Does students’ performance in the formative CLIPP examination predict their scores in the NBMEPSE?

Lolowa Almekhaini, Ahmad R. Alsuwaidi, Khaula Khalfan Alkaabi, Sania Al Hamad and Hassib Narchi
Department of Pediatrics, UAE University, Al Ain, United Arab Emirates

Abstract
Purpose – Computer-Assisted Learning in Pediatrics Program (CLIPP) and National Board of Medical Examiners Pediatric Subject Examination (NBMEPSE) are used to assess students’ performance during pediatric clerkship. International Foundations of Medicine (IFOM) assessment is organized by NBME and taken before graduation. This study explores the ability of CLIPP assessment to predict students’ performance in their NBMEPSE and IFOM examinations.
Design/methodology/approach – This cross-sectional study assessed correlation of students’ CLIPP, NBMEPSE and IFOM scores. Students’ perceptions regarding NBMEPSE and CLIPP were collected in a self-administered survey.
Findings – Out of the 381 students enrolled, scores of CLIPP, NBME and IFOM examinations did not show any significant difference between genders. Correlation between CLIPP and NBMEPSE scores was positive in both junior ($r = 0.72$) and senior ($r = 0.46$) clerkships, with a statistically significant relationship between them in a univariate model. Similarly, there was a statistically significant relationship between CLIPP and IFOM scores.
In an adjusted multiple linear regression model that included gender, CLIPP scores were significantly associated with NBME and IFOM scores. Male gender was a significant predictor in this model. Results of survey reflected students’ satisfaction with both NBMEPSE and CLIPP examinations.
Originality/value – Although students did not perceive a positive relationship between their performances in CLIPP and NBMEPSE examinations, this study demonstrates predictive value of formative CLIPP examination scores for their future performance in both summative NBMEPSE and IFOM. Therefore, students with poor performance in CLIPP are likely to benefit from feedback and remediation in preparation for summative assessments.

Keywords Pediatrics clerkship, Assessments, Medical education, UAE

Paper type Research paper

Introduction
The College of Medicine and Health Sciences (CMHS) at the United Arab Emirates University (UAEU), a public federal university, offers a 6-year training program consisting of a 2-year premedical and a 4-year MD program that is divided into two phases of 2 years each: preclinical and clinical training, which center on different specialties. An 8-week pediatrics clerkship is offered in year five (junior clerkship), and a further 4-week clerkship in the final year (senior clerkship). Each clerkship program has a structured teaching program that includes bedside teaching in affiliated teaching hospitals. At the end of each clerkship...
program, students undergo an assessment that includes mini-clinical assessments (MiniCLEX) and the National Board of Medical Examiners Pediatrics Subject Examination (NBMEPSE).

The National Board of Medical Examiners (NBME) is an independent, nonprofit and web-based American organization established to develop countrywide examination regulations that allow medical accrediting authorities to evaluate candidates for licensing issuance. The NBME subject examination is used internationally to assess and train students in medical schools (Wright & Baston, 2017). It offers standardized and objective examinations in various disciplines, and these examinations are used throughout all phases of medical studies, particularly in the final assessments after clerkships (Johnson, Khalil, Peppler, Davey, & Kibble, 2014; National Board of Medical Examiners, 2016). Various studies have identified the NBME subject examination as the best intervention to evaluate low-performing students before graduation (Johnson et al., 2014; Guiot & Franqui-Rivera, 2018). It also serves as a global resource and a model for assessing approaches and appraisals in medicine, with a wide array of assessments that are well suited for medical schools (Guiot & Franqui-Rivera, 2018). These assessments allow medical schools to appraise their students' achievement and compare their performance with a sizable, global and representative group of medical students at the same training stage (Bakoush, Al Dhanhani, Alshamsi, Grant, & Norcini, 2019). They assess both basic and clinical science knowledge (National Board of Medical Examiners, 2016) and comprise 100 multiple-choice questions (MCQs), each constructed in the form of clinical vignettes with one best answer that focuses on the application and integration of comprehension (National Board of Medical Examiners, 2019, 2020). Each of these examinations evaluates students' comprehension of normal development (5%–10% of the questions), as well as a majority of cardiovascular, neurological, renal and other diseases, with each domain evaluated with 15% of the questions. These examinations also assess the important tasks of physicians, including health maintenance and promotion (5%–10% of questions), comprehension of disease mechanisms (25%–30% of questions), establishment of a diagnosis (40%–45% of questions) and the application of management principles (10%–15% of questions). The content area classification is as follows: foundation, diagnosis, management and prevention. The scores are compared across examinations, schools and examinee performance records. Individual performance profiles of correct percentage scores are reported. This web-based examination is administered in a secured setting in our college. In the final clerkship assessments administered, the NBMEPSE scores account for 30% of the total assessments (Narchi, 2013, 2014). In this study, NBME 1 refers to the examination offered at the end of the junior clerkship level (Year 1), while NBME 2 denotes the assessment administered at the end of the senior clerkship (Year 2). Students are provided with a score interpretation guide and a performance profile to facilitate their self-assessment (National Board of Medical Examiners, 2019, 2020).

