

Correlation of Performance of Students in Comprehensive Examinations, General Weighted Average Grade to the Physician Licensure Examinations in a Medical School with an Organ System Integrated Curriculum

Marissa N. Valbuena, Teresita R. Castillo, Josefina I. Lapeña and Tita L. Dimaano

College of Medicine University of the Philippines Manila

ABSTRACT

Objectives

1. To describe the performance of medical students in comprehensive examinations given after each learning unit or year level.
2. To correlate the scores of the students in the comprehensive examinations to the general weighted average grade (GWAG) for the same learning unit.
3. To correlate the scores in the subject-based part of the comprehensive examinations to the scores in the Physicians Licensure Examination (PLE).

Methods. The scores of graduates of the University of the Philippines Manila College of Medicine (UPCM) for the years 2009-2011 in the comprehensive examinations, their general weighted average grade per learning unit, and scores in the PLE were retrieved and analyzed. Correlation coefficients of comprehensive examination score and GWAG, and comprehensive examination score and PLE score were computed.

Results. A total of 434 out of 463 graduates from the UPCM were included in this study. There was no trend observed in the performance of students in the comprehensive examinations as they progressed from one learning unit to the next. Performance in both portions of the comprehensive examinations was observed to have significant correlation to students' LU GWAG for most of the learning units, with the exception of LU 6 and LU 7 for Class B. Performance in the subject-based portion of the comprehensive examinations was observed to have significant correlation to students' PLE scores.

Conclusions. The students' scores in the comprehensive examinations had moderate to strong correlation with the GWAG for the corresponding learning unit. Similarly, scores in the subject-based portion of the comprehensive examination had moderate to strong correlation with PLE scores.

Key Words: comprehensive examinations, physicians licensure examinations, student performance, organ system integration, medical education

Corresponding author: Marissa N. Valbuena, MD, MHPEd
College of Medicine
University of the Philippines Manila
547 Pedro Gil St., PO Box 593, Ermita, Manila 1000 Philippines
Telephone: +632 5361392
Fax No: +632 5260371
Email: mnvalbuena@post.upm.edu.ph

Introduction

Filipino graduates of Philippine medical schools are required to take and pass the Physician Licensure Examinations (PLE) given by the Board of Medicine of the Professional Regulatory Commission (PRC) after a one-year internship training to be able to practice medicine as a profession in the Philippines.¹ The performance of the graduates of a medical school (passing rate) is often used as a yardstick of the quality of education provided by the school. The PLE is a multiple choice objective written examination on 12 subjects, namely, Biochemistry, Anatomy and Histology, Microbiology, Physiology, Legal Medicine, Pathology, Pharmacology and Therapeutics, Surgery, Orthopedics, Ophthalmology and Otorhinolaryngology, Medicine, Obstetrics and Gynecology, Pediatrics and Nutrition, and Preventive Medicine and Public Health.

In 2004, the University of the Philippines Manila College of Medicine (UPCM) implemented the organ system integrated (OSI) medical curriculum. The OSI curriculum consists of five learning units (LU) that revolves around a theme, as presented in Table 1.

Table 1. Description of Learning Units of OSI Medical Curriculum of the UP Manila, College of Medicine²

LEARNING UNIT	THEME
LU 3	Human development, structure and function
LU 4	Human pathophysiology and Therapeutics
LU 5	Ambulatory Medicine (Integrated Clinical Clerkship I)
LU 6	Hospital based Medicine (Integrated Clinical Clerkship II)
LU 7	Comprehensive Health Care (Internship)

During the pre-implementation phase of the new curriculum, the curriculum committee and the college administrators conceptualized the adoption of the comprehensive examination. The comprehensive examination is a four-hour written examination, consisting of 200 multiple choice questions given at the end of the second semester of each academic year. It consists of two parts, a subject-based examination and an integrative examination. The subject-based examination covered subjects given in the Physicians Licensure Examinations (PLE). The subject-based portion of the comprehensive

examination was intended to prepare the students for the PLE. The integrative portion was intended to serve as an assessment tool to evaluate how well students have integrated and combined knowledge and topics taken in various courses and disciplines. At the time this study was conducted, there had been three batches of graduates who had taken their yearly comprehensive examinations and the PLE.

