Mini Surgical Simulation, Role Play, and Group and Behavioral Interviews in Resident Selection

Dotun Ogunyemi, MD, FACOG, MFM
Carolyn Alexander, MD
Edward Tangchitnob, MD
David Seil Kim, MD, PhD, MBA

ABSTRACT

Background A robust selection process is critical to residents' "cultural fit" and success in their program. Traditional selection methods have shortcomings.

Objective We describe a novel residency interview process for obstetrics-gynecology residents that incorporates behavioral, group, and surgical simulation multiple mini interviews (MMIs).

Methods In 2010, the Cedars-Sinai Medical Center obstetrics-gynecology residency program developed surgical simulation, role play, ethics group interview, and Accreditation Council for Graduate Medical Education competency-based behavioral interview stations.

Results From 2010 to 2012, a total of 199 applicants were interviewed, 62 ranked in the top 20, and 18 matched into the program. The MMI scores for interview stations were used in compiling our rank list and were found to adequately differentiate candidates. The MMI mean scores for role play, ethics interview, surgical simulation, and the behavioral interview for the top 20 ranked candidates were statistically significantly higher than those for other applicants. Standardized tests minimally correlated with various interview modalities. Applicants found the interview process acceptable. Implementing these MMI stations increased the total applicant interview time for the day by 15% (from 5.5 to 6.5 hours) and increased the face-to-face interview time from 2 to 4 hours. Approximately 42 hours of coordinator time was required for the yearly interview cycle.

Conclusions A multifaceted interview process utilizing MMI, group interview, and surgical simulation MMI is feasible and acceptable. The approach may decrease subjectivity and reliance on traditional interview methods and facilitate the selection of "compatible" residents into the program.

Introduction

The traditional resident selection processes rely on curricula vitae, United States Medical Licensing Examination (USMLE) scores, letters of recommendation, and personal interviews. However, residents' success in their training program is based on "cultural fit" within the institution, trainability on the Accreditation Council for Graduate Medical Education (ACGME) competencies, sound judgment, and the ability to work in teams. Inherent problems of traditional interviews are the variability in the interviewers' communication skills, bias, tolerance, and personality.[1–5] Many studies have shown that traditional interviews have poor interrater reliability and perceived gender and cultural bias, with inadequate reliability and questionable validity.[1–6]

The multiple mini interview (MMI) is a standardized, validated selection process that consists of several stations, each of which assesses different aspects of the applicants; the MMI is used in medical schools in Canada and has spread to other nations.[1–8] The behavioral format utilizes questions that relate to past behaviors and performance as a way to identify latent professionalism characteristics, and this approach is predictive of future behavior.[9] To identify desirable attributes in residency applicants (cultural fit, communication skills, teamwork, trainability, and professionalism), we developed a multifaceted interview process. Our objective is to describe this novel MMI that incorporates (1) an ACGME competency-based behavioral interview; (2) an ethics group interview; (3) role play; and (4) surgical simulation. Our hypothesis is that a competency-based MMI process, if used to create the rank list, can measure ACGME competencies other than cognitive abilities and can provide an additional tool in applicant selection.

DOI: http://dx.doi.org/10.4300/JGME-D-15-00203.1

Editor's Note: The online version of this article contains a description of the development and process of the multiple mini interview, selection criteria for the position of obstetrics-gynecology intern head, CT role play scenario tool, and social media and ethics station scoring sheet.

Methods

MMI Development

The Cedars-Sinai Medical Center obstetrics-gynecology residency program is an urban academic program located in Los Angeles, California, with 6 categorical

residency positions for each postgraduate level class, for a total of 24 residents. On average, the program receives over 400 applicants for 6 categorical positions each year. Between 2009 and 2010, the curriculum committee developed an innovative admission process for intern selection (provided as online supplemental material).

What was known and gap

Traditional resident selection processes have focused on intellectual attainment and may have undervalued other important attributes and elements of "cultural fit."

What is new

A multiple mini interview (MMI) for resident selection, including surgical simulation, role play, and behavioral interview stations.

Limitations

Single program study limits generalizability; interview stations and survey tools lack validity evidence.