In 2017, the Computer-Assisted Learning in Pediatrics Program (CLIPP) was added to both junior and senior pediatrics clerkships as a midterm formative assessment taken voluntarily. This is an online, validated assessment of 100-item MCQs created by the Council on Medical Student Education in Pediatrics (COMSEP), which is an independent organization of pediatric medical student education (Schifferdecker, Berman, Fall, & Fischer, 2012; Fall et al., 2005; Sox et al., 2018; COMSEP, 2020). The examination is delivered via a secure browser; it complements the performance assessment of students during their clerkship, enhances students' self-directed learning and helps develop appropriate remediation when needed (Aquifer pediatrics, 2020). It also helps students prepare for the summative NBMEPSE at the end of their clerkship training. The students receive their CLIPP reports with their respective scores well before their NBMEPSE examination. In addition, the CLIPP reports also help ascertain the effectiveness of the curriculum as well as teaching techniques.
employed by instructors. In this study, CLIPP 1 refers to the examination administered during junior clerkship (Year 1) and CLIPP 2 to the one administered during senior clerkship (Year 2).

At the end of their last year of study, just before graduation, students sit for the International Foundation of Medicine (IFOM) examination, which covers all aspects of clinical medicine as opposed to the NBME Subject Examination and the CLIPP. The IFOM examination comprises 160 MCQs from basic clinical science in the following clinical areas: pediatrics, medicine, family medicine, surgery, psychiatry, and obstetrics and gynecology (IFOM, 2019a, b). In addition, it allows the comparison of CMHS-UAEU students’ performance prior to graduation with that of final-year medical students around the world. Therefore, the CMHS-UAEU uses the IFOM Clinical Science Examination as an essential part of assessing students prior to graduation. In this study, IFOM 2 refers to the examination administered in the final year prior to graduation.

The objectives of this study were as follows: (1) to evaluate, in each clerkship, the predictive value of the CLIPP formative assessment on students’ performance in the NBMEPSE summative assessment as well as on their performance in the final IFOM assessment and (2) to survey the students’ satisfaction and perceptions regarding the benefits of both of these evaluation tools. The null hypothesis was that the formative CLIPP scores do not predict the NBMEPSE or the IFOM scores. If that hypothesis was rejected for both independent summative assessments (NBMEPSE and IFOM) in both clerkships, the value of the formative CLIPP assessment to predict the two latter scores would be firmly validated. This would therefore suggest that identifying the students performing poorly in the CLIPP assessment would pave the way for the implementation of remediation measures that can improve students’ future performance in both the NBMEPSE and the IFOM.

Methods
This study, undertaken at the CMHS-UAEU, is a retrospective observational review of the marks obtained by all junior and senior students as well as students who recently graduated after completing both their voluntary formative CLIPP and the mandatory summative NBMEPSE at the end of each of the two pediatric clerkships between October 2017 and June 2020.

In addition, at the end of each NBMEPSE, students were asked to participate in a 5-min online survey via a self-administered questionnaire regarding their perceptions about the value of the CLIPP examination and any correlation they think it has with the NBMEPSE (the complete survey is available as Supplemental Digital Appendix 1). The survey was anonymous and did not have any impact on the students’ progression throughout their medical studies. By completing the online survey, the students confirmed their voluntary consent to participate in this study. Participants’ confidentiality was maintained, and no identifying data were included in data analysis. The participants were informed that they reserved the right to withdraw at any stage of the study without facing any penalties.

Statistical analysis
After testing for normality (Wilks–Shapiro test), the examination scores were expressed as mean value and standard deviation (SD). The group mean scores were compared using the unpaired t-test, and the correlation between the students’ CLIPP and NBMEPSE scores was tested using the Pearson product–moment correlation. The correlation between the CLIPP and NBMEPSE scores in each clerkship was tested using a univariate linear multivariate regression model, and the IFOM scores were tested using a linear multivariate regression.
model including gender. The results of the survey were expressed as percentages of students selecting a particular option on a 5-point Likert scale. All the statistical tests were performed using the STATA version 15 software package (StataCorp, Texas, USA), and a two-sided $p$ value < 0.05 defined statistical significance.

