Annals of Medicine and Surgery 71 (2021) 103019

Contents lists available at ScienceDirect

Annals of Medicine and Surgery

journal homepage: www.elsevier.com/locate/amsu

Cross-sectional Study

A pilot project to introduce the multiple mini interview (MMI) at a Borneo medical school: The universiti Malaysia Sabah experience- a cross-sectional study

Nicholas Pang [a], Fairrul Kadir [a,*], Assis Kamu [b], Ho Chong Mun [b], Jiann Lin Loo [c], Fatimah Ahmedy [a], Mohammad Saffree Jeffree [a]

[a] Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[b] Faculty of Science and Natural Resources, Universiti Malaysia Sabah, Jalan UMS, 88400, Kota Kinabalu, Sabah, Malaysia
[c] Ty Derbyn, Wrexham Maelor Hospital, Betsi Cadwaladr University Health Board, 17 Dolydd Road, Wrexham, LL13 7TE, UK

ARTICLE INFO

Keywords:
Interview
Admission
Borneo

ABSTRACT

Introduction: The Multiple Mini Interview (MMI) demonstrates efficacy and superiority over traditional medical interviews in assessing non-cognitive domains during the recruitment of medical undergraduates. At Universiti Malaysia Sabah (UMS), a five-station MMI was piloted in 2019, featuring a mix of three examiner-driven stations (assessing professionalism, ethics, and motivation to study medicine), and two roleplayer-driven stations (assessing empathy and science communication specifically, and communication skills in general).

Methods: 260 candidates were grouped into two separate geographical groups – urban and suburban/rural. Descriptive analysis, skewness and kurtosis were performed for normality assessment, whereas Cronbach's alpha, McDonald's omega, and Greatest lower bound assessed internal consistency. For validity measures, correlations were calculated between scores for separate stations, overall scores, urban and suburban/rural status. Also, exploratory factor analysis was performed on the five stations as validity measures. Difficulty and discrimination indices were calculated as quality measures. Qualitative analysis was performed on "red flag" comments detailing grossly unsuitable candidates.

Results: Roleplayer-driven stations yielded more red flags than examiner-driven stations. The three examiner-driven stations were significantly and moderately correlated (rho between 0.602 and 0.609, $p < 0.001$). The Empathy roleplayer-driven station was not correlated with two examiner-driven stations and only weakly correlated with the Ethics examiner-driven and the Science Communication roleplayer-driven station. Factor analysis suggests a three-factor model. The two roleplayer-driven stations stood as independent factors, and the three examiner-driven stations coalescing as one factor provided the best explanatory model. Quality measures suggest all five stations had suitable discriminatory properties (all >0.530), whereas the stations were distributed equally in difficulty index.

Conclusion: The UMS MMI has identified specific skillsets that may be in short supply in our incoming medical students. Also, it illustrates the yawning gap between academic knowledge and 'translational' scientific knowledge and communication skills.

1. Introduction

The Multiple Mini Interview (MMI) was first conceived by Kevin Eva and the team at McMaster University in 2004 [1]. It was initially developed as a backlash towards traditional interviews, which focused on getting candidates to justify their selection of medicine orally but did not objectively test personal qualities and skills that candidates possessed [1]. The MMI was initially conceived as an OSCE (Objective Structured Clinical Examination) style encounter [1]. Each station in this OSCE would assess objectively non-cognitive and specific cognitive skills that would be desirable to the practice of medicine. MMIs aim to reduce the individual bias of a traditional interview conducted by a single examiner and even out multiple biases across multiple examiners by averaging scores across various domains and areas [2]. The evidence

* Corresponding author.
E-mail address: fairrul.kadir@ums.edu.my (F. Kadir).

https://doi.org/10.1016/j.amsu.2021.103019
Received 29 September 2021; Received in revised form 1 November 2021; Accepted 1 November 2021
Available online 5 November 2021

suggests that MMIs used to select undergraduate health programs appear reasonable feasibility, acceptability, validity, and reliability [3]. Furthermore, MMIs appear to be non-biased for age, gender, or socio-economic status; nevertheless, specific ethnic and social backgrounds demonstrated low performance in a minimal number of published studies [3].

