Abstract

Background: Education literature worldwide is replete with studies evaluating the effectiveness of Multiple Mini Interviews (MMIs) in admissions to medicine but less than one percent of published studies have been conducted in selection to nursing and midwifery programmes.

Objectives: To examine the predictive validity of MMIs using end of programme clinical and academic performance indicators of pre-registration adult, child, and mental health nursing and midwifery students.

Design and setting: A cross-sectional cohort study at one university in the United Kingdom.

Participants: A non-probability consecutive sampling strategy whereby all applicants to the September 2015 pre-registration adult, child, mental health nursing and midwifery programmes were invited to participate. Of the 354 students who commenced year one, 225 (64%) completed their three-year programme and agreed to take part (adult 120, child 32, mental health nursing 30 and midwifery 43).

Methods: All applicants were interviewed using MMIs with six and seven station, four-minute models deployed in nursing and midwifery student selection respectively. Associations between MMI scores and the cross-discipline programme performance indicators available for each student at this university at the end of year three: clinical practice (assessed by mentors) and academic attainment (dissertation mark) were explored using multiple linear regression adjusting for applicant age, academic entry level, discipline and number of MMI stations.

Results: In the adjusted models, students with higher admissions MMI score (at six and seven stations) performed better in clinical practice (p<0.001) but not in academic attainment (p=0.122) at the end of their three-year programme.

Conclusion: These findings provide the first report of the predictive validity of MMIs for performance in clinical practice using six and seven station models in nursing and midwifery programmes. Further evidence is required from both clinical and academic perspectives from larger, multi-site evaluations.
1. INTRODUCTION

The World Health Organisation (WHO, 2016) state that patients’ ‘experience’ of care is critical and not just complementary to routine clinical practices. In the context of the burgeoning needs of health service provision worldwide, nurses and midwives require expert clinical skills as well as the attributes and values necessary to provide the highest standard of person-centred care. The delivery of this care is dependent on identifying the ‘right’ students (Great Britain, Department of Health, DoH, 2013).

Selecting the ‘right’ students for admission onto healthcare programmes is extremely challenging. A myriad of factors potentially impact on student’s value choices and consequential care provision in clinical situations after programme commencement (XXXX et al 2018) where values are not static constructs (Pattison and Pill, 2004). Extraneous influences and organisational pressures at macro and micro levels can impact on values behaviour (Parks and Guay, 2009) and thereby programme performance. Despite these complexities, healthcare student admissions staff are responsible for making transparent, defensible decisions about whom they accept onto their programmes informed by best available evidence.

To inform the evidence-base for healthcare student selection and aiming to identify those most likely to excel in caring professions, ‘values-based recruitment’ (VBR) (recruitment and selection according to pre-defined values and attributes, Health Education England, HEE, 2014) approaches like multiple mini interviews (MMIs) (Rees et al. 2016) have been widely adopted. Originating in medical student admissions (Eva et al, 2004), MMIs have been increasingly used in nursing and midwifery student selection processes worldwide. However, a recent narrative synthesis systematic review (XXXX, 2018) revealed that 57 out of 63 published studies evaluating the effectiveness of MMIs were in a medical or allied health professional student selection context with less than one percent (n=6) having been conducted specifically in admissions to nursing and midwifery programmes. This study redresses this deficit by evaluating the predictive validity of MMIs in selection of pre-registration adult, child and mental health nurses and midwifery students at one university in the United Kingdom (UK).

2. BACKGROUND

MMIs were first deployed in medical student selection in Canada (Eva et al. 2004) but have since been adopted in admissions to nursing, midwifery, paramedic practice, pharmacy, dental, veterinary medicine, dietetics, occupational therapy, physiotherapy, emergency medicine and physician associate programmes internationally (McAndrew et al. 2012, Perkins et al. 2013, Grice et al. 2014, Tavares & Mausz, 2015, Soares et al. 2015, Cameron et al., 2017, Kumar et al. 2018).

MMIs feature a number of stations, typically between five and twelve (Rees et al. 2016) in a timed circuit. At each station, candidates answer either ‘situational’ or ‘behavioural’ scenario questions. A situational MMI station might ask ‘what would you do in this situation and why’ in a response to a hypothetical scenario (Figure 1). A behavioural station question requests the candidate to reflect on something that has previously happened to them (Yoshimura, 2015). Scenarios are designed to assess pre-defined personal domains (attributes and/or values). Each candidate’s response is scored on a structured pro forma. The aggregate of assessments generated by the process is designed to be less susceptible to unconscious bias than a traditional 1:1 or panel interview (Eva et al. 2004, Roberts et al. 2014, Reiter et al. 2018).
The number of published studies evaluating the effectiveness and acceptability of MMIs in the last decade has risen exponentially, but meta-analysis is not possible due to heterogeneity in relation to study design, MMI models and participant groups. Published evidence suggests however that MMIs are a reliable admissions approach with the potential to identify those most likely to excel in their education programmes (Rees et al. 2016). However, the majority of research has been conducted in admissions to medicine, with less than one percent in nursing and midwifery student selection context (XXXX, 2018). This paper is novel, as the first published evaluating the predictive validity of MMIs specifically in a nursing and midwifery student selection. It also builds on previously reported first year findings (XXXX et al, 2018) thereby broadening the evidence base for MMIs outside the field of medicine.