One of the adages in medical education is that "assessment drives learning". For integrated learning to be achieved, there should be integrated assessment. Case-based assessment methods attempt to assess integration of students' knowledge. When multiple choice questions are focused on assessment of process of application, analysis, synthesis and application, they can provide opportunities for students to demonstrate they can integrate their knowledge.³

Comprehensive examinations assess student's knowledge from several or all courses within a period of study. It is sometimes referred to as a cumulative examination.⁴

It is the purpose of this study to assist school administrators in assessing the value of comprehensive examinations in evaluating preparedness of graduates for the PLE. The study also aimed to determine if any correlation exists between academic and comprehensive examination performance. Specifically, the objectives of this study were as follows: to describe performance of medical students in comprehensive examinations given after each learning unit or year level; to correlate the scores of the students in the comprehensive examinations to the general weighted average grade (GWAG) for the same learning unit; and to correlate the scores of the subject based part of the comprehensive examinations to the scores in the Physicians Licensure Examination (PLE).

Methods

Subjects consisted of students accepted to the University of the Philippines Manila College of Medicine (UPCM) who graduated in 2009 to 2011 who fulfilled the following criteria: (a) completed a Doctor of Medicine degree within the standard 5-year period since entrance into the program; (b) took all the comprehensive examinations with their class; and (c) took the Physicians Licensure Examinations (PLE) immediately upon completion of the program, in the same year they graduated. To maintain confidentiality of student data, class and student names and other identifiers were concealed prior to data analysis.

Data collected consisted of examination scores for each of the comprehensive examinations administered at the end of each academic year and the Physicians Licensure Examination (PLE). General weighted average of grades (GWAG) for each learning unit and Medical Proper GWAG were also tabulated for each subject. The GWAG is based on the grades obtained in each course and the credit hours for

the course. Scores and GWAG were subjected to descriptive statistical analysis. Analysis of variance (ANOVA) of all scores and GWAG across classes was performed to determine if the performance of subjects across the different classes was comparable. In order to determine the presence of any relationship between students' performance in the comprehensive examinations and their corresponding GWAG, Pearson's correlation coefficients (r) were computed using Microsoft Excel Data Analysis Toolpack. The relationship of performance in the year-end comprehensive examinations to performance in the PLE was likewise analyzed with Pearson's Correlation. Average correlation coefficients were determined by transforming Pearson's Correlation Coefficient (r) to Fisher Z values.

Results

Subjects. Of the total number of 463 students who graduated from the UPCM for the years 2009 to 2011, only 434 students satisfied the inclusion criteria for the study (Table 2).

Table 2. Number of graduates and subjects per class included in study

Class	Total No. of Graduates	No. Included in Study
A	158	149
B	154	147
C	151	138
TOTAL	463	434

Comprehensive Examinations. The performances of the students in the comprehensive examinations given at the end of each Learning Unit (LU) are summarized in Tables 3 and 4. The mean of scores across the three classes ranged from 48.75 (Class B, Subject-based Examination for LU 4) to 86.84 (Class B, Integrated Examination for LU 5). Class B student scores were noted to be higher in all integrative examinations when compared to their corresponding subject-based examinations. The opposite trend was observed for Class A and C where the student scores were higher in the majority of the subject-based examinations (Class A: LU 4, 6, 7 and Class C: LU 4, 7) in comparison to their corresponding integrated examinations. Note that an objective structured clinical examination (OSCE) is administered to LU 5 students instead of a subject-based examination.