Literature suggests there is no "the MMI" [4]. There is vast flexibility as there is no "one correct way" to conduct an MMI. Of course, there are specific founding and fundamental principles that are helpful to adhere to, not to make an MMI rely on a carousel of traditional interviews [1,5]. With that, these principles were espoused in the efforts of Universiti Malaysia Sabah (UMS) to devise, construct, and administer its own Borneo-flavoured MMI. It was hypothesized that roleplayer-driven encounters would be of better efficacy than merely a series of examiner-driven encounters, which would become a series of truncated traditional interviews. Hence, a concerted effort was made to train reliable and authentic roleplayers who would provide an assessment with discriminatory ability.

2. Methods

2.1. MMI station setup

After comparing literature across the board for MMIs from multiple universities, it was decided to include the following five stations. Three interviewer-led, scenario-driven stations were designed: a traditional **motivation** and preparation station; an **ethics** dilemma; and a **professionalism** issue with an aspect of teamwork. Two roleplayer-led, observer-assessed stations followed: a **breaking bad news** scenario focusing on empathy; and a **"science communication"** station focusing on the ability to perform simple mathematics under duress and communicate with a panicky allied healthcare worker. All five stations were trialled with a pilot group to ensure feasibility and iron out glaring administration difficulties.

Each station was designed to assess two separate domains, with certain domains assessed across multiple stations. The domains assessed included: motivation and preparation, teamwork and leadership (2 stations), ethics, professionalism, empathy, communication skills (3 stations), and logical thinking (2 stations). Each station also included a third "general impression" domain. All domains were scored out of 5, and rubrics were provided to guarantee reliability. Each station also contained a "red flag" feature to flag up candidates who cause the interviewer or observer great concern about the person's qualities to do medicine. Kevin Eva et al.'s recent 12 Tips article was used as the guiding point [5].

Three innovations were employed, differing from the literature. Firstly, fourth-year medical students instead of professionals were consciously chosen as roleplayers for both observer-rated stations. Secondly, candidates were permitted to use English or Bahasa Malaysia (the national language) to answer all five stations. This decision was consciously made to mimic the unique reality of pursuing an English-language undergraduate degree but working in a Malay-language society. Thirdly, a Science Communication station was set up to simulate a panicking nurse and doctor encounter. No prior knowledge of pharmacology, medication names, or science knowledge was required to perform the calculation. This station differed from existing stations in other MMIs in literature and was hypothesized to test translational rather than traditional cognitive abilities coupled with basic communication techniques.

2.2. Statistical analysis of MMI

Descriptive statistics, skewness, and kurtosis were performed to assess for normality. T-tests were performed between urban and sub-urban/rural groups to assess for significant differences in station scores.

2.3. Quality, validity and reliability of MMI

A few statistical methods were employed to assess further the MMI level, including quality, validity, and reliability measures. **Quality measures** involved calculating the difficulty and discrimination indices of each of the stations. Then, the indices will be classified based on classifications which are further elaborated in the "Results" section.

Difficulty index (D) was calculated as such:

$$D = \frac{No.\ of\ candidates\ with\ the\ total\ score\ 80\%\ (12\ out\ of\ 15)\ and\ above}{Total\ candidates} \times 100$$

The Discrimination index (R) was calculated as such:

$$D = \frac{H - L}{n}$$

where.

H is the number of top 27% candidates with a total score of 80% (12 out of 15) and above.

L is the number of bottom 27% candidates with a total score of 80% (12 out of 15) and above.

n is the total number of 27% of the candidates.