Bottom line

The MMI reduces subjectivity and reliance on standardized scores and increases selection based on cultural fit.

Behavioral Interview Station

For each ACGME competency, 3 to 4 subtitles were constructed, with 3 to 4 open-ended questions for each subtitle. We created 3 behavioral questionnaires titled as follows:

1. Practice-based learning and improvement and interpersonal and communication skills
2. Professionalism and systems-based practice (provided as online supplemental material)
3. Patient care and medical knowledge

Each applicant had 3 one-on-one interviews with faculty or residents for exposure to all ACGME competency-based behavioral items. The program director's behavioral interview questionnaire focused on the domains of persistence, adaptability, self-/stress management, action orientation, teamwork and support, interpersonal communication, and strategic focus.

Role Play Clinical Scenario Station

Applicants were assigned to a 10-minute clinical scenario station in groups of 5, with 2 facilitators/raters. Applicants were provided with instruction cards and randomly assigned role play as nurse, chief resident, obstetrics-gynecology intern, and 2 observers. The scenario was a woman at 30 weeks gestation presenting with head trauma. The goal was to ensure that she received a head computed tomography (CT), despite challenges purposely proposed to the applicants, such as an uncooperative radiology technician. Applicants were evaluated on (1) collaboration; (2) effective communication; (3) situational awareness; (4) proactive behavior; (5) use of chain of command; and (6) professionalism, all on a 1 to 5 Likert scale. A rubric was created that described the Likert scale 1 through 5 scores for each of the behaviors evaluated to ensure consistency and reliability of scoring (provided as online supplemental material). Applicants were not tested on their clinical knowledge.

Social Media Ethics Group Station

Applicants were interviewed in groups of 5 by a faculty and a resident. A vignette was presented about a resident who posted pictures of pathology specimens, patients, and colleagues on a social media site. The following questions were asked: What do you think? What are ethical issues? Is it appropriate? Should resident take photos of patients? Should we use social networks?

The applicants were measured on leadership behavior, assertive communication, consideration, social and ethical insights, and acting in a proactive and nondisruptive way, using a 1 to 5 Likert scale. The rubric was graded on a Likert scale 1 through 5 scores for each of the behaviors evaluated to ensure consistency and reliability of the scoring process (provided as online supplemental material).

Surgical Technical Simulation Station

Applicants were assigned in groups of 3 to the simulation station with resident facilitator/raters. Interviewees spent about 5 minutes each on a LAP Mentor high-fidelity virtual reality simulator (Simbionix USA Corp, Cleveland, OH) performing a peg transfer exercise; a peg transfer exercise on a standard laparoscopic box trainer; and playing the Super-MonkeyBall2 video game on the XBox360. The rater assessed the applicants on (1) baseline technology prowess; (2) ability to understand new technology; (3) general attitude; (4) dexterity with video monitors; (5) attitude toward technology; (6) virtual reality simulator performance; and (7) video game skills.

Rank list meeting occurred after the interviews; applicants were ranked based on scores from all MMI stations and by consensus (provided as online supplemental material).

Acceptability was measured empirically using applicant responses. Applicants were invited to anonymously complete the applicant survey at the end of the interview day. To assess if MMIs can help differentiate candidates for a rank list, we compared the 20 top-ranked applicants and matched applicants that matriculated into the program to the remainder

TABLE 1
Interview Day Activity and Time Required

Interview Day Activity for Each Applicant	Time Required	MMI Extra Time
1. Orientation, presentations, and breakfast	60 min	
2. Departmental grand rounds	60 min	
3. Three one-on-one interviews with residents or faculty	90 min	
4. Interview with the program director	30 min	
5. Medical center tour	30 min	
6. Lunch with residents and faculty	60 min	
7. Surgical technical simulation station		30 min
8. Role play clinical scenario station		15 min
9. Social media ethics group station		15 min
Total time per candidate	5 hr 30 min	1 hr
Total time per candidate for interview-only events (items 3, 4, and 7–9, with the last 3 items requiring 2 interviewers)	2 hr	2 hr

Abbreviation: MMI, multiple mini interview.

of the applicants. Chi-square test, Student *t* tests, and Pearson correlation coefficient were used as indicated. A *P* value of $< .05$ was taken as significant.

