

# A Comparison of the Objective Structured Clinical Examination Results across Campuses of The University of the West Indies (2001 and 2002)

FW Hickling<sup>1</sup>, KAD Morgan<sup>1</sup>, W Abel<sup>1</sup>, CE Denbow<sup>2</sup>, Z Ali<sup>3</sup>, GD Nicholson<sup>4</sup>, C Siquee<sup>5</sup>

## ABSTRACT

**Objective:** To compare the performance of medical students in the Objective Structured Clinical Examination (OSCE) of the final MBBS Examination across the four campuses of The University of the West Indies, over a two-year period.

**Design and Methods:** All final examination results of the Medicine and Therapeutics OSCE were collected from the Faculty of Medical Science at the four campuses of The University of the West Indies and analyzed using both parametric (*t*-tests and ANOVAs) and non-parametric tests (chi-squared tests).

**Results:** Results indicated that students achieved significantly higher mean scores in the 2002 examination than in 2001 ( $t = 3.85$ ,  $df = 415$ ,  $p = 0.000$ ). There were no significant differences between campuses with regards to the mean corrected score in 2001. Also in 2001, in adult stations, all campuses achieved significantly higher scores than Jamaica. However, in Jamaica, mean child health station scores were significantly higher than all other campuses and, the mean score in Trinidad and Tobago was higher than the Bahamas and Barbados. In 2002, all other campuses achieved significantly higher scores than Trinidad and Tobago and females performed significantly better than males with regards to overall mean scores ( $t = 2.814$ ,  $df = 189$ ,  $p = 0.005$ ). Also in 2002, Barbados achieved significantly higher mean corrected scores than Trinidad and Tobago ( $F = 4.649$ ,  $df = 3,191$ ;  $p = 0.004$ ) and Barbados and Trinidad and Tobago both obtained significantly higher mean child health station scores than Jamaica.

**Conclusions:** The important conclusion from this study is that the OSCE scores in Medicine and Therapeutics are generally uniform across the four campuses of the University, thereby confirming the consistency of the approach to teaching and helping to validate the efficacy and veracity of the medical graduate being produced by The University of the West Indies.

# Una Comparación del Examen Clínico Objetivo Estructurado Resultados en los Campus de la Universidad de West Indies (2001 y 2002)

FW Hickling<sup>1</sup>, KAD Morgan<sup>1</sup>, W Abel<sup>1</sup>, C Denbow<sup>2</sup>, Z Ali<sup>3</sup>, G Nicholson<sup>4</sup>, C Siquee<sup>5</sup>

## RESUMEN

**Objetivo:** Comparar el rendimiento académico de los estudiantes de medicina en el examen clínico objetivo estructurado del examen final de MBBS, en los cuatro campus de La Universidad de West Indies, en un período de dos años.

**Diseño y métodos:** Se recopilaron todos los resultados del examen final de ECOE de Medicina y Terapéutica de la Facultad de Ciencias Médicas, en los cuatro campus de la Universidad de West Indies. Los datos fueron analizados usando tanto tests paramétricos (tests *t* y ANOVAs) como tests no paramétricos (tests de chi-cuadrado).

From: Departments of Community Health and Psychiatry<sup>1</sup> and Medicine<sup>2</sup>, The University of the West, Kingston 7, Jamaica West Indies; Department of Paediatrics<sup>3</sup>, The University of the West Indies, St Augustine, Trinidad and Tobago; Department of Medicine<sup>4</sup>, The University of the West Indies, Cave

Hill, Barbados and Department of Paediatrics<sup>5</sup>, Princess Margaret Hospital, Nassau, Bahamas.