It was decided to use Cronbach's alpha, McDonald's omega, and greatest lower bound to test the internal consistency for reliability measures. For validity, exploratory factor analysis (EFA) was performed together with Spearman's correlations for concurrent validity. Spearman's rho coefficients were calculated to identify correlations between scores for separate stations, overall scores, and urban and suburban/rural status. Exploratory factor analysis was also performed to assess if the five stations measured five separate constructs or whether a more parsimonious factor structure was more suitable statistically. Principal component analysis was used with the Varimax rotation method employing Kaiser normalization. Eigenvalues >1 and scree plots were used to identify an optimal number of factors. All correlations less than 0.3 were eliminated in the rotated component matrix, and all communalities of correlations >1.0 were identified. This work was reported in line with the STROCSS criteria [6]. It was also registered with a Research Registry (unique research number: researchregistry7254).

3. RESULTS.

In total, 260 candidates attempted the MMI examination in 2019.72 candidates were from an urban background (the Klang Valley, referring to the Greater Kuala Lumpur conurbation of 5 million population). In contrast, the remaining 188 were from semi-urban or rural backgrounds from the rest of Malaysia.

Table 1 demonstrates the descriptive means of the totals for each station and total red flags per station. There were more red flag candidates in the non-urban group and for roleplayer-driven stations. The urban group only notched 4 red flags (5.56%), whereas the non-urban group clocked up 17 red flags (9.04%). Only 4 red flags in total were reported for the "Professionalism" station, and 1 red flag in total for the remaining 2 stations. This compares unfavourably with the Empathy station (8 red flags total) and Science Communication station (8 red flags total). Analysis of red flags included "angering actors" and "repeatedly breaking news for the wrong diagnosis despite a reminder by an examiner".

As observed in Table 2, minimum total scores for all stations ranged from 3 to 6, while the median scores for all stations were 12, except for the breaking bad news station where it was 9. Mean total scores for all stations ranged from 10.00 to 12.39 out of a possible total of 15. Cronbach alpha for the MMI was 0.824, suggesting good internal consistency. Skewness and kurtosis for all individual component scores, all total station scores, and total scores for the MMI were all within $\pm$ 2, suggestive of normal distribution. Statistically, performing a *t*-test, only "Teamwork and Leadership" for Station 3 significantly from all the measured domains.

Table 1
Descriptive statistics

	Semi-urban/Rural (N = 188)		Urban (N = 72)		Overall (N = 260)			
	Mean	Count	Mean	Count	Mean	Lower CL for Mean	Upper CL for Mean	Count
Station 1: Motivation	4.03		4.07		4.04	3.93	4.15	
Station 1: Teamwork	3.83		3.83		3.83	3.73	3.94	
Station 1: General	3.97		3.92		3.96	3.85	4.06	
Total score - Station 1	11.35		11.82		11.48	11.1	11.86	
Station 1: Red flag		0		0				0
Station 2: Ethics	4.09		4.18		4.12	4.03	4.21	
Station 2: Communication	4.07		4.21		4.11	4.01	4.22	
Station 2: General	4.13		4.24		4.16	4.06	4.26	
Total score - Station 2	11.71		12.62		11.97	11.59	12.34	
Station 2: Red flag		0		1				1
Station 3: Professionalism	4.09		4.29		4.15	4.05	4.25	
Station 3: Teamwork	4.01		4.26		4.08	3.98	4.18	
Station 3: General	4.02		4.04		4.03	3.94	4.12	
Total score - Station 3	11.54		12.6		11.83	11.46	12.2	
Station 3: Red flag		3		1				4
Station 4: Empathy	3.18		3.33		3.22	3.09	3.35	
Station 4: Communication	3.39		3.46		3.41	3.33	3.5	
Station 4: General	3.42		3.44		3.43	3.32	3.54	
Total score - Station 4	9.99		10.24		10.06	9.76	10.36	
Station 4: Red flag		8						8
Station 5: Logical thinking	3.59		3.67		3.61	3.48	3.74	
Station 5: Communication	3.84		3.82		3.83	3.72	3.94	
Station 5: General	3.78		3.69		3.75	3.64	3.86	
Total score - Station 5	11.2		11.18		11.19	10.87	11.52	
Station: Red flag		7		2				9

Table 2
Individual station scores.