The institution's Institutional Review Board gave exempt approval for the study.

Results

This sample encompasses the 199 applicants who interviewed for the intern classes between 2010 and 2012. A total of 62 (2 applicants had the same score) were ranked as the top 20 applicants, and 18 matched into the program. Of all applicants, 28 (14%) were men and 171 (86%) were women, versus the top 20 ranked applicants with 14 men (23%) and 48 women (77%; $P = .024$). Of the applicants, 139 (70%) were non-Hispanic White and 60 (30%) were other ethnicities (mainly Asian American). There were no statistical differences between the mean USMLE Step 1 and 2 scores for the 20 top-ranked applicants versus others. For residents who matched to the program versus others, USMLE Step 1 score was lower (221.61 [SD = 15.99] versus 231.49 [SD = 14.17]; $P = .02$), with no statistical differences in USMLE Step 2 scores.

Per each applicant, the MMI increased the time for the interview day by 15% (from 5.5 to 6.5 hours). The MMI stations increased face-to-face interview time for applicants 100%, from 2 to 4 hours (TABLE 1), and the MMI required approximately 23% (33 hours) more faculty/resident time than the prior traditional interview process, which had required about 110 total faculty/resident hours for each interview year season. Approximately 42 hours of coordinators' time were required per interview cycle for the year. All of the interview modalities were internally developed by the program directors with support from the coordinators.

Behavioral Interview Station

The mean scores of our 20 top-ranked candidates were significantly higher than those for other candidates, whereas for the matched group, only the score for patient care–medical knowledge was significantly higher (TABLE 2). Significant correlations were found between the

TABLE 2
Significant Associations of the Behavioral Interview Process With Top-Ranked and Matched Candidates

	Other Applicants (n = 134) Mean (SD)	Top 20 Ranked (n = 63) Mean (SD)	*P* Value
PBLI-ICS	3.42 (0.94)	3.98 (0.73)	$< .001$
PROF-SBP	3.62 (0.80)	3.90 (0.82)	.03
PC-MK	3.66 (0.89)	3.96 (0.78)	.03
Total	3.69 (0.79)	3.99 (0.57)	.007
	Other Applicants (n = 178)	**Matched Applicants (n = 18)**	***P* Value**
PC-MK	3.68 (SD = 0.83)	4.12 (SD = 0.65)	.015

Abbreviations: PBLI-ICS, practice-based learning and improvement and interpersonal and communication skills; PROF-SBP, professionalism and systems-based practice; PC-MK, patient care and medical knowledge.

TABLE 3
Significant Associations of Role Play Clinical Scenario Station With Top-Ranked and Matched Candidates

	Top-Ranked Applicants Versus Others, Mean (SD)			Matched Applicants Versus Others, Mean (SD)		
	Other Applicants (n = 134)	Top 20 Ranked (n = 63)	*P* Value	Other Applicants (n = 152)	Matched Applicants (n = 16)	*P* Value
Collaboration	4.01 (0.88)	4.45 (0.55)	.004	4.14 (0.83)	4.59 (0.49)	.003
Effective communication	3.90 (0.87)	4.51 (0.55)	< .001	4.04 (0.84)	4.59 (0.52)	.001
Situational awareness	3.97 (0.90)	4.61 (0.48)	< .001	4.13 (0.86)	4.53 (0.70)	.015
Proactive behavior	3.96 (0.88)	4.54 (0.62)	< .001	4.09 (0.86)	4.53 (0.70)	.029
Use of chain of command	3.94 (0.92)	4.52 (0.53)	< .001	4.08 (0.87)	4.41 (0.66)	.08
Professionalism	4.23 (0.89)	4.74 (0.37)	< .001	4.35 (0.82)	4.72 (0.45)	.008
Total score	24.05 (4.88)	27.37 (2.63)	< .001	24.8 (4.63)	27.41 (3.03)	.005

behavioral interview scores and with the USMLE Step 2 scores (r = 0.20, *P* = .015), program director's assessment (r = 0.28, *P* = .001), role play (r = 0.24, *P* = .002), social media ethics group interview (r = 0.16, *P* = .04), and surgical skills simulation (r = 0.26, *P* < .001).

