Relationship between Online Formative Assessment and Summative Assessment in the Pediatric Module among Fifth-Year Medical Students

Varisa Piriyakitphaiboon MD, MSc[1]
Thanyaros Sinsophonphap MD[1]
Ornatcha Sirimongkolchaiyakul MD[1*]

[1] Department of Pediatric, Faculty of Medicine Vajira Hospital, Navamindradhiraj University, Bangkok, Thailand
* Corresponding author, e-mail address: ornatcha@nmu.ac.th
Vajira Med J. 2022; 66(3): 181-8
http://dx.doi.org/10.14456/vmj.2022.18

Abstract

Objectives: The present study aimed to explore the relationship between online formative outcomes and summative performance in the pediatric module among fifth-year medical students.

Methods: We retrospectively collected and reviewed the learning data of 84 fifth-year medical students who enrolled in the pediatric module between August 2020 and July 2021.

Results: Students who received individualized feedback had a higher mean summative score than their counterparts (61.17 ± 9.5 vs. 59.24 ± 9.6), and their summative outcome was significantly higher than their formative outcome (61.17% vs. 47.9%). The online formative scores were not correlated with summative scores, but there appeared a strong positive correlation between cumulative grade point average (CGPA) and summative scores (r = 0.685, *p*-value = 0.001) and a weak negative correlation between pediatric clerkship evaluation and summative scores (r = -0.380, *p*-value = 0.004). Univariate analysis revealed that lower average cumulative grade point average was significantly associated with an unsatisfactory grade in the pediatric module.

Conclusion: The online formative assessment was not correlated with the summative assessment. Therefore, clinical teachers should consider other factors, especially CGPA to identify students who have a tendency for academic failure.

Keywords: formative assessment, medical education, pediatrics module, summative evaluation, feedback

Received: 30 December 2021, **Revised:** 11 April 2022, **Accepted:** 9 May 2022

ความสัมพันธ์ของการประเมินผลระหว่างเรียนแบบออนไลน์ต่อผลการเรียนวิชากุมารเวชศาสตร์ของนักศึกษาแพทย์ชั้นปีที่ 5

วริษา พิริยะกิจไพบูลย์ พ.บ.[1]
ธันยรส สินโสภณภาพ พ.บ.[1]
อรอัชฌา ศิริมงคลชัยกุล พ.บ.[1*]

[1] ภาควิชากุมารเวชศาสตร์ คณะแพทยศาสตร์วชิรพยาบาล มหาวิทยาลัยนวมินทราธิราช กรุงเทพมหานคร ประเทศไทย

* ผู้ติดต่อ, อีเมล: ornatcha@nmu.ac.th

Vajira Med J. 2022; 66(3): 181-8
http://dx.doi.org/10.14456/vmj.2022.18

บทคัดย่อ

วัตถุประสงค์: เพื่อศึกษาความสัมพันธ์ของการประเมินผลระหว่างเรียนแบบออนไลน์ต่อการประเมินผลสรุปในวิชากุมารเวชศาสตร์ของนักศึกษาแพทย์ชั้นปีที่ 5

วิธีดำเนินการวิจัย: การศึกษาวิจัยชนิดย้อนกลับโดยการเก็บรวบรวมและวิเคราะห์ข้อมูลผลการเรียนของนักศึกษาแพทย์ชั้นปีที่ 5 ที่ลงทะเบียนเรียนในภาควิชากุมารเวชศาสตร์ระหว่างสิงหาคม 2563 ถึง กรกฎาคม 2564 จำนวน 84 คน

ผลการวิจัย: นักศึกษาแพทย์ที่ได้รับข้อมูลป้อนกลับเฉพาะบุคคลมีค่าเฉลี่ยของผลคะแนนประเมินสูงกว่าค่าเฉลี่ยคะแนนของนักเรียนทั้งหมด (61.17 ± 9.5 vs. 59.24 ± 9.6) และยังมีค่าเฉลี่ยของคะแนนการประเมินผลระหว่างเรียนสูงกว่าคะแนนประเมินผลหลังเรียนอย่างมีนัยสำคัญ (61.17% vs. 47.9%) แต่ไม่พบความสัมพันธ์ของคะแนนประเมินระหว่างเรียนแบบออนไลน์กับคะแนนประเมินหลังเรียน อย่างไรก็ตามพบว่าเกรดเฉลี่ยสะสมมีความสัมพันธ์เชิงบวกระดับสูงกับคะแนนประเมินหลังเรียน (r = 0.685, p-value = 0.001) และคะแนนประเมินด้านการปฏิบัติงานสัมพันธ์เชิงลบกับคะแนนประเมินหลังเรียน (r = -0.380, p-value = 0.004) จากการศึกษาครั้งนี้พบว่ามีเพียงเกรดเฉลี่ยสะสมที่เป็นปัจจัยสัมพันธ์กับเกรดของวิชากุมารเวชศาสตร์