	Min Score	Max Score	Mean	Median	SD
Station 1 (Motivation)	5	15	11.48	12	3.12
Station 2 (Ethics)	6	15	11.97	12	3.02
Station 3 (Professionalism)	4	15	11.3	12	2.95
Station 4 (Empathy)	5	15	10.06	9	2.44
Station 5(Science Communication)	3	15	11.19	12	2.64

2.4. Validity: correlations between station scores

Spearman's rho coefficients were calculated for all total scores per station and correlated with total scores too (see Table 3). As demonstrated, Stations 1 to 3 were significantly and moderately correlated to each other (ρ between 0.602 and 0.609, $p < 0.001$). Station 4 (Empathy) was not correlated with Station 1 and 3 and only weakly correlated with the Ethics and the Science Communication station results. Station 5 (Science Communication) was only weakly correlated with all 4 other stations. This result suggests that Empathy and Science Communication might stand alone as separate factors in this assessment tool, which will be further explored during factor analysis.

Also, total communication scores were moderately correlated with the 3 stations where it was explicitly measured (rho between 0.506 and 0.661, $p < 0.001$). It was only weakly correlated with Motivation and Professionalism stations (rho between 0.348 and 0.349, $p < 0.001$). Total scores were correlated moderately with all 5 station scores.

2.5. Differences between urban and rural groups

When examining differences between the Urban and Suburban/Rural groups, there was a difference between both groups for Ethics ($t = -2.599$, $p = 0.010$) and Professionalism ($t = -3.27$, $p = 0.011$) stations. There was no significant difference between both groups for Motivation and Preparation, the Breaking Bad News, and Science Communication stations.

2.6. Exploratory factor analysis

2.6.1. Five-factor model

Initially, as there were 5 factors on the examination theorized to fall into and measure separate constructs (based on literature review from other universities' MMIs), a five-factor model was used. However, based on the eigenvalues shown in Table 4, only 2 factors had an eigenvalue >1. This corresponded with the kink in the scree plot, which was after the 2nd factor (see Fig. 1). After removing all components <0.3, only 2 stations fell into the 2 factors with eigenvalue >1, whereas one factor (Professionalism station) fell into two separate factors. Also, a two-factor model only accounted for 64% of the variance. When a 3rd factor was

Table 3
Spearman's rho coefficient between individual stations and total.

Domains	Communication	General Impression	Station 1 (Motivation)	Station 2 (Ethics)	Station 3 (Professionalism)	Station 4 (Empathy)	Station 5 (Science Comm)
Total score	.854**	.972**	.602**	.609**	.609**	.469**	.551**
Communication	–	.822**	.348**	.661**	.349**	.506**	.552**
General impression	–	–	.588**	.615**	.579**	.474**	.500**
Station 1 (Motivation)	–	–	–	.227**	.316**	.084	.271**
Station 2 (Ethics)	–	–	–	–	.280**	.176**	.177**
Station 3 (Professionalism)	–	–	–	–	–	.055	.179**
Station 4 (Empathy)							.151*

**: $p < 0.05$, *: $p < .001$.

Table 4
Eigenvalues for a five-factor model.

Component	Total	Initial eigenvalues	
		% of variance	Cumulative %
1	2.241	44.826	44.826
2	1.079	21.589	66.415
3	.882	17.649	84.064
4	.435	8.709	92.773
5	.361	7.227	100.000

considered, the 3rd factor also had an eigenvalue approaching 1, and a 3-factor model accounted for 84% of the variance. The rotated component matrix for a 5-factor model did not cleanly separate, with Professionalism cross-loading into two separate factors (see Table 5). Thus, it was proposed to run a further analysis based on a 3-factor model.

2.6.2. *Three-factor model*

As suggested in Table 6, the items fell into 3 factors in this factor loading model, corresponding to the three groups – all the scenario-driven stations as one factor, with the breaking bad news station and the science communication station as separate factors. There were also no factors with loading >1.0; hence there was no demonstrable communality. As with the earlier 5-factor model, the eigenvalues remained similar, with 3 factors having eigenvalues all >0.88 (see Fig. 2) Therefore it was found that a 3-factor model was the best to explain the distribution of the stations.