Role Play Clinical Scenario Station

The 20 top-ranked candidates scored significantly higher in role play compared to other applicants. Likewise, the matched group scored significantly higher in the total score and in all but 1 component compared to other applicants (TABLE 3). There was a strong positive correlation between role play and ethics group interview (r = 0.81, *P* < .001), with weaker positive correlations with the behavioral interview (r = 0.24, *P* = .002), surgical skills simulation (r = 0.21, *P* = .007), and USMLE Step 1 scores (r = 0.17, *P* = .029).

Social Media Ethics Group Station

Both the 20 top-ranked candidates and the matched group scored significantly higher in the ethics group interview than other applicants (TABLE 4). There were weak positive correlations with the surgical skills simulation (r = 0.19, *P* = .014) and male sex (r = 0.17, *P* = .029).

Surgical Technical Simulation Station

The mean overall scores of our 20 top-ranked candidates (3.7 [SD = 0.71] versus 3.33 [SD = 0.76]; *P* = .002) and the candidates who matched to our program (3.84 [SD = 0.5] versus 3.42 [SD = 0.78]; *P* = .004) were significantly higher than those of other candidates. There was a weak positive correlation with the male sex (r = 0.25, *P* < .001).

Program Director Assessment

The mean overall score of our 20 top-ranked candidates was significantly higher than that of other candidates (3.44 [SD = 0.80] versus 2.90 [SD = 0.79]; *P* < .001). There were no significant differences between those that matched versus others.

TABLE 4
Group Interview on Facebook Posting: Top-Ranked Applicants' and Matched Applicants' Scores Versus Others' Scores

	Top-Ranked Applicants Versus Others, Mean (SD)			Matched Applicants Versus Others, Mean (SD)		
	Other Applicants (n = 115)	Top 20 Ranked (n = 50)	*P* Value	Other Applicants (n = 152)	Matched Applicants (n = 16)	*P* Value
Leadership	3.81 (0.82)	4.48 (0.58)	.004	3.98 (0.82)	4.20 (0.65)	NS
Assertiveness	3.78 (0.83)	4.48 (0.52)	< .001	3.94 (0.83)	4.30 (0.56)	.034
Consideration	4.09 (0.85)	4.64 (0.45)	< .001	4.20 (0.81)	4.73 (0.37)	.012
Ethics and social insight	4.00 (0.80)	4.54 (0.69)	< .001	4.11 (0.80)	4.57 (0.53)	.021
Proactive	3.98 (0.85)	4.53 (0.56)	< .001	4.10 (0.80)	4.50 (0.53)	.017
Not aggressive	4.13 (0.85)	4.65 (0.47)	< .001	4.23 (0.81)	4.70 (0.41)	.029
Total score	23.77 (4.43)	27.33 (2.72)	< .001	24.56 (4.64)	27.00 (2.46)	.003

Abbreviation: NS, not significant.

TABLE 5
Applicants Acceptability Survey Using a Likert Scale[a]

Survey Question	Interview Process Assessed by Question	Mean Score
Communications with the program were courteous, and my questions were answered promptly.	Program directors and coordinators	4.78
The information packet contained useful information.	Orientation of applicants to interview process	4.41
The interview length was adequate.	Duration of all interview day stations	4.33
Faculty asked appropriate questions during the interviews, and provided useful additional information about the program.	Behavioral interview process and program director assessment	4.46
The luncheon provided a good opportunity to interact with other candidates and to talk in more detail with the residents, faculty, and fellows.	Social skills of faculty and residents	4.53
The residents with whom I spoke were courteous, asked appropriate questions, and provided useful insights into the operation of the program.	Behavioral interview process, surgical technical simulation, role play clinical scenario, and social media ethics group interview	4.67
The tour of the facilities was beneficial.	Social skills of residents	4.27
I was treated fairly throughout the interview process and had an opportunity to express myself during my time with the faculty and residents.	Behavioral interview process and program director assessment	4.75
I enjoyed participating in the procedure workshop segment of the interview.	Surgical technical simulation	4.57
I enjoyed participating in the group interactive segment of the interview.	Role play clinical scenario and social media ethics group interview	4.04
The interview and visit has increased my interest in Cedars program.	Effectiveness of the interview process	4.69
The program's Facebook page is more helpful to me than the program's website.	Efficacy of program's social media site	4.53
The program's Facebook page gave me a positive perception of the program.	Orientation of applicants to program	4.55
Overall rating of your interview		4.68

[a] Likert scale of 1 to 5 (1, strongly disagree; to 5, strongly agree).