While there was note of improvement of scores in the subject-based examinations for Class A as students progressed from one learning unit to the next, no trend was observed for the other two classes. Mean scores for the integrated comprehensive examinations did not exhibit any trend in any of the classes. Comparison of the mean scores for each learning unit using single factor analysis of variance (ANOVA) revealed that these differed significantly from class to class.

Table 3. Performance in Subject-based Comprehensive Examinations

BATCH	LEARNING UNIT	MEAN	STD DEVIATION	CONFIDENCE INTERVAL
A n = 149	3	52.74	7.82	51.48, 54.01
	4	59.29	6.60	58.22, 60.86
	5	-	-	-
	6	59.77	5.85	58.82, 60.72
	7	61.74	6.52	60.68, 62.79
B n = 147	3	56.71	8.17	55.38, 58.04
	4	48.75	6.27	47.72, 49.77
	5	-	-	-
	6	59.50	6.20	58.49, 60.51
	7	55.80	6.28	54.17, 56.82
C n = 138	3	-	-	-
	4	64.50	7.47	63.24, 65.76
	5	-	-	-
	6	52.37	5.64	51.42, 53.32
	7	62.48	7.32	61.25, 63.71

Table 4. Performance in Integrated Comprehensive Examinations

BATCH	LEARNING UNIT	MEAN	STD DEVIATION	CONFIDENCE INTERVAL
A n = 149	3	59.69	6.15	58.70, 60.69
	4	52.66	7.10	51.51, 53.81
	5	59.70	4.88	58.91, 60.50
	6	55.06	4.90	54.27, 55.85
	7	53.62	7.54	52.40, 54.84
B n = 147	3	61.13	7.50	59.91, 62.35
	4	68.15	8.08	66.84, 69.47
	5	86.84	7.06	85.69, 87.99
	6	61.65	6.18	60.65, 62.66
	7	56.35	6.71	55.25, 57.14
C n = 138	3	56.71	6.95	55.54, 57.88
	4	58.43	7.56	57.16, 59.70
	5	67.76	6.61	66.64, 68.87
	6	54.85	6.62	53.74, 55.97
	7	60.34	6.20	59.30, 61.38

Table 5. General Weighted Average of Grades (GWAG) for Medicine Proper and for each Learning Unit

BATCH	LEARNING UNIT	MEAN	STD DEVIATION	CONFIDENCE INTERVAL
A n = 149	3	2.19	0.22	2.15, 2.22
	4	2.28	0.25	2.24, 2.32
	5	1.85	0.10	1.84, 1.87
	6	1.80	0.13	1.78, 1.82
	7	1.68	0.16	1.66, 1.71
	Med Proper	2.00	0.15	1.98, 2.03
B n = 147	3	2.13	0.24	2.10, 2.17
	4	2.06	0.21	2.03, 2.10
	5	1.74	0.13	1.72, 1.76
	6	1.92	0.98	1.76, 2.08
	7	1.87	0.24	1.83, 1.91
	Med Proper	1.92	0.15	1.90, 1.94
C n = 138	3	2.13	0.25	2.09, 2.18
	4	2.12	0.26	2.07, 2.16
	5	1.82	0.13	1.80, 1.84
	6	1.85	0.14	1.83, 1.88
	7	1.87	0.28	1.82, 1.92
	Med Proper	1.96	0.17	1.93, 1.99

General Weighted Average of Grades. Mean GWAG for the various learning units and GWAG upon course completion of the three classes included in the study are listed in Table 5. The lowest mean for learning unit GWAG was observed for LU 4 of class A (mean=2.28, CI=2.24, 2.32) while the highest learning unit GWAG was observed for LU 7 of class A (mean=1.68, CI=1.66, 1.71). Note that for all classes included in the study, mean GWAGs were higher for LU 5, 6 and 7 compared with GWAGs for LU 3 and 4. Analysis of variance of GWAGs for each learning unit revealed that only the mean GWAGs for LU 3 and LU 6 were comparable across the three classes. Mean GWAGs for the other learning units significantly differed from each other. Upon comparing mean Med Proper GWAG for the different classes, ANOVA revealed that they differed significantly from class to class.