สรุปผลการวิจัย: การศึกษาครั้งนี้ไม่พบความสัมพันธ์ระหว่างผลการประเมินระหว่างเรียนและผลการประเมินหลังเรียน ดังนั้น อาจารย์ทางคลินิกควรพิจารณาถึงปัจจัยอื่นโดยเฉพาะอย่างยิ่งเกรดเฉลี่ยสะสม เพื่อช่วยบอกแนวโน้มนักเรียนที่อาจไม่ประสบความสำเร็จในการเรียน

คำสำคัญ: การประเมินผลระหว่างเรียน แพทยศาสตรศึกษา กุมารเวชศาสตร์ การประเมินผลหลังเรียน การให้ข้อมูลสะท้อนกลับ

วันที่รับบทความ 30 ธันวาคม 2564 **วันแก้ไขบทความ** 11 เมษายน 2565 **วันตอบรับบทความ** 9 พฤษภาคม 2565

Introduction

Formative assessment is a part of teaching strategies that aim to improve the communication between teachers and students and enhance academic achievement[1]. This type of assessment benefits both teachers and students. To an extent, teachers can provide informative feedback to clarify the learning objectives and key elements for learning achievement, and students can use this feedback to gain a deeper understanding and plan their own revision[2]. Teachers can provide feedback in various forms, such as computer-based feedback or assessment and verbal communication. Although emerging evidence suggests that feedback improves learning outcomes, the best manner to offer feedback has remained controversial[3].

Online learning has been increasingly associated with medical education in various ways, particularly online formative assessment. This approach provides a flexible platform that allows students to study anywhere at their own pace, and they can also review the assessment questions to gain more understanding[4]. Online formative assessment may help to identify students who tend to fail examinations and to improve their summative performance[5-6]. However, the reports of correlation between formative and summative scores remain inconsistent, and these studies were mainly conducted in subjects other than the pediatric module[5,7].

Therefore, our study aimed to evaluate the correlation between the summative assessment with formative assessment in pediatric module and other factors, including cumulative grade point average (CGPA), pediatric clerkship evaluation among fifth-year medical students. We also explored the factors that associated with unsatisfactory pediatric grade.

Study design

This study was a retrospective analytical study. The data of 84 fifth-year medical students who studied clinical clerkship in the pediatric department at the Faculty of Medicine, Vajira Hospital, Navamindradhiraj University, between August 2020 and July 2021 were collected and analyzed. The duration of the pediatric module was 10 weeks. The data that we collected comprised sex and assessment data, namely formative assessment scores (multiple-choice questions [MCQs]), CGPA, clerkship evaluation, summative assessment (SA), and pediatric basic knowledge grade. All elements of assessment data are described below.

Formative assessment (FA)

The formative assessment comprised 20 MCQs, and the scope of examination covered the core topics in general pediatrics module, including nutrition, infectious diseases, nephrology, allergy, neonatology, cardiology, growth and development, hematology, dermatology, respiratory, neurology, endocrinology, and gastroenterology. This examination was conducted online via Google Forms in the middle of the course period, and students had the option to repeat the test as many times as they desired. However, the first score of attempts will be used to evaluate. All students received overall feedback on the results, whereas students who had the three lowest scores received individualized feedback. In terms of overall feedback, students received their scores as well as advise from teachers on important topics that they should focus. For individualized feedback, teachers provided advice about the module objectives and learning achievement strategies, and students had the opportunity to ask and discuss with the teachers about their specific learning problems apart from formative outcomes. The content validity of the examination was 0.85, as evaluated by three pediatric staff members, and the reliability of the examination was 0.83, as assessed using the Kuder–Richardson formula 20. The results of the examination were non-normally distributed; therefore, we used the median (13) as the

minimum score for evaluating the association with a satisfactory grade.

Summative assessment (SA)

The summative assessment comprised 100 MCQs, and the scope of examination covered the same topics in general pediatrics as in the formative assessment. These scores accounted for 40% of the pediatric basic knowledge grade. We used similar methods as in the formative assessment to evaluate the quality of the summative examination. The content validity of all examinations was 0.84, and the reliability of the examination in each rotation was 0.82, 0.80, 0.68, and 0.85. These exams took place within the last weeks of each rotation, which were at least four weeks before formative assessment.