Applicant Acceptability

The results of a postinterview survey given to the applicants during the second year of the new interview process suggested that applicants found all aspects of the interview process acceptable. The total score for acceptability was 4.7 (out of a 1 to 5 scale, with 5 being the highest score), with the lowest acceptable component, the group interviews, having a score of 4.0. The survey completion rate was 78% (47 of 60 applicants). Additionally, applicants' open-ended comments were very positive (TABLE 5).

Discussion

We have shown the feasibility of implementing a multifaceted interview with MMI group interviews, clinical role play scenario, gynecological surgical simulation, and behavioral-based interviews in obstetrics-gynecology residency interviews. There were significant opportunity costs associated with substantial time commitments by faculty, residents, and coordinators. For programs that do not have simulators, laparoscopic box trainers, or video game boxes, acquiring these devices could add significant financial costs.

Since top-ranked and matched applicants scored significantly higher on the MMI stations, we infer that our integrated MMI processes can be used to provide additional assessments that further differentiate among applicants than is possible with traditional methods.

It was reassuring that the applicants found the interview process to be acceptable, without evidence of gender or cultural bias. Previous studies have shown that US residency applicants highly value personal connections made during the interview

day.[6,10] We tried to make it "fun" and provide sources of personal connections for the applicants.

The group interview and the role play allowed the assessment of interpersonal and communication skills, professionalism, and systems-based practice, which was important to the culture of the program. The simulation allowed the assessment of potential surgical skills. Our current interview process omits the ethics group interview, as it provides the same information as the role play station.

Our approach has limitations. Although the interview modalities were standardized, the specific MMI stations did not have demonstrated reliability and validity evidence.[10–20] Our survey lacks validity evidence, and obtaining feedback from applicants on the day of the interview may have introduced social desirability bias.

Conclusion

The MMI requires additional faculty, program director, and coordinator time, as well as access to a simulation. However, the information resulting from it benefits resident selection and fit and provides value for the added effort.