Performance in the Physicians Licensure Examinations. Mean PLE scores for Classes A, B and C were 82.18, 81.90 and 83.61, respectively (Table 6). ANOVA revealed that the scores from the three classes were statistically different from each other ($F=17.746$, $p<0.0001$, $F_{critical}=3.017$).

Table 6. Physician's Licensure Examination Scores

BATCH	n	MEAN	STD DEVIATION	CONFIDENCE INTERVAL
A	149	82.18	2.38	81.79, 82.56
B	147	81.90	2.84	81.43, 82.36
C	138	83.61	2.35	83.21, 84.00

Correlation between Comprehensive Exam Scores and LU GWAG. Performance in both portions of the comprehensive examinations was observed to have significant correlation to students' LU GWAG (Tables 7 and 8) for most of the learning units, with the exception of LU 6 and LU 7 for Class B. Note that in the grading system of the UPCM, the highest grade is 1.0 and the lowest passing grade is 3.0, conditional failure is 4.0 and failure is 5.0. The reported negative correlation coefficients would therefore imply that as LU GWAG decreases, scores in the comprehensive examinations correspondingly increases.

Statistically significant correlation coefficients between subject-based comprehensive examinations and corresponding GWAG for the learning units ranged from -0.330 to -0.769 with the strongest correlation observed for LU 6 of Class C and the weakest correlation for LU 7 of Class A. Note that no relation was established between scores and LU GWAG for LU 6 and LU 7 of Class B ($r=-0.096$, $p=0.119$ and $r=-0.094$, $p=0.128$, respectively).

Similarly, statistically significant correlation between integrated examinations and corresponding LU GWAG ranged from -0.443 to -0.766. Strongest correlation was observed between scores and GWAG of LU 6 of Class C and weakest correlation was observed between scores and GWAG of LU 7 Class A. As with the subject-based

examinations, no relationship was established between scores and LU GWAG for LU 6 ($r=-0.080$, $p=0.166$) and LU 7 ($r=-0.028$, $p=0.368$) of Class B.

Average correlation between examination scores of each LU and their corresponding GWAG for the different classes was calculated using the Fisher transformation of r to z which resulted in average correlation values presented in Tables 7 and 8. All averages calculated were significant for both portions of the comprehensive examinations. Further statistical calculations revealed that the correlation coefficients was not significantly different for the three classes for all the learning units (LU 3, 4, 6 and 7) of the subject-based examination group and LU 3 and 7 of the integrated examination group.

Table 7. Correlation between Subject-based Exam and LU GWAG

LU	Batch	n	r	z
3	A	149	-0.679 ^s	-0.827
	B	147	-0.614 ^s	-0.715
	Ave		-0.648	-0.772
4	A	149	-0.563 ^s	-0.637
	B	147	-0.581 ^s	-0.664
	C	138	-0.551 ^s	-0.620
Ave		-0.566	-0.641	
6	A	149	-0.733 ^s	-0.935
	B	147	-0.098 ^{ns(1)}	-0.098
	C	138	-0.769 ^s	-1.018
Ave		-0.590	-0.678	
7	A	149	-0.330 ^s	-0.343
	B	147	-0.094 ^{ns(2)}	-0.094
	C	138	-0.500 ^s	-0.549
Ave		-0.313	-0.324	

s- statistically significant, $p<0.001$, *ns(1)* – not significant, $p=0.119$, *ns(2)*-not significant, $p=0.128$