Correspondence: Professor FW Hickling, Department of Community Health and Psychiatry, The University of the West Indies, Kingston 7, Jamaica, West Indies. Fax: (876) 927-2116, e-mail: frederick.hickling@uwimona.edu.jm

**Resultados:** Los resultados indicaron que los estudiantes alcanzaron puntuaciones significativamente más altas en el examen del 2001 que en el del 2002 ( $t = 3.85$ ,  $df = 415$ ,  $p = .000$ ). No hubo diferencias significativas entre los distintos campus con respecto a los resultados corregidos promedios 2001. También en 2001, en las unidades asistenciales de adultos, todos los campus lograron resultados significativamente más altos que Jamaica. Sin embargo, en Jamaica, los resultados promedios de las unidades pediátricas fueron significativamente más altos que en todos los otros campus, y el resultado promedio en Trinidad y Tobago fue más alto que en Bahamas y Barbados. En 2002, todos los otros campus lograron resultados significativamente más altos que Trinidad y Tobago, y las mujeres obtuvieron rendimientos significativamente mejores que los de los hombres, con respecto a los resultados promedios generales ( $t = 2.814$ ,  $df = 189$ ,  $p = .005$ ). También en 2002, Barbados alcanzó resultados corregidos promedios significativamente más altos que Trinidad y Tobago ( $F = 4.649$ ,  $df = 3,191$ ;  $p = .004$ ), mientras que Barbados así como Trinidad y Tobago, obtuvieron resultados significativamente más altos en la unidades pediátricas, en comparación con Jamaica.

**Conclusiones:** La conclusión principal de este estudio es que los resultados del OSCE en Medicina y Terapéutica son generalmente uniformes en los cuatro campus de la Universidad, confirmando de ese modo la solidez del enfoque de la enseñanza, y contribuyendo a validar la eficacia y calidad del graduado de medicina egresado de la Universidad de West Indies.

West Indian Med J 2005; 54 (2): 140

## INTRODUCTION

The Objective Structured Clinical Examination (OSCE) has been widely used, in Germany, Japan, the United Kingdom and North America for the purpose of evaluating clinical competence of nurses, medical students and residents at various levels of training (1). Students are evaluated in a number of skills, which include history taking, communication skills and the ability to undertake practical procedures. The validity and reliability of the OSCE have been confirmed and extensively reported in the literature (1–4). Nevertheless, there has been some controversy surrounding the issue of its predictive validity due to the fact that it provides an indirect measure of performance because it involves virtual and not real clinical situations (5). In addition to this, there seems to be high content-specificity of the clinical skills being evaluated, therefore students show great variation in performance depending on the assigned case or task (6).

In 2001, the Faculty of Medical Sciences, The University of the West Indies introduced the OSCE format in the clinical evaluation of students in the Medicine and Therapeutics examination of the Final Bachelor of Medicine, Bachelor of Surgery (MBBS) programme. The OSCE format of the Medicine and Therapeutics examination consists of 21 to 25 clinical stations. These encompass several disciplines, such as: internal medicine, paediatrics/child health, dermatology, community health and psychiatry. Local and external examiners supervise and evaluate the examination process. The OSCE format was implemented across all four campuses (Mona, Jamaica; Cave Hill, Barbados; St Augustine, Trinidad; and Nassau, The Bahamas). Teelucksingh *et al* have since shown that the process was perceived as being fairer to the students (7). The first examination was held in November/December 2000. In an attempt to standardize the

process and thereby the results, the examiners from all campuses are present at each campus during the examination period.

In 2001, the internal and external examiners of the Medicine and Therapeutics Examination mandated that the results of the examinations of all the campuses be collated and analyzed. The objective of this study was to compare the mean scores between the first two years of the examination, the pass rates among the four campuses, the mean scores of adult and child health stations across the four campuses, and the mean scores by gender. The hypothesis advanced is that there are no significant differences between pass rates across campuses, stations or between the years 2001 and 2002.