The following quality measures were performed as per Table 7. The difficulty index suggested that the Empathy and Breaking Bad News station was the most difficult, whereas the Professionalism and Ethics stations were the easiest (see Table 9). All five stations had good discrimination indices >0.53 suggesting they were very good items (see Table 8). For reliability, on all three reliability measures, scores were between 0.629 and 0.721 as per Table 10, suggesting acceptable internal consistency.

3. Discussion

The results suggest that the MMI has reasonable reliability and validity, with acceptable internal consistency using three separate measures. The five stations yield a three-factor model employing EFA. There is no disparity between geographical regions for the Science Communication station and Empathy stations. However, descriptively, the mean scores for both stations appear to be significantly lower than those of the preceding three scenario-driven stations. Interestingly, there was a statistical difference between urban and rural areas for the Ethics and Professionalism station. This may be linked to higher levels of education, and hence better eloquence and confidence in communication, for these two non-interviewer driven stations. Conversely, the Motivation station demonstrated no significant difference, suggesting the motivation to study medicine is similar between urban and rural areas.

The factor analysis demonstrates another important finding – that Science Communication and Empathy stations fell into distinct factors. This is corroborated from correlation testing that indicates that the Science Communication and Empathy stations were somewhat weakly or not correlated with the other three interviewer-driven stations. Hence, this demonstrates the roleplayed stations were able to separate candidates based on their ability to perform skills or have attributes that were conducive or favourable to the practice of medicine.

Due to the stringent entry criteria for interviews to study medicine, students from different geographical regions have roughly similar entry academic preuniversity results and Malaysian English University Test proficiency results. Our results hence alleviate a concern that urban candidates might perform far better with a largely roleplay-based examination. Correlations suggest that only Ethics and Professionalism

Table 5
Rotated Component Matrix for the five-factor model.

Station	Component				
	1	2	3	4	5
Motivation			.921		
Ethics	.920				
Professionalism	.305				.906
Empathy				.996	
Science Comm		.993			

Table 6
Rotated Component Matrix for the three-factor model.

Station	Component		
	1	2	3
Motivation	.823		.
Ethics	.847		
Professionalism	.878		
Empathy			.993
Science Comm		.987	

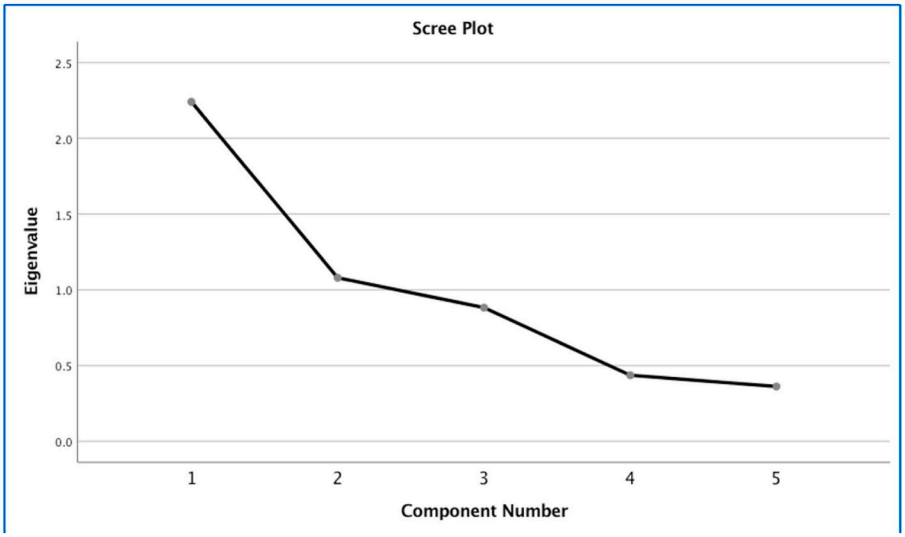

Fig. 1. Scree Plot for the five-factor model.