3. METHODS

3.1 Study objectives

To examine the predictive validity of MMIs using end of programme clinical and academic performance of pre-registration adult, child, and mental health nursing and midwifery students taking into consideration: age; academic entry level; discipline and MMI design.

3.2 Design

A cross-sectional cohort study was conducted at one university in the United Kingdom (UK) with the 2015 entry students to the three fields of nursing (adult, child, mental health) and midwifery programmes. Students’ year three clinical practice performance was assessed by their clinical mentors. This assessment process is standardised across these programmes using a validated rubric with specific criteria set by the UK Nursing and Midwifery Council (NMC, 2009, 2010). Additionally, in the nursing programmes a short practice-based objective structured clinical examination (OSCE) (Harden and Gleeson, 1979, Mitchell et al. 2015) was completed by students where an aggregate final mark was then calculated to determine ‘overall’ clinical practice performance by the clinical mentor. Academic performance was measured using dissertation mark as this was the only generic and consistently assessed academic outcome indicator available across disciplines, at the end of year three, at this university.

In midwifery student selection a seven station, four-minute MMI model with one interviewer at each station was deployed. This is congruent with published evidence evaluating the effectiveness of different MMI models which states that 7-12 stations offer optimal reliability (XXXX et al. 2014, 2015, Rees et al. 2016). However, due to limited availability of interviewers, the seven-station model was not possible in nursing student selection. Consequently, a pragmatic decision was made by admissions staff at this university to deploy a six station, four-minute model with one interviewer at each station for selection to nursing programmes. The systematic development of the scenarios used to assess values/attributes at each station was underpinned by the principles of DeVellis’s (2003) framework for scale development and informed by previous work by XXXX (2014, 2015). Communication skills and kindness, compassion and empathy were assessed as generic attributes at each station complemented by station-specific attributes (Table 1). Interviewers were trained in MMI administration prior to interviewing to optimise inter-rater reliability.

3.3 Data collection and participants

A non-probability consecutive sampling strategy was employed whereby in all applicants to the September 2015 pre-registration adult, child, mental health nursing and midwifery programmes were
invited to participate. Of the 354 students who were offered and accepted a place on one of the nursing or midwifery programmes, 301 (85%) agreed to take part and completed year one (the characteristics of year one of this sample are reported elsewhere, XXXX, 2018) and 225 (75%) went onto complete the whole year three year programme (adult 120, child 32, mental health nursing 30 and midwifery 43). The 76 students for whom data is not reported (between years one and three) are accounted for: 8 students were removed from the analysis as their dissertation was supervised by the PI; 20% (n=15) and 33% (n=25) had extensions or were on ‘interrupt’ respectively at the time of analysis so the data were incomplete; 23% (n=17) exited the programme with a Certificate of Higher Education (CHEd and did not undertake a dissertation and 14% (n=11) were programme terminated. Due to ethical restrictions it was not possible to follow up the students to find out why they exited the programme with a CHEd or the reasons for programme termination. Additionally, there were insufficient numbers to conduct statistical analysis in relation to those 11 who were programme terminated.

The sample was sufficient to detect a correlation coefficient of 0.5 with 80% power in a two-sided significance test between MMI score and end of year three clinical practice and academic (dissertation) performance.

3.4 Statistical analysis

Participant responses at each MMI station were scored against ten criteria on a seven-point Likert scale from unsatisfactory to excellent. This resulted in a total maximum score of 490 and 420 for the seven (midwifery) and six (nursing) station models respectively. Details of the scoring criteria are withheld from publication as this would compromise test security for future MMIs. Scores for midwifery students were rescaled to conform to the same scale as nursing MMI models thus bearing the same weight in statistical analysis.

Firstly, Pearson’s correlation coefficient (r) was employed to explore the relationship of participants’ MMIs scores with clinical practice and academic performance (dissertation mark). Then single and multiple linear regression was used to examine the relationship. The multiple models were used to account for the effect of age at recruitment, academic entry level and discipline in the regression models. Collinearity did independent variables in the multiple models did not exceed 0.6 fulfilling the assumption for inclusion.