## METHOD

### Sample

All final examination results of the Medicine and Therapeutics OSCE were collected from the Faculty of Medical Sciences at the four campuses, in addition to relevant demographic data, and were entered into the Statistical Programme for the Social Sciences (SPSS), version 10. There were a total of 215 students who took the examination in 2001 and 195 (100 males and 95 females) in 2002. In 2001, there were 22 stations in Trinidad and Tobago, 23 stations in The Bahamas and Barbados, and 16 stations in Jamaica. While in 2002, there were 17 stations in Trinidad and Tobago, The Bahamas and Jamaica and 18 stations in Barbados.

### Procedure

The minimum competency scores were pre-determined by the modified Angoff technique (8) and obtained for all stations on each of the four campuses. The total raw score achieved by each candidate for all the stations was corrected to a passing score of 100 by using the total minimal com-

petency score for the examination. This pass mark was utilized on all four campuses. The scores for each station were entered into a spreadsheet database.

Cross-tabulations were made among the four campuses and their respective mean corrected scores. In addition, pass rates by campus for each station were collated, and mean scores by campus for the combined child health and adult stations were aggregated. The mean corrected score for each station in a sub-specialty, using the same technique applied to the overall mean corrected score, was added to obtain this aggregate score. In the 2002 examination, a communication skills station developed by the Department of Community Health and Psychiatry was introduced into the process and this was combined with the psychiatry station and compared across campuses. A one-way analysis of variance (ANOVA) was conducted on the data regarding mean corrected scores due to the quantitative interval nature of the data. A chi-square test for independence was further utilized to analyze the remaining factors due to the nominal/categorical characteristics of both the independent (campus) and dependent variables (pass rates). In addition to this, a pass rate by gender analysis was done via chi-square.

An independent samples t-test was also performed in order to compare the average corrected scores, child health and adult medicine scores between the years 2001 and 2002. Because of the low numbers of candidates in The Bahamas, the results from The Bahamas were combined with those from Trinidad and Tobago to facilitate the statistical analysis. Also, The University of the West Indies considers both these campuses as a single unit from an administrative perspective.

## RESULTS

### OSCE (2001)

Results indicated that there were no significant differences between campuses with regards to the mean corrected score. In a comparison of the performance in the child health stations across campuses, mean scores in Jamaica ( $x = 109.41$ ) were significantly higher ( $p = 0.000$ ) than all other campuses while, the average overall score in Trinidad ( $x = 107.69$ ) was higher than the Bahamas ( $x = 104.85$ ) and Barbados ( $x = 102.95$ ) (Table 1). However, in the adult stations, all other campuses achieved significantly higher scores than Jamaica ( $x = 53.11$ ,  $p = 0.000$ ). In the psychiatry station, a significant difference was noted in pass rates across campuses ( $p = 0.032$ ).

Statistically significant differences were found in nine stations (Table 2).

### OSCE (2002)

A one-way analysis of variance indicated that there was a significant difference between overall means by campuses (Table 3). The post-hoc analysis revealed that Barbados ( $x = 117.62$ ) achieved significantly higher scores than Trinidad and Tobago ( $x = 110.08$ ,  $p = 0.004$ ). No other comparisons were significant.

Table 1: Total pass rates, average paediatrics and adult scores across campuses (2001) and distribution of scores

Campus	Pass	n	Mean corrected	Standard deviation	Child health	Adult
Jamaica	68 (81%)	84	109.41	14.55	48.96	53.11
Trinidad	73 (75%)	98	107.69	11.84	38.71	63.88
Bahamas	8 (80%)	10	104.85	5.90	33.65	65.78
Barbados	19 (83%)	23	102.95	23.54	33.96	67.6
<b>Total</b>	<b>175 (79%)</b>	<b>215</b>	<b>107.70</b>	<b>14.42</b>		

Table 2: Pass rates by stations across campuses (2001)

