Effectiveness of Examiner Training in Preparing Medical Students as Mock OSCE Examiners

LEE KONG CHIAN SCHOOL OF MEDICINE

NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

Imperial College London

YT Yap[1], JH Koo[1], KY Ong[1], YQ Goh[1], SHW Neo[1], KY Tham[1,2]
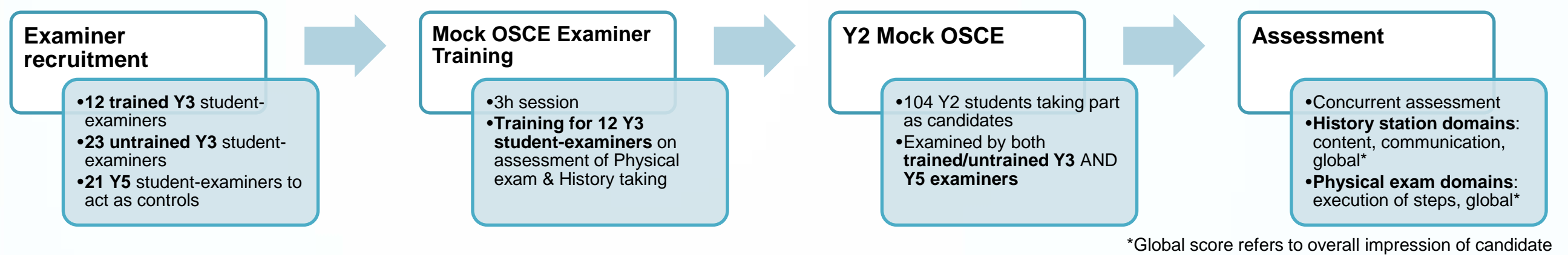1 Lee Kong Chian School of Medicine, 2 Tan Tock Seng Hospital

Introduction

• Mock OSCEs are simulations of clinical exam situations with simulated patients and senior students as examiners.

• Previous studies concluded that student-examiners are a viable alternative to faculty examiners in OSCEs, but there are mixed reports on the accuracy of peer assessment. There is also limited research about formal attempts to develop the teaching skills of medical students.

Research Questions

• Are Year 3 (Y3) student-examiners as effective as final Year 5 (Y5) student-examiners in assessing Year 2 (Y2) candidates in a Mock OSCE?

• Is a student-examiner training programme for Y3 student-examiners effective?

Methods

• Trained and untrained Y3 and Y5 student-examiners were recruited to assess Y2 candidates in a student-led Mock OSCE

Examiner recruitment
- **12 trained Y3** student-examiners
- **23 untrained Y3** student-examiners
- 21 **Y5** student-examiners to act as controls

→

Mock OSCE Examiner Training
- 3h session
- **Training for 12 Y3 student-examiners** on assessment of Physical exam & History taking

→

Y2 Mock OSCE
- 104 Y2 students taking part as candidates
- Examined by both **trained/untrained Y3 AND Y5 examiners**

→

Assessment
- Concurrent assessment
- **History station domains**: content, communication, global*
- **Physical exam domains**: execution of steps, global*

*Global score refers to overall impression of candidate

Results and Discussion

Table 1. Comparison of Y3 vs Y5 examiners for physical exam stations

Intraclass Correlation Coefficients	Execution	Global
All Y3 vs Y5 examiners (N=66)	0.903	0.750
Trained Y3 vs Y5 examiners (N=26)	0.961	0.819
Untrained Y3 vs Y5 examiners (N=40)	0.851	0.718

<0.4 = poor agreement; 0.4-0.75 = moderate agreement; >0.75 = excellent agreement

P-value of mean scores	Execution	Global
All Y3 vs Y5 examiners	0.816	0.906
Trained Y3 vs Y5 examiners	0.774	0.418
Untrained Y3 vs Y5 examiners	0.552	0.271

Independent samples t-tests, *p<0.05

Table 2. Comparison of Y3 vs Y5 examiners for history taking stations

Intraclass Correlation Coefficients	Content	Communication	Global
All Y3 vs Y5 examiners (N=108)	0.864	0.780	0.730
Trained Y3 vs Y5 examiners (N=53)	0.846	0.820	0.771
Untrained Y3 vs Y5 examiners (N=55)	0.876	0.771	0.718

P-value of mean scores	Content	Communication	Global
All Y3 vs Y5 examiners	0.230	0.036*	0.121
Trained Y3 vs Y5 examiners	0.064	0.012*	0.024*
Untrained Y3 vs Y5 examiners	0.999	0.609	0.902

Physical Exam

• Y3s were as effective as Y5 student-examiners for both execution and global domains, as shown by good correlation and no significant difference in mean scores.

• Training further improved reliability of examiners, with highest correlation between trained Y3 and Y5 examiners.

• Physical exam stations are less susceptible than history taking stations to bias as steps are prescriptive. Even so, there is room for future training programmes to emphasize on physical exam assessment and formulating objective and well-structured mark schemes.

History Taking

• Compared to physical exam stations, there is lower correlation and there are significant differences between the scores of Y3 and Y5 student-examiners.

• However, examiner training resulted in better correlation between Y3 student-examiners and the more experienced Y5 student-examiners in both the communication and global domains. This indicates that examiner training was effective in improving assessment of these domains.

• This could be because history taking and clinical communications are more susceptible to bias, hence training is needed to improve judgement and possibly increase reliability of student-examiners.

• Significant differences in the scores of trained Y3 and Y5 student-examiners despite a higher correlation could be explained by outliers in the data set.

• Lower correlations between the trained Y3 and Y5 student-examiners in the content domain could be explained by a lack of focus on content during training. Future training should also focus on the assessment of content of history taking stations.

Conclusion

• Y3 student-examiners are mostly effective substitutes for Y5 student-examiners in a Y2 Mock OSCE.

• Examiner training improved examiner reliability and effectiveness in physical exam stations as well as communication and global domains of history taking stations. Training should be targeted at this to reduce bias and improve global judgement of examiners.

Limitations

• Our study was conducted in the context of a pre-clinical OSCE and findings should be cautiously applied to clinical OSCEs where candidates perform physical examinations and take histories from real patients.