In addition, the effect of age and academic entry level on clinical practice and academic performance was examined with Pearson’s correlation and t-test. A series of t-test were conducted to evaluate differences in MMI scores, clinical and academic performance across disciplines. Complete-case analysis were undertaken and participants who had incomplete data (n=40) were excluded from the analysis. The cut-off for statistical significance was considered at p = 0.05. Statistical analyses were carried out using SPSS version 22 (SPSS Inc., Chicago, USA) and R version 3.0.2 (R Development Core Team, Austria).

The findings reported in this article build on previously published data (XXXX, 2018) which examined associations between MMI scores and clinical practice performance at the end of year one of this study participants at this university, as well as the reliability (internal consistency) of the MMI model.

3.5 Ethical approval

This study received a favourable ethical opinion from the university’s Research Ethics Committee (EC/2011/12/18/FHMS).

4. RESULTS
Participant demographics

Across professional groups, all participants where English was not their first language had IELTS (International English Language Test System) scores of seven or above demonstrating proficiency in English language. 76% percent of participants were Caucasian British (71% adult, 75% child, 55% mental health nursing and 95% midwifery).

Participants demographic characteristics (age and academic entry level) as well as outcomes of year three (clinical practice and academic performance) assessment are presented in Table 2. There was a statistically significant difference in age and academic entry level between professions. Mental health nursing students had the highest median age of all disciplines 24 (interquartile range: 18 – 35).

Across all disciplines, no significant correlation was found between age and clinical practice performance (Pearson correlation coefficient $r = 0.08, p = 0.148$). Similarly, there were no statistically significant differences between academic entry level and clinical practice outcomes (independent samples t-test = 0.152, $p=ns$).

Differences in MMI scores, clinical practice and academic performance between disciplines.

MMI scores varied by discipline; child nursing and midwifery student MMI scores were higher than students studying adult and mental health nursing, $p <0.001$ (Table 2). There was no relationship between discipline and clinical performance ($p=0.385$) and academic performance ($p=0.124$). There was a similar trend to MMI scores evident with child nursing and midwifery students achieving higher academic performance scores than those on adult and mental health nursing programmes.

Correlation between MMI scores, clinical practice and academic performance

MMI scores were found to relate to clinical performance at the end of year three for adult, mental health and midwifery students (Table 3). The relationship was strongest (showing moderate, statistically significant) correlations for mental health (0.43 $p=0.01$) and midwifery students (0.32 ($p=0.038$). No correlations are reported in relation to academic performance across disciplines (Table 3).

Regression analysis

The relationships between MMI scores and clinical practice performance and academic performance were modelled (Table 4 and 5 respectively). Correlation of independent variables in the multiple models did not exceed 0.6 fulfilling the assumption of no collinearity.

A positive, statistically significant relationship was found between MMI scores and clinical practice performance. The simple and multiple regression coefficients were 0.07 ($p=0.005$) and 0.10($p<0.001$) respectively (Table 4). Figure 2 shows the univariate relationship in a scatter plot.

A positive, statistically significant relationship between MMI score and academic performance was found in a simple regression model (coefficient = 0.06, $p = 0.008$) (Table 5). The relationship between MMI score and academic performance became not significant in the multiple regression model. This is potentially due to the effect of discipline which had a statistically significant effect on MMIs. Child nursing and midwifery students had higher MMI scores ($p<0.001$) (as well as dissertation scores) than adult and mental health nursing. This is also clear on the scatter plot in Figure 3. The scores for child nursing and midwifery students cluster in the top right-hand corner of the plot.
5. DISCUSSION

This is the first evaluation of the predictive validity of MMIs in nursing and midwifery student selection showing that scores arising from robustly designed MMIs (XXXX, 2014, 2015) relate to future practice performance (independently assessed by clinical mentors).

The relationship reported between MMI score at the point of student selection and end of programme clinical performance (when age, academic entry level and discipline were accounted for as potential confounders) confirms that MMIs have value in identifying students most likely to offer higher standards of care in nursing and midwifery contexts.

The significant associations reported between MMI score and end of programme clinical practice performance in adult and mental health student cohorts in a six-station model and midwifery student selection in a seven-station model are consistent with previously published findings of participants’ performance at the end of year one of this study (XXXX, 2018). Systematic reviews evaluating the effectiveness of different MMI models outside the fields of nursing and midwifery (Pau et al. 2013, Rees et al. 2016) conclude 7 to 12 stations offer greatest predictive validity in clinical practice performance. Establishing optimal numbers of stations is essential for identifying those most likely to excel in clinical practice as well as streamlining processes thereby reducing costs. The findings in this study in relation to the predictive validity of the six-station model are therefore important and merit further research. The reason for a non-significant correlation between MMI score and clinical practice performance reported in the child nursing cohort where a seven station MMI was employed is unclear. The authors suggest a larger sample would offer additional insights.