Station	Bahamas	Barbados	Trinidad and Tobago	Jamaica	$\chi^2$	p
	Pass % (n)	Pass % (n)	Pass % (n)	Pass % (n)		
Paed tech	100% (10)	65% (15)	65% (64)	74% (62)	1.07	0.585
Paed abd	50% (5)	65% (15)	78% (76)	71% (60)	1.15	0.563
Paed cvs	40% (4)	87% (20)	79% (77)	58% (49)	10.22	0.006
Paed cns	100 (10)	78% (18)	88% (86)	73% (61)	8.34	0.015
Paed data	N/A	18% (4)	N/A	63% (53)	14.15	0.000
Paed growth	20% (2)	N/A	N/A	74% (62)		
Paed hist	80% (8)	48% (11)	63% (62)	74% (62)	5.87	0.053
Paed lab	40% (4)	52% (12)	78% (76)	92% (77)	19.27	0.000
Derm	30% (3)	64% (15)	64% (63)	52% (44)	1.92	0.383
X-ray	100% (10)	43% (10)	49% (48)	N/A	0.77	0.38
Fundus	60% (6)	78% (18)	61% (60)	N/A	2.6	0.107
CNS	N/A	N/A	N/A	64% (54)	N/A	N/A
History	60% (6)	74% (17)	79% (77)	86% (72)	2.82	0.245
Psychiatry	40% (4)	65% (15)	36% (35)	37% (31)	6.91	0.032
Rheum	70% (7)	91% (21)	69% (68)	83% (70)	8.36	0.015
Cranial nrv	100% (10)	96% (22)	62% (61)	N/A	8.62	0.003
Respiratory	30% (3)	91% (21)	81% (80)	79% (66)	2.57	0.277
Abdomen	60% (6)	87% (20)	41% (40)	55% (46)	15.23	0.000
Thyroid	90% (9)	96% (22)	68% (67)	–	6.53	0.011
CVS	80% (8)	83% (19)	52% (51)	62% (52)	6.53	.038
CNS motor	100% (10)	70% (16)	84% (82)	–	3.19	.074
Biochem	90% (9)	22% (5)	30% (29)	–	1.55	.214
Haem	60% (6)	44% (10)	85% (83)	–	15.9	.000
Sensation	60% (6)	44% (10)	35% (34)	–	3.48	.062
Snellen	60% (6)	70% (16)	47% (46)	–	.358	.55

\* NB Bahamas and Trinidad and Tobago were combined to perform the actual analysis

\* Rows in italics have cells with expected frequencies < 5. Results are not 100% reliable.

Results further indicated that Barbados ( $x = 45.12$ ) and Trinidad and Tobago ( $x = 42.71$ ) are both obtaining significantly higher scores in child health than Jamaica ( $x = 40.76$ ) ( $p = 0.001$ ) while, in adult stations, all campuses achieved significantly higher scores than Trinidad and Tobago ( $p = 0.000$ ), but the remaining three campuses were not statistically different from each other (Table 3).

No significant differences were found between campuses with regards to performance in Community Health and Psychiatry when the two stations were combined. Results also showed that females in general performed significantly

Table 3: Total pass rates, average paediatrics and adult scores across campuses (2002), and distribution of scores

Campus	Pass	n	Mean corrected	Standard deviation	Child health	Adult
Jamaica	73 (95%)	77	113.58	9.22	40.76	54.85
Trinidad	77 (83%)	93	110.08	12.82	42.71	60.31
Bahamas	7 (100%)	7	118.82	7.68	43.58	60.87
Barbados	16 (89%)	18	117.62	10.09	45.12	62.4
<b>Total</b>	<b>173 (89%)</b>	<b>195</b>	<b>112.48</b>	<b>10.24</b>		

better than males ( $p = 0.005$ ) with regards to mean corrected scores.

Many stations produced data that when subjected to statistical analysis had expected frequencies of less than 5, making those results not 100% reliable. Reliable statistically significant differences were found within the paediatric dysmorphic, the adult cardiovascular and the community health stations (Table 4).