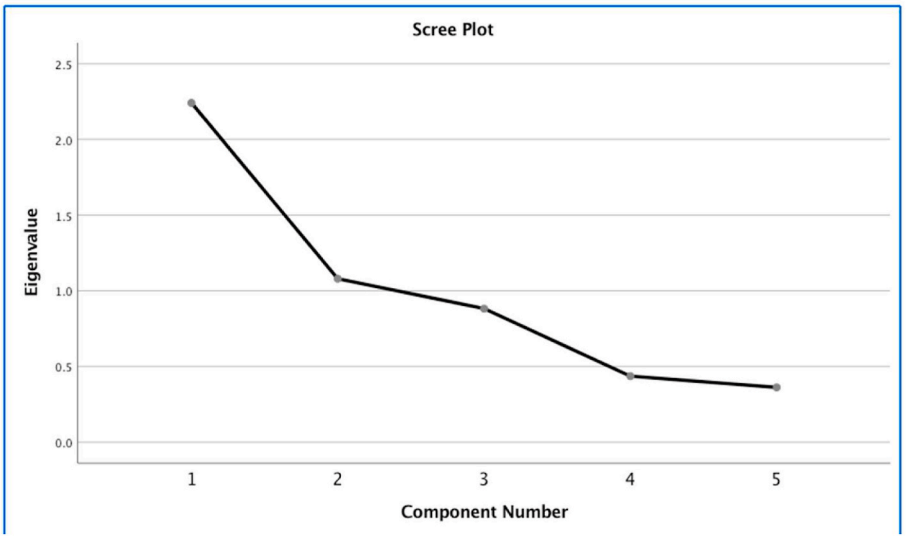

Fig. 2. Scree Plot for the three-factor model.

Table 7
Quality measures.

Quality measure	Station 1	Station 2	Station 3	Station 4	Station 5
Difficulty index	59.23%	65.00%	71.54%	30.00%	51.15%
Interpretation of difficulty index	Average	Easy	Easy	Difficult	Average
Discrimination index	0.63	0.63	0.56	0.53	0.74
Interpretation of discrimination index	Very good item	Very good item	Very good item	Very good item	Very good item

Table 8
Interpretation of the difficulty index.

Range	Difficulty level
20% and below	Very difficult
21–40%	Difficult
41–60%	Average
61–80%	Easy
More than 80%	Very easy

Table 9
Interpretation of the discrimination index.

Range	Verbal description
0.40 and above	Very good item
0.30–0.39	Good item
0.20–0.29	Fair item
0.09–0.19	Poor item

Table 10
Scale reliability statistics.

	Cronbach's α	McDonald's ω	Greatest lower bound
Scale	0.629	0.683	0.721

stations (both interviewer-rather than scenario-driven stations) showed a significant difference between urban and rural candidate performances. One factor contributing to the parity is allowing candidates to answer in either English or Bahasa Malaysia. This would have taken away the edge urban candidates might have had when dealing with English-based roleplay scenarios.

This stands us apart from the crowd in other contemporaneous MMIs in Malaysia, as many of them still have a more significant proportion of English language stations [7]. This decision was consciously made to mimic the reality of working in a Malay language society. Even though medical school education and doctor-level discussions, conferences, and symposia are done exclusively in English, our MMI hoped to reflect our values as a patient-centred medical school. Allowing the use of Bahasa Malaysia from start to finish in an examination, we felt, did not compromise our "English language proficiency assessment" – as that had been adequately assessed through the Malaysian University English Test (MUET), where a rigorous set of group interview speaking tests had already been conducted [8].

The probable lack of discriminatory factors between the interviewer-based stations is further suggested by factor analysis. Factor analysis demonstrates that the three interviewer-based stations measure a statistically similar construct. This raises the question of whether there is any value in having more than one interviewer-based station in an MMI. Multiple universities have utilized a mix of interviewer-based and roleplay-based assessments; the consensus across the literature is that half of the stations would fall into either category. However, suppose factor analysis suggests that interviewer-based stations measure the same construct. In that case, medical education designers of MMIs need to pause to consider whether it is better to "convert" the ethics and professionalism stations to a roleplay-based, rather than an interviewer-led, station.