The cross-discipline differences in MMI score and academic (but not clinical practice) performance are noteworthy. Variance in dissertation marks within disciplines could have impacted on the multiple regression model; in simple regression, a statistically significant relationship was found between MMI score and dissertation mark, but this effect did not remain significant in the multiple regression model. The authors suggest this can be attributed to the relative competitiveness of child and midwifery admissions which result in higher mean MMI scores. The University programme academic entry requirements were also higher in the September 2015 cohorts in child and midwifery programmes possibly explaining a correspondingly higher mean dissertation mark. In a nursing selection context, Traynor et al. (2017) concluded that MMIs showed better differentiation of more able students in terms of academic performance; conversely, Gale et al (2016) suggest MMIs are weakly predictive of academic success. Given the primary aim of MMIs is to assess values and attributes, it might be expected that predictive validity of MMIs in relation to academic performance is uncertain and further research is needed.

These novel findings progress the empirical evidence-base for MMIs in a nursing and midwifery student selection context. When considering the outcomes of this study, the authors also acknowledge the possible effects of curricula and pressured clinical practice environments on students’ programme performance over three years. While selection approaches like MMIs are designed to assess values and attributes, conservancy of such personal domains cannot and should not be assumed (XXXX et al. 2018). XXXX et al. (2017, 2018) conceptualised healthcare students’ expression of movement and change in their values as a ‘values journey’ during their education programmes. In their recent consensus statement Patterson et al. (2018) identify fundamental theoretical gaps in published evidence around selection methods/methodologies including MMIs across professions. Patterson et al. (2018) call for more sophisticated cross-disciplinary evaluation approaches to address current knowledge deficits. The authors plan further multi-method studies.
designed to explore students’ ‘values journey’ in the context of MMI scores and programme performance.

5.1 Study rigour

The Principal Investigator (PI) in this study was personal tutor to a number of the midwifery participants (10). This was anticipated at the commencement of the study. The PI was not involved in directly interviewing any of the midwifery applicants. The PI was unavoidably supervisor to eight midwifery participants for their dissertation; consequently, these students were excluded from the analysis to avoid any possible conflict of interest. The practice assessments were carried out in the clinical area, independently from academic staff at the university.

Two independent individuals checked the accuracy of a random sample of 40% of the data. 98% accuracy was reported in the first check and 100% accuracy in the second. Therefore, no additional data checking was considered necessary.

5.2 Limitations and recommendations

This study is limited by its single case study design. It is not possible to determine how representative the participants were of the total student cohort due to ethical restrictions on the use of data from non-participants. A mixed-methods in depth case study approach would be required to fully unpick the complex relationship between individual performance of students selected using MMIs and work environmental/organisational factors which might impact on their programme performance.

Mentor grading is used as a proxy indicator for clinical practice performance. The authors acknowledge the potential limitations associated with the accuracy of mentor grading (Fisher et al, 2017). To mitigate this, mentors attended regular ‘mentor update’ meetings at their Hospital sites as well as having an annual snapshot of their marks moderated by an external tutor from the university. The complexity associated with clinical practice performance assessment using the rubric is also acknowledged. While it was informed by the UK Nursing and Midwifery Council criteria (NMC, 2009, 2010) there is a lack of universally understood definitions of competence and commonly accepted assessment measures. Therefore, potential inbuilt biases and assumptions may remain.

Dissertation mark was used in the study as a measure of academic performance. This was because it was the only theoretical assessment consistently made across disciplines at this university in the third year.

6. Conclusion

This preliminary evidence supports the potential of MMIs to identify those most likely to excel in clinical practice in the selection of students to nursing and midwifery programmes when six and seven stations are employed. These are the first such cross-discipline findings to be reported thereby strengthening the published evidence base in this context.

Selection methodologies and methods should be as robust as possible in identifying ‘the right’ applicants. Those involved in these processes have a responsibility to adopt evidence-based approaches appropriate to their discipline. In the future, innovative, multi-site, multi-method MMI evaluation studies are planned specifically in nursing and midwifery student selection context, to progressively strengthen our understanding of complex interview processes and how these relate to subsequent care provision.
Further details regarding MMI station content are withheld due to potential issues with test security. Please contact XXXX for further details.

References


Reiter, H., Eva, K. 2018. Vive la difference: the freedom and inherent responsibilities when designing and implementing multiple mini interviews. Academic Medicine, 93, 7.