Table 4: Pass rates by stations across campuses (2002)

Station	Bahamas	Barbados	Trinidad and Tobago	Jamaica	$\chi^2$	p
	Pass % (n)	Pass % (n)	Pass % (n)	Pass % (n)		
Paed H <sub>x</sub>	100% (7)	78% (14)	61% (57)	55% (42)	3.85	0.146
Paed cvs	71% (5)	83% (15)	85% (79)	65% (50)	9.3	0.01
Paed CNS	71% (5)	89% (16)	67% (62)	88% (68)	12.8	0.002
Paed tech	100% (7)	100% (18)	95% (88)	70% (54)	25.2	0.000
Paed Dys	71% (5)	100% (18)	37% (34)	53% (41)	22.8	0.000
Paed data	100% (7)	89% (16)	90% (84)	96% (74)	2.16	0.34
Ad Hist	57% (4)	67% (12)	67% (62)	90% (69)	13.9	0.001
Ad CVS	100% (7)	89% (16)	57% (53)	74% (57)	7.86	0.02
Ad CNS	86% (6)	61% (11)	57% (53)	90% (69)	20.9	0.002
Neuro	71% (5)	67% (12)	N/A	N/A	–	–
Ad X-ray	100% (7)	78% (14)	66% (61)	67% (52)	0.762	0.683
Derm	100% (7)	83% (15)	89% (83)	78% (60)	4.9	0.086
Rheum	71% (5)	83% (15)	80% (74)	77% (59)	0.424	0.809
Abdomen	86% (6)	83% (15)	68% (63)	86% (66)	7.31	0.026
Respiratory	100% (7)	94% (17)	72% (67)	91% (70)	10.6	0.005
Comm Hlt	57% (4)	50% (9)	56% (52)	75% (58)	8.43	0.015
Psychiatry	86% (6)	72% (13)	87% (81)	84% (65)	2.56	0.278
Fundus	N/A	100% (18)	81% (75)	78% (60)	5.84	0.211

\* Rows in italics have cells with expected frequencies < 5. Results are not 100% reliable

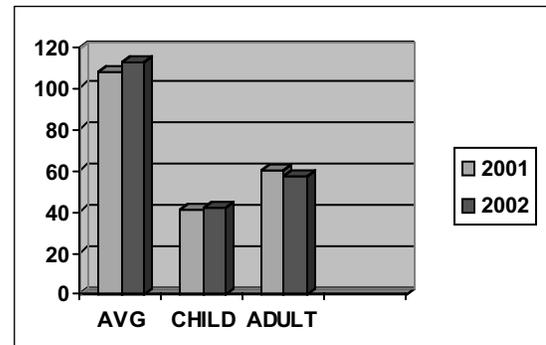
### OSCE 2001 vs 2002

Results indicated that students achieved significantly higher mean scores in the 2002 examination than in 2001 ( $t = 3.85$ ,  $p = 0.000$ ). In addition, although the mean score for child health was higher in 2002, the difference was not statistically significant. However, with regards to the adult stations, scores in 2001 were significantly higher than those in 2002 ( $t = 3.23$ ,

$p = 0.001$ ) (Table 5). All campuses registered an increase in overall mean scores from 2001 to 2002.

Table 5: Comparison of OSCE 2001 and 2002 mean corrected scores, child health and adult medicine stations

OSCE	Mean corrected	Child health	Adult
2001	107.70	41.88	60.26
2002	112.48	42.19	57.83



### DISCUSSION

There was general improvement from Year 1 to Year 2 in the mean scores of the OSCE, however, this did not hold true for the adult medicine stations which saw a decline in scores from Year 1 to Year 2. Nevertheless, the child health stations saw a slight increase, though not a statistically significant one.