Table 8. Correlation between Integrative Exam and LU GWAG

LU	Batch	n	r	z
3	A	149	-0.599 ^s	-0.692
	B	147	-0.656 ^s	-0.786
	C	138	-0.704 ^s	-0.875
Ave		-0.654	-0.782	
4	A	149	-0.544 ^s	-0.610
	B	147	-0.520 ^s	-0.589
	C	138	-0.724 ^s	-0.916
Ave		-0.604	-0.700	
5	A	149	-0.573 ^s	-0.652
	B	147	-0.460 ^s	-0.497
	C	138	-0.684 ^s	-0.837
Ave		-0.577	-0.658	
6	A	149	-0.611 ^s	-0.711
	B	147	-0.080 ^{ns(1)}	-0.080
	C	138	-0.766 ^s	-1.011
Ave		-0.532	-0.593	
7	A	149	-0.443 ^s	-0.476
	B	147	-0.028 ^{ns(2)}	-0.028
	C	138	-0.510 ^s	-0.563
Ave		-0.338	-0.352	

s-statistically significant, $p<0.001$, *ns(1)* – not significant, $p=0.116$, *ns(2)*-not significant, $p=0.368$

Correlation between Comprehensive Exam Scores and PLE Scores. Performance in the subject-based portion of the comprehensive examinations was observed to have significant correlation to students’ PLE scores (Table 9). Correlations generally ranged from moderate to strong. An exception to this was correlation between scores in LU 7 of Class B where only a very weak correlation was observed.

Table 9. Correlation between Subject-based Exam and PLE

LU	Rank	n	r	z	X ²	p value
3	A	149	0.567 ^s	0.643		
	B	147	0.489 ^s	0.535		
	Ave		0.53	0.590	0.851	0.356
4	A	149	0.419 ^s	0.446		
	B	147	0.541 ^s	0.606		
	C	138	0.563 ^s	0.637		
Ave		0.509	0.561	2.985	0.225	
6	A	149	0.655 ^s	0.784		
	B	147	0.751 ^s	0.975		
	C	138	0.722 ^s	0.912		
Ave		0.711	0.889	2.751	0.253	
7	A	149	0.430 ^s	0.460		
	B	147	0.167 ^{ns}	0.169		
	C	138	0.701 ^s	0.869		
Ave		0.455	0.491	11.755	0.001	

S-statistically significant, $p<0.001$, *S**-statistically significant, $p=0.022$

With the exception of the correlation between integrated examination and PLE score for LU 7 of Class B, all other correlations between integrated examinations and PLE were statistically significant (Table 10). These correlations ranged from moderate to strong with the strongest correlation observed for LU 6, Class B.

Average correlation between examination scores of each LU and students’ PLE scores were likewise calculated using the Fisher transformation of r to z (Tables 9 and 10). Average correlations were significant for both portions of the comprehensive examinations. Significant differences in correlations between the different classes were observed for LU 7 of the subject-based examination group and LU 4 and 6 of the integrated examination group.

Discussion

The medical students’ scores for both portions of the comprehensive examination were observed to be low in all learning units, with the majority of mean scores falling below 60%. In addition to this, it was observed that there was no improvement of scores as students progressed from one learning unit to another with the exception of Class A in the subject-based comprehensive examinations. While adequacy of knowledge base is a major factor that can affect performance in examinations, other factors such as preparedness and motivation to perform well must also be considered. One possible cause for the poor performance of students in the comprehensive examinations is their lack of motivation to prepare for the said examination since their outcome will not affect their grades as this was not included

Table 10. Correlation between Integrated Exam and PLE

LU		n	r	z	X ²	p value
3	A	149	0.533 ^s	0.594		
	B	147	0.551 ^s	0.620		
	C	138	0.650 ^s	0.775		
	<i>Ave</i>		0.579	0.661	0.266	0.265
4	A	149	0.447 ^s	0.481		
	B	147	0.530 ^s	0.590		
	C	138	0.677 ^s	0.824		
	<i>Ave</i>		0.556	0.627	8.526	0.014
5	A	149	0.558 ^s	0.630		
	B	147	0.573 ^s	0.652		
	C	138	0.630 ^s	0.741		
	<i>Ave</i>		0.587	0.673	0.966	0.617
6	A	149	0.534 ^s	0.596		
	B	147	0.724 ^s	0.916		
	C	138	0.692 ^s	0.852		
	<i>Ave</i>		0.656	0.786	8.304	0.016
7	A	149	0.522 ^s	0.579		
	B	147	0.123 ^{NS}	0.124		
	C	138	0.605 ^s	0.701		
	<i>Ave</i>		0.433	0.464	1.042	0.307