Cumulative grade point average (CGPA)

The CGPA was defined as the average grade point obtained in all subjects from the first to fourth year. These data were normally distributed; therefore, we used the mean of CGPA (3.26 ± 0.37) as the benchmark for evaluating the association with a satisfactory grade.

Clerkship evaluation (CE)

The clerkship evaluation was an assessment of students' performance of ward duties, consisting of two main parts as described below, and the maximum total score was 90. The scores were non-normally distributed; therefore, we used the median (70.82, interquartile range [IQR] = 5.6) as the minimum score for evaluating the association with a satisfactory grade.

1. professional habits, morals, and ethics (maximum total score 30)
    1.1. Responsibility for duties
    1.2. Accountability
    1.3. Professional personality
2. Knowledge and clinical skills (maximum total score 60)
    2.1. History taking
    2.2. Physical examination
    2.3. Investigation
    2.4. Interpretation of clinical data
    2.5. Clinical reasoning
    2.6. Plan of management

Staff members assigned these scores by direct observation of medical students' performance during ward rotations.

Pediatric basic knowledge grade

This letter grade was calculated on the basis of five components of assessment, including writing patient reports, case discussion presentations, self-study assignments, MCQ scores, and modified essay question scores. There were eight grade tiers, including A, B+, B, C+, C, D+, D, and F. We used criteria from a previous study by Luvira et al.[7] to categorize grades into two groups: grades above C were considered satisfactory, and grades C and below were considered unsatisfactory.

Statistical analysis

Results were reported as numbers and percentages for categorical variables and mean with range for continuous variables. Formative and summative scores were converted to percentage using the medial of each score and the difference in percentages was analyzed by Wilcoxon signed-rank test. The relationship between continuous variables (FA, CGPA, CE, and SA) was analyzed using Pearson's correlation. Univariate logistic regression analysis was used to identify the potential factors for unsatisfactory grade in pediatric knowledge among medical students. All data were collected and analysed using IBM SPSS Statistics for Windows, version 21.0 (IBM Co., Armonk, NY, USA), and *p*-values <0.05 were considered to indicate statistical significance.

Ethics statement

This study was approved by the Institutional Review Board at Navamindradhiraj University (COA 118/2562).

Results

A total of 84 fifth-year medical students were included in this study. Of these, 38 (45.2%) students were male, and the average CGPA was 3.26 ± 0.37. Formative assessment scores ranged from 3 to 20 out of 20, with a median of 13 (IQR = 4.75), respectively. For the clerkship performance rating, student scores ranged between 61.67 and 81.10 out of 90, with a median 70.82 (IQR = 5.6), respectively.

Regarding pediatric module learning outcomes, summative assessment MCQ scores ranged between 31 and 76 out of 100, with a mean of 59.51 (±9.5). Specifically, 12 (14.3%) students who received individualized feedback had a higher mean summative score (61.17 ± 9.5) than their counterparts (59.24 ± 9.6). The number of students who obtained grades A, B+, B, C+, C, D+, and D were 17 (20.2%), 8 (9.5%), 18 (21.4%), 21 (25.0%), 16 (19.0%), 3 (3.5%), and 1 (1.2%), respectively. According to the criteria described previously, 20 (23.8%) students received an unsatisfactory grade in pediatric knowledge.

For all students, the percentage of formative scores was significantly higher than that of summative (67.8% vs. 59.5%). In contrast, in students who obtained individualized feedback, the formative outcome was significantly lower than the summative outcome (47.9% vs. 61.2%). Pearson's correlation analysis was performed to determine the relationship between various types of assessment and summative assessment. Formative assessment scores were not correlated with summative assessment scores (r = 0.064, *p*-value = 0.561). However, a strong positive correlation was found between CGPA and summative assessment scores (r = 0.685, *p*-value = 0.001), whereas a weak negative correlation (r = -0.380, *p*-value = 0.004) was found between pediatric clerkship evaluation and summative assessment scores (table 1 and fig. 1a, 1b). We also used univariate analysis to determine the association between an unsatisfactory grade in pediatric knowledge and the other four variables. Only CGAP ≤3.26 was significantly associated with an unsatisfactory grade in pediatric knowledge (table 2).

Table 1:

Correlation analyses between formative assessment and summative assessment , cumulative grade point average and summative assessment, pediatric clerkship evaluation and summative assessment.