This is only the second year since The University of the West Indies has implemented the OSCE as part of the MBBS final examination. Teelucksingh and his colleagues (7) have put to rest any speculation and unease about the veracity of the process in the Faculty of Medical Sciences at The University of the West Indies. The highly quantifiable nature of the scoring process of the OSCE has opened up important avenues for critical evaluation of the performance of the students across the campuses, within each campus, within each medical specialty, and indeed within each station. Of course, this puts the veracity of the process and efficacy of the actual teaching process in medicine, and indeed the clinical examination skills of the teachers of medicine in every campus under speculation. The process provides an important method for testing the quality of medical graduate produced in the Caribbean, and provides an important tool for quality assessment for the Faculty of Medical Sciences in The University of the West Indies.

The important conclusion from this study is that the OSCE scores in Medicine and Therapeutics are generally uniform across the four campuses of the University, thereby confirming the efficacy and veracity of the teaching process and by extension, the uniform standard of the medical

graduate being produced by The University of the West Indies. There are minor variations across campuses and within campuses that would be expected, and a few significant variations that require discussion and discourse for adequate explanation. A number of changes which occurred in the examination process from Year 1 to Year 2 need to be taken into consideration. Some of those include: the fact that stations were uneven, there was greater familiarity with the OSCE process for both students and examiners, minimal competencies were changed and each campus maintained its own style of administration.

The only significant difference across campuses was seen in 2002 where the candidates from Barbados achieved higher mean scores than those from Trinidad and Tobago. A number of different factors may have accounted for the statistical differences between these campuses. Some of these may have included inter-examiner reliability, varying teaching practices, number of repeaters, student preparation for the OSCE and entry qualifications into the medical programme. The possibility also exists that the position in the sequence in which each campus is examined in the three weeks of examinations across the Caribbean may affect the performance of the campus. It is possible that the campus examined last may score higher grades as information about the examination circulates among students across the Caribbean.

The results also suggest that students at the Mona campus seem to have achieved significantly higher scores in the child health stations in 2001, but performed worse across the campuses in 2002. With regards to the adult stations, Mona, Jamaica, again changed its positioning by one incre-

ment, moving from the lowest average score in 2001 to the 2<sup>nd</sup> lowest score in 2002. There were also significant differences in the performance of the OSCE by gender, with females performing better than males. Worldwide trends have shown that females in medical school tend to do better than males and the results of this study certainly support this contention. The results of this study must at best be seen as preliminary in nature, and an attempt to provide a baseline for future evaluation as the OSCE process is refined and strengthened in The Faculty of Medical Sciences.

## REFERENCES

1. Carraccio C, Englander R. The objective structured clinical examination: a step in the direction of competency-based evaluation. *Arch Pediatr Adolesc Med* 2000; **154**: 736–41.
2. Mavis BE, Henry RC, Ogle KC, Hoppe, RB. The emperor's new clothes: the OSCE reassessed. *Acad Med* 1996; **71**: 447–53.
3. Harden RM. What is an OSCE? *Med Teach* 1988; **10**: 19–22.
4. Woodburn J, Sutcliffe N. The reliability, validity, and evaluation of the objective structured clinical examinations in podiatry (chiroprody). *Assess Eval High Educ* 1996; **21**: 131–46.
5. Gorter S, Rethans J, van der Hejide D, Scherpbreer A, Houben H, van der Vleuten et al. Reproducibility of clinical performance assessment in practice using incognito standardized patients. *Med Ed* 2002; **36**: 827–32.
6. Kramer AW, Jansen JJ, Zuithoff P, Dusman H, Tan LH, Grol RP et al. Predictive validity of a written knowledge test of skills for an OSCE in postgraduate training for general practice. *Med Ed* 2002; **36**: 812–19.
7. Teelucksingh S, Ali Z, Fraser HS, Denbow CE, Nicholson GD. Fifty years of clinical examination at the University of the West Indies. *West Indian Med J* 2001; **50** (Suppl 4): 50–2.
8. Kaufman DM, Mann KV, Muijtjens AMM, Van der Vleuten CPM. A comparison of standard-setting procedures for an OSCE in undergraduate medical education. *Acad Med* 2000; **75**: 267–71.