S-statistically significant, $p < 0.001$, NS-not statistically significant, $p = 0.069$

in the computation of final grades for LU 3 to LU 5. Since the comprehensive examinations were given at the same time as the final examinations for individual subjects, students may have instead focused on studying for the final examinations as their outcome will affect their final grades. The lack of improvement in student performance in the comprehensive examinations may also be attributed to the variability in the construction of the examinations as these are made by different faculty committees.

Students who are good academically (higher GWAG in the LU) tend to perform comparatively better in the comprehensive examinations. Students with higher GWAGs understandably retained more knowledge compared with the rest of the students. The study of Wanvarie and Prakunhungsit⁵ among Thai medical students showed moderately strong correlation between academic performance and the step 1 Medical Licensing Examination of Thailand (MLET) given during the third year of medical school and a stronger correlation between the scores of comprehensive clinical science examination and the step 2 MLET. Baker et al.⁶ also found a strong association of academic performance in the first 2 years of osteopathic medical school and the performance in Comprehensive Osteopathic Medical Licensure Examination in the United States.

The observed positive correlation between the subject-based comprehensive examinations and the PLE may be explained in part by the similarity in the content and structure of the two examinations. What is disturbing is that despite the positive correlation, comprehensive examinations scores of students were generally low.

While the subject-based portion of the comprehensive examinations were intended to prepare students for the PLE

and assess final competencies of students, the examinations may also serve as a means to identify which students would require attention and assistance to ensure that they achieve competencies required for licensure examinations.

Conclusion

The students' scores in the comprehensive examinations had moderate to strong correlation with the GWAGs for the corresponding learning units. Similarly, scores in the subject-based portion of the comprehensive examination had moderate to strong correlation with PLE scores. Based on current results, it can be inferred that subject-based comprehensive examinations adequately prepare students for the PLE. Future studies can be undertaken to determine if performance in the comprehensive examination can actually be predictive of future performance in the PLE.

Acknowledgment

The authors thank Ms. Katherine Salazar of the Department of Clinical Epidemiology, College of Medicine, University of the Philippines Manila for her assistance in the statistical analysis.

References

1. Bellosillo JN, Castro BC, Mapili EL, et al. Basics of Philippine Medical Jurisprudence and Ethics. Quezon City, Philippines: Central Book Supply, Inc.; 2010. pp. 650-668.
2. Aguilin-Dalisay G, Cantiller JA, Yakat JA. The UP College of Medicine Organ System Integrated Curriculum: Evaluating First Returns. Pambansang Samahang Sikologiyang Pilipino, Nov 2009 (Unpublished).
3. Prideaux D. Integrated Learning. In: Dent JA, Harden RM, eds. A Practical Guide for Medical Teachers, 2nd ed. Edinburg: Elsevier; 2005. pp. 151-159.
4. Ryan GJ, Nykamp D. Use of cumulative examinations at U.S. Schools of Pharmacy. Am J Pharm Educ. 2000; 64:409-412.
5. Wanvarie S, Prakunhungsit S. Step-2 Thai Medical Licensing Examination result: a follow-up study. Ann Acad Med Singapore. 2008; 37(12):1024-6.
6. Baker HH, Foster RW, Bates BP, et al. Relationship between academic achievement and the COMLEX-USA Level 1 performance: a multisite study. J Am Osteopath Assoc. 2000; 100(4):238-42.