References

1. Fraga JD, Oluwasanjo A, Wasser T, Donato A, Alweis R. Reliability and acceptability of a five-station multiple mini-interview models for residency program recruitment. *J Community Hosp Intern Med Perspect*. 2013;17(3):3–4.
2. Finlayson HC, Townson AF. Resident selection for a physical medicine and rehabilitation program: feasibility and reliability of the multiple mini-interview. *Am J Phys Med Rehabil*. 2011;90(4):330–335.
3. Razack S, Faremo S, Drolet F, Snell L, Wiseman J, Pickering J. Multiple mini-interviews versus traditional interviews: stakeholder acceptability comparison. *Med Educ*. 2009;43(10):993–1000.
4. Uijtdehaage S, Doyle L, Parker N. Enhancing the reliability of the multiple mini-interview for selecting prospective health care leaders. *Acad Med*. 2011;86(8):1032–1039.
5. Humphrey S, Dowson S, Wall D, Diwakar V, Goodyear HM. Multiple mini-interviews: opinions of candidates and interviewers. *Med Educ*. 2008;42(2):207–213.
6. Soares WE 3rd, Sohoni A, Hern HG, Wills CP, Alter HJ, Simon BC. Comparison of the multiple mini-interview with the traditional interview for U.S. emergency medicine residency applicants: a single-institution experience. *Acad Med*. 2014;90(1):76–81.
7. Eva KW, Reiter HI, Rosenfeld J, Norman GR. The relationship between interviewers' characteristics and ratings assigned during a multiple mini-interview. *Acad Med*. 2004;79(6):602–609.
8. Knorr M, Hissbach J. Multiple mini-interviews: same concept, different approaches. *Med Educ*. 2014;48(12):1157–1175.
9. Easdown JL, Castro PL, Shinkle EP, Small L, Algren J. The behavioral interview, a method to evaluate ACGME competencies in resident selection: a pilot project. *J Educ Perioper Med*. 2005;7(1):1–10.
10. Pau A, Jeevaratnam K, Chen YS, Fall AA, Khoo C, Nadarajah VD. The multiple mini-interview (MMI) for student selection in health professions training—a systematic review. *Med Teach*. 2013;35(12):1027–1041.
11. Eva KW, Reiter HI, Trinh K, Wasi P, Rosenfeld J, Norman GR. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ*. 2009;43(8):767–775.
12. Campagna-Vaillancourt M, Manoukian J, Razack S, Nguyen LH. Acceptability and reliability of multiple mini interviews for admission to otolaryngology residency. *Laryngoscope*. 2014;124(1):91–96.
13. Hofmeister M, Lockyer J, Crutcher R. The acceptability of the multiple mini interview for resident selection. *Fam Med*. 2008;40(10):734–740.
14. Dore KL, Kreuger S, Ladhani M, Rolfson D, Kurtz D, Kulasegaram K, et al. The reliability and acceptability of the multiple mini-interview as a selection instrument for postgraduate admissions. *Acad Med*. 2010;85(suppl 10):60–63.
15. Liao SC, Hsiue TR, Lin CH, Huang AM. Multiple mini-interviews combined with group interviews in medical student selection. *Med Educ*. 2014;48(11):1104.
16. Tavares W, Mausz J. Assessment of non-clinical attributes in paramedicine using multiple mini-interviews. *Emerg Med J*. 2015;32(1):70–75.
17. Jones PE, Forister JG. A comparison of behavioral and multiple mini-interview formats in physician assistant program admissions. *J Physician Assist Educ*. 2011;22(1):36–40.
18. Nuthalapaty FS, Jackson JR, Owen J. The influence of quality-of-life, academic, and workplace factors on residency program selection. *Acad Med*. 2004;79(5):417–425.
19. Eva KW, Reiter HI, Trinh K, Wasi P, Rosenfeld J, Norman GR. Predictive validity of the multiple mini-interview for selecting medical trainees. *Med Educ*. 2009;43(8):767–775.

20. Dodson M, Crotty B, Prideaux D, Carne R, Ward A, de Leeuw E. The multiple mini-interview: how long is long enough? *Med Educ*. 2009;43(2):168–174.

Dotun Ogunyemi, MD, FACOG, MFM, is Professor, Department of Obstetrics & Gynecology, Oakland University William Beaumont School of Medicine System, Vice Chair for Women's Quality and Safety, Beaumont Health, and Clinical Professor, Department of Obstetrics & Gynecology, David Geffen School of Medicine, University of California, Los Angeles (UCLA); **Carolyn Alexander, MD,** is Attending Physician, Southern California Reproductive Center & Cedars-Sinai Medical Center, and Associate Professor, The David Geffen School of Medicine, UCLA; **Edward Tangchitnob, MD,** is Medical Director of Minimally Invasive Gynecologic & Robotic Surgery, Citrus Valley Medical Center; and **David Seil Kim, MD, PhD, MBA,** is Assistant Residency Program Director, Division of Minimally Invasive Surgery, Department of Obstetrics & Gynecology, Cedars-Sinai Medical Center, and Assistant Clinical Professor, The David Geffen School of Medicine, UCLA.

Funding: The authors report no external funding source for this study.

Conflict of interest: The authors declare they have no competing interests.

The authors would like to thank Jennifer Kim for her contributions with the data entry and collection.

Corresponding author: David Seil Kim, MD, PhD, MBA, Cedars-Sinai Medical Center, 160 West Tower, 8635 West 3rd Street, Los Angeles, CA 90048, 310.423.2914, fax 310.423.0140, david.kim3@cshs.org

Received May 18, 2015; revisions received October 30, 2015, and February 16, 2016; accepted February 22, 2016.