC Med Educ*. 2016;**16**:87. [PubMed ID: 26957002]. [PubMed Central ID: PMC4784338]. https://doi.org/10.1186/s12909-016-0606-4.

24. Shaw A. A perspective on situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Pers Individ Differ*. 2021;**177**. https://doi.org/10.1016/j.paid.2021.110850.

25. Tiffin PA, Paton LW, O'Mara D, MacCann C, Lang JWB, Lievens F. Situational judgement tests for selection: Traditional vs construct-driven approaches. *Med Educ*. 2020;**54**(2):105–15. [PubMed ID: 31872483]. https://doi.org/10.1111/medu.14011.

26. Patterson F, Baron H, Carr V, Plint S, Lane P. Evaluation of three short-listing methodologies for selection into postgraduate training in general practice. *Med Educ*. 2009;**43**(1):50–7. [PubMed ID: 19140997]. https://doi.org/10.1111/j.1365-2923.2008.03238.x.

27. Soemantri D, Findyartini A, Yolanda S, Morley E, Patterson F. Evaluation of Situational Judgment Tests in student selection in Indonesia and the impact on diversity issues. *BMC Med Educ*. 2022;**22**(1):239. [PubMed ID: 35366862]. [PubMed Central ID: PMC8976983]. https://doi.org/10.1186/s12909-022-03247-4.

28. Patterson F, Lopes S, Harding S, Vaux E, Berkin L, Black D. The predictive validity of a situational judgement test, a clinical problem solving test and the core medical training selection methods for performance in specialty training. *Clin Med (Lond)*. 2017;**17**(1):13–7. [PubMed ID: 28148572]. [PubMed Central ID: PMC6297587]. https://doi.org/10.7861/clinmedicine.17-1-13.

29. Lievens F, Peeters H, Schollaert E. Situational judgment tests: a review of recent research. *Pers Rev*. 2008;**37**(4):426–41. https://doi.org/10.1108/00483480810877598.

30. Lievens F, Patterson F. The validity and incremental validity of knowledge tests, low-fidelity simulations, and high-fidelity simulations for predicting job performance in advanced-level high-stakes selection. *J Appl Psychol*. 2011;**96**(5):927–40. [PubMed ID: 21480685]. https://doi.org/10.1037/a0023496.

31. Al Hashmi W, Klassen RM. Developing a situational judgement test for admission into initial teacher education in Oman: An exploratory study. *Int J Sch Educ Psychol*. 2019;**8**(sup1):187–98. https://doi.org/10.1080/21683603.2019.1630042.