Correlation	Pearson's r	p-value
FA x SA	0.064	0.561
CGPA x SA	0.685	0.001
CE x SA	-0.380	0.004

CE: clerkship evaluation; CGPA: cumulative grade point average; FA: formative assessment; SA: summative assessment

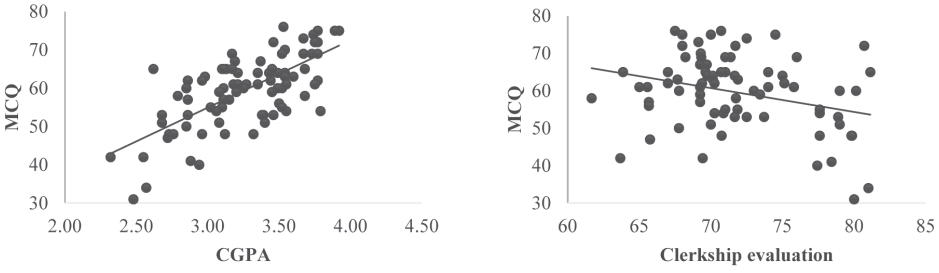

Figure a Figure b

Figure 1: Scatter plots of the correlations between Cumulative Grade Point Average and Summative Assessment (figure a) and Pediatric Clerkship Evaluation and Summative Assessment (figure b).

Table 2:

Univariate analysis of factors for the Unsatisfactory Pediatric Grade

Factors	Satisfactory Pediatric Grade (N = 64)	Unsatisfactory Pediatric Grade (N = 20)	Odd ratio (95%CI)	p-value
Sex (male)	27 (42.2)	11 (55.0)	1.579 (0.644-3.754)	0.301
Formative score ≤13	28 (43.7)	8 (40.0)	1.286 (0.540 – 3.060)	0.570
Cumulative GPA ≤3.26	22 (34.4)	17 (85.0)	10.818 (2.857-40.960)	< 0.001
Clerkship evaluation ≤70.8	35 (54.7)	8 (40.0)	0.686 (0.290-1.623)	0.391

Discussion

Formative assessment is a crucial tool for teachers in evaluating the academic performance of students, as evidenced by previous findings that formative outcomes were correlated with summative outcomes and could identify students who were likely to fail the course[6, 8]. In contrast, the present study did not find a correlation between these two outcomes, and formative performance was not a predictive factor for an unsatisfactory grade in the pediatric module among medical students. One of the reasons may be the limitation of our formative assessment method. Although we tried to create a formative examination that covered the same essential topics in the pediatric module as the summative examination, there were only 20 items in the formative examination, which was a considerably smaller number than the 100 MCQ items in the summative examination (construct underrepresentation). Therefore, it was difficult to provide comprehensive knowledge in all aspects of the topics. Furthermore, although the online formative test was available for students to retake during the course, we did not find any students who did so. This may be the reason why they did not gain a benefit from this online platform.

The individualized feedback in this study may also be one of the factors that contributed to negative outcomes. As previously stated, students who obtained the three lowest scores received personal advice. This strategy may encourage students to perform better in the summative examination, which was consistent with our results that their summative examination scores were significantly higher than their formative examination scores, and the mean of their scores was also slightly higher than that of other students. This finding is also supported by previous reports. According to Nolette et al., using a focus group as a formative evaluation in an accelerated pharmacy curriculum helped students feel more comfortable asking questions and better comprehend difficult concepts[9]. Luvira et al. also did not find a correlation between formative and summative outcomes; however, they reported that undergraduate students who underwent formative evaluation were more likely to achieve a satisfactory grade in epidemiology[7].

Academic achievement among medical students was measured by CGPA, which is influenced by a number of factors, such as learners' attitude, learning motivation, and daily habits[10-12]. This indicates that medical students who obtained a decent CGPA have learning characteristics

(e.g., high learning concentration) that assist them in achieving expected learning outcomes. Our results demonstrated a strong correlation between CGPA and summative examination scores, and medical students who had CGPA ≤3.26 tended to obtain an unsatisfactory grade in the pediatric module. As a result, CGPA remains an important tool for teachers in determining which medical students are most likely to fail to achieve satisfactory outcomes.