This illustrates the problem with many scenario-based interview methods. Short scenarios are not unique to the MMI and were not invented by Eva et al. formal interviews since time immemorial have featured questions about what individuals would do in certain situations beyond asking people their personal motivations to study medicine to a greater or lesser degree. However, being able to "say one would do something" is not the same as being able or willing to "do something".

Hence there is a possibility that, if the 3 stations (motivation, ethics, and professionalism) were converted from scenario-based to roleplay stations, there might be a more precise separation of candidates who had the desirable qualities to study medicine. Standardized scenario-based stations always run the risk of scripted answers and bias from the interviewer. No doubt there are also biases from roleplayers, as naturally empathetic or naturally friendly candidates might be unconsciously assisted.

3.1. Qualitative feedback from all stakeholders

Candidates appeared by and large to enjoy their examination. This was echoed back to us from the feedback received after examinations from candidates who had sat for it. Most respondents surveyed informally post-exam felt that despite being difficult, it was pretty enjoyable. Many thought that it had been a reasonable assessment as it has tested their ability to do things rather than say something.

Examiners felt that the MMI format took away a lot of the positive confounding effect of practice and rehearsal upon preparation for an interview. Each station had a stem that began deliberately vaguely, and no two candidates received the same MMI experience as they were redirected away from a vague stem. This, in one fell swoop, eliminated much of the benefit that people of higher socioeconomic classes had by more extended training and practice. Statistically, this was borne out by the fact that candidates from all subgroups scored somewhat similar grades in the Ethics and Professionalism station. There was no apparent benefit as the station was redirected suitably, and there was an equal chance of scoring well.

The roleplayers felt satisfied, too, as they were involved democratically in the selection of their colleagues. This also gave them an insight into administrative workings of selection of professionalism-based criteria and taught them the pattern recognition techniques of making quick heuristic decisions regarding which candidates were better in a short period – pattern recognition techniques that are eminently translatable into the practice of bedside medicine.

The Science Communication station proved to be unique in both construction and outcome: it was not initially designed to test resilience. However, to our surprise, both in the pilot and in the actual MMI, it proved to be the station that teased out the more resilient students under pressure. Students separated pretty clearly into the ones who could do it without much panicking within a matter of 30 s, the majority who only completed it after much prompting from the nurse, and the odd one who simply was overwhelmed and received a red flag. Feedback from candidate's post-exam confirmed that they had experienced substantial cognitive dissonance during the task. They recognized how simplistic the task was, yet they could not complete it within the stipulated time, let alone the 1 min it would normally take under non-stress conditions. The performance on the Science Communication station is not reflective of the actual difficulty of the calculation, as the calculation in question merely involves basic knowledge of arithmetic and ratios and has been trialled on upper primary school students and members of the public of average mathematical ability.

3.2. Recommendations and future directions

Further research in the pipeline will focus on various aspects of the MMI captured in 2019. This information is not available currently, as the students' identities in the MMI cannot be released so far due to confidentiality restrictions. However, there is a lot of potential research that can be undertaken. This includes but is not limited to: the psychometric properties of the current MMI as an instrument (reliability and validity), the ethical dilemma of involving current medical students as part of the recruitment process, further analyses of specific sociodemographic factors that might predict or affect performance on particular domains of the MMI, and focusing on analyses of certain aspects, e.g. ethics, professionalism, empathy and science communication that have shown variability or unexpected results in preliminary results discussions.

One of the essential things longitudinally is to assess if our MMI has any predictive validity. There will be interesting results after one, two, and the first batch has undertaken five years of medical studies to receive the MMI. It will be interesting to observe whether our 2019 MMI can predict future performance in examinations (both clinical and preclinical), especially OSCEs, and whether it will be able to predict dropout rates, future professional behaviour, and resilience – all the qualities that the MMI claims to measure at intake. This longitudinal review will look at the validity of MMI not merely as an admission tool, either within or outside the health profession education context, but as a prediction tool. Its stated claim to purpose is one of both sensitivity (the ability to detect those who are suitable for healthcare professions) and specificity (the ability to pick up those who are grossly unsuitable), which multiple longitudinal research studies have born out. Only time will tell whether the inaugural UMS MMI has achieved its initial purpose.

Also, as described above, median scores for Empathy and Communication Skills are much lower, suggesting it is an aspect that was seriously lacking. Median and mean scores for empathy were significantly lower than scores for the three interviewer-led stations, averaging above 4 for all three geographical regions assessed. This is worrying as empathy is a core skill expected of a doctor to have. This provides us with further direction in which to orient our undergraduate medical curriculum. There will be a new integrated curriculum rolled out in 2022. Hopefully, this curriculum will allow a higher proportion of "soft skills" and empathy-building to be delivered, as our medical faculty also recognizes that MMI has identified a remediable deficit.

4. Conclusion

The MMI is a worldwide trend in the reform of admissions processes to medical schools worldwide. Much research has demonstrated its utility, both within and without the healthcare professions and in different socio-cultural contexts in other continents. The UMS MMI has identified that certain specific skill sets may be in short supply in our incoming medical students. These results give us hope, rather than despair – for with the aid of our MMI as a score at baseline, we are now no longer guiding them in the dark, blindly, towards professional moulding into doctors, but rather, with intimate knowledge of their individual personal qualities pre-medicine, and purpose and vision to make them better versions of their current selves throughout the five years' time they have been granted with UMS Medical Faculty.

Provenance and peer review

Not commissioned, externally peer reviewed.

Ethical approval

No ethical approval is required for this cross-sectional study.

Sources of funding

This research did not receive any specific grant from funding agencies in the public, commercial, or non-profit sectors.

Consent

Not applicable as this study does not involve any patient data.

Author contribution

Nicholas Pang and Jiann Lin Loo were responsible for conceptualizing the study and reviewing the literature. Fairrul Kadir, Assis Kamu, Ho Chong Mun were involved in data collection, interpretation of studies, and manuscript drafting. Mohammad Saffree Jeffree and Fatimah Ahmedy were critically analyzing literature and expert input in synthesizing knowledge and finalizing the manuscript's content.

Declaration of competing interest

The authors report no conflict of interest nor proprietary or commercial interest in any product mentioned, or concept discussed in this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.amsu.2021.103019.

References

[1] K.W. Eva, J. Rosenfeld, H.I. Reiter, G.R. Norman, An admissions OSCE: the multiple mini-interview, Med. Educ. 38 (2004) 314–326.
[2] K.W. Eva, H.I. Reiter, K. Trinh, P. Wasi, J. Rosenfeld, G.R. Norman, Predictive validity of the multiple mini-interview for selecting medical trainees, Med. Educ. 43 (2009) 767–775.
[3] E.L. Rees, A.W. Hawarden, G. Dent, R. Hays, J. Bates, A.B. Hassell, Evidence regarding the utility of multiple mini-interview (MMI) for selection to undergraduate health programs: a BEME systematic review: BEME Guide No. 37, Med. Teach. 38 (2016) 443–455.
[4] M. Knorr, J. Hissbach, Multiple mini-interviews: same concept, different approaches, Med. Educ. 48 (2014) 1157–1175.
[5] Kevin W. Eva, Macala Catherine, Fleming Bruce, Twelve tips for constructing a multiple mini-interview, Med. Teach. 41 (2019) 510–516.
[6] R. Agha, A. Abdall-Razak, E. Crossley, N. Dowlut, C. Iosifidis, G. Mathew, for the STROCSS Group, The STROCSS 2019 guideline: strengthening the reporting of cohort studies in surgery, Int. J. Surg. 72 (2019) 156–165.
[7] A.F.A. Rahim, M.S.B. Yusoff, Validity evidence of a multiple mini interview for selection of medical students: Universiti sains Malaysia experience, Educ. Méd. J 8 (2016).
[8] S. Rethinasamy, K.M. Chuah, The Malaysian University English Test (MUET) and its use for placement purposes: a predictive validity study, Electron journal of foreign language teaching 8 (2011) 234–245.