Notably, we found a negative correlation between CE and summative assessment scores. Not only were knowledge and clinical skills assessed during the clerkship but also professionalism, which could not be evaluated through the MCQs. Therefore, this component may be the factor that influenced this negative correlation. Previously, Greenburg et al. also reported that USMLE step 1 score could not predict the professionalism scores among medical students during internship[13]. Moreover, although the pediatric staff used a structured form to evaluate students, they still performed the assessment subjectively, which may have affected the reliability and validity of the rating. This can be explained by two effects, namely halo and leniency errors; for example, evaluators may use one area of performance to judge other areas or may have various factors influencing their judgment, such as fear of negative feedback from students, fear of the impact on the student–teacher relationship, and a lack of evaluation skills. Specifically, the structured mark sheet could not eliminate the leniency effect[14].

However, there were certain limitations to our study. Medical students from different clerkship rotations were asked the same questions in the formative assessment, but they were provided different sets of MCQs for the summative assessment. As a result of the examinations' varying reliability, the students may have encountered different levels of difficulty in the MCQ examination. Furthermore, the effectiveness of online formative assessment may not be observed because of the limitation of the retrospective analytical design. A well-designed randomized controlled trial study should be performed in the future.

Conclusion

Our study did not find the correlation between the online formative and the summative assessment. Therefore, clinical Teachers should consider other factors, especially CGPA, to identify students who have a tendency for academic failure. Besides, individualized feedback could be an essential component of formative assessment for improving learning outcomes in the pediatric module.

Conflict of interest

The authors report no conflict of interest.

Acknowledgment

The authors thank Mrs. Woranit Samranjai for assistance in data collection. The authors also thank the Medical Education Division of the Faculty of Medicine at Vajira Hospital, Navamindradhiraj University, that helped in providing the data.

References

1. Bell B, Cowie B. The characteristics of formative assessment in science education. Sci Ed 2001; 85:536–53.
2. Black P, Wiliam D. Developing the theory of formative assessment. Educ Asse Eval Acc 2009;21:5-33.
3. Bing-You R, Hayes V, Varaklis K, Trowbridge R, Kemp H, McKelvy D. Feedback for Learners in Medical Education: What Is Known? A Scoping Review. Acad Med 2017;92(9):1346-54.
4. Ellaway R, Masters K. AMEE Guide 32: e-Learning in medical education Part 1: Learning, teaching and assessment. Med Teach 2008;30(5):455-73.
5. Velan GM, Jones P, McNeil HP, Kumar RK. Integrated online formative assessments in the biomedical sciences for medical students: benefits for learning. BMC Med Educ 2008;8:52. doi: 10.1186/1472-6920-8-52.

6. Azzi AJ, Ramnanan CJ, Smith J, Dionne E, Jalali A. To quiz or not to quiz: Formative tests help detect students at risk of failing the clinical anatomy course. Anat Sci Educ 2015;8(5):413-20.
7. Luvira V, Bumrerraj S, Srisaenpang S. Formative Evaluation and Learning Achievement in Epidemiology for Preclinical Medical Students. Indian J Community Med 2018;43(4):298-301.
8. Mitra NK, Barua A. Effect of online formative assessment on summative performance in integrated musculoskeletal system module. BMC Med Educ 2015;15:29. doi: 10.1186/s12909-015-0318-1.
9. Nolette S, Nguyen A, Kogan D, Oswald C, Whittaker A, Chakraborty A. Efficacy of formative evaluation using a focus group for a large classroom setting in an accelerated pharmacy program. Curr Pharm Teach Learn 2017;9(4):633-8.
10. Mandal A, Ghosh A, Sengupta G, Bera T, Das N, Mukherjee S. Factors affecting the performance of undergraduate medical students: a perspective. Indian J Community Med 2012;37(2):126-9.
11. Al Shawwa L, Abulaban AA, Abulaban AA, Merdad A, Baghlaf S, Algethami A, et al. Factors potentially influencing academic performance among medical students. Adv Med Educ Pract 2015;6:65-75.
12. Kwankajonwong N, Ongprakobkul C, Qureshi SP, Watanatada P, Thanprasertsuk S, Bongsebandhu-Phubhakdi S. Attitude, but not self-evaluated knowledge, correlates with academic performance in physiology in Thai medical students. Adv Physiol Educ 2019;43(3): 324-31.
13. Greenburg DL, Durning SJ, Cohen DL, Cruess D, Jackson JL. Identifying medical students likely to exhibit poor professionalism and knowledge during internship. J Gen Intern Med 2007;22(12):1711-7.
14. McKinstry BH, Cameron HS, Elton RA, Riley SC. Leniency and halo effects in marking undergraduate short research projects. BMC Med Educ 2004;4:28. doi: 10.1186/1472-6920-4-28.