Results

Participants
A total of 381 students (299 females, 78.5%) were enrolled during the study period. Out of these, 286 students had completed both junior and senior clerkships. In the junior clerkship, the CLIPP-1 assessment was taken by 200 students and the NBME-1 by 353 students, with 200 taking both examinations. The CLIPP-2 assessment and the NBME-2 assessment in the senior clerkship were taken by 223 and 286 students, respectively, with 223 taking both examinations. The IFOM-2 examination was taken by 167 senior students (43.8%). However, in total, only 128 junior and senior students completed the online survey.

Assessment scores
There were no significant differences in the CLIPP, NBME and IFOM scores based on gender (Table 1). Similarly, there were no significant differences in the NBME and IFOM scores between the students who voluntarily took the CLIPP assessment and those who did not (Table 2).

Correlation between the CLIPP and NBME scores
While there was a good correlation ($r = 0.72$) between both scores and in both genders in the junior clerkship, it was more modest ($r = 0.46$) in the senior clerkship where it was weak in females ($r = 0.44$) but remained strong ($r = 0.73$) in males (Table 3).

Table 1. Examination scores of students in the two clerkships

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>$p$ value*</th>
<th>All students</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLIPP 1</td>
<td>39.7 ± 11.9</td>
<td>39.1 ± 11.0</td>
<td>0.7</td>
<td>39.3 ± 11.2</td>
</tr>
<tr>
<td>NBME 1</td>
<td>51.0 ± 13.0</td>
<td>52.6 ± 1.9</td>
<td>0.2</td>
<td>52.3 ± 12.2</td>
</tr>
<tr>
<td>CLIPP 2</td>
<td>38.2 ± 13.5</td>
<td>42.6 ± 13.4</td>
<td>0.06</td>
<td>41.8 ± 13.5</td>
</tr>
<tr>
<td>NBME 2</td>
<td>56.1 ± 14.4</td>
<td>54.4 ± 14.5</td>
<td>0.4</td>
<td>54.8 ± 14.5</td>
</tr>
<tr>
<td>IFOM 2</td>
<td>529.6 ± 95.8</td>
<td>522.8 ± 82.6</td>
<td>0.6</td>
<td>524.3 ± 85.4</td>
</tr>
</tbody>
</table>

Note(s): $n = 200$ in junior clerkship and $223$ in senior clerkship, expressed as mean ± standard deviation (SD)
Assessments 1 and 2 represent junior and senior pediatric clerkships, respectively
*unpaired $t$-test
Source(s): Authors’ work

Table 2. Comparison of NBME and IFOM examination scores between students who voluntarily took the CLIPP assessment and those who did not

<table>
<thead>
<tr>
<th></th>
<th>CLIPP taken</th>
<th>CLIPP not taken</th>
<th>$p$ value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total students</td>
<td>200</td>
<td>143</td>
<td>0.17</td>
</tr>
<tr>
<td>NBME 1</td>
<td>53.0 ± 12.7</td>
<td>51.2 ± 11.3</td>
<td>0.63</td>
</tr>
<tr>
<td>Total students</td>
<td>223</td>
<td>63</td>
<td>0.63</td>
</tr>
<tr>
<td>NBME 2</td>
<td>54.6 ± 14.9</td>
<td>55.5 ± 13.1</td>
<td>0.27</td>
</tr>
<tr>
<td>Total students’</td>
<td>151</td>
<td>16</td>
<td>0.27</td>
</tr>
<tr>
<td>IFOM 2</td>
<td>526.6 ± 86.4</td>
<td>502.3 ± 74.2</td>
<td>0.27</td>
</tr>
</tbody>
</table>

Note(s): The results are expressed as mean ± standard deviation (SD), unless stated otherwise
Assessments 1 and 2 represent junior and senior pediatric clerkships, respectively
*unpaired $t$-test
Source(s): Authors’ work
Modeling of the relationship between the CLIPP and NBME scores
In both clerkships, there was a statistically significant linear relationship between both scores (Table 4). In Junior Clerkship 1, for each increase by 1 mark in the CLIPP score, the resulting increase in the NBME mark was 0.82 (Table 3) regardless of gender. In Senior Clerkship 2, for each increase by 1 mark in the CLIPP score, the resulting increase in the NBME mark was 0.54, with an additional effect of gender; that is to say, the CLIPP scores of male students were expected to increase by 8.7 marks when compared with those of female students (Table 4).

Modeling of the relationship between the CLIPP and IFOM scores
In the unadjusted simple linear regression model, CLIPP 2 scores were significantly associated with the final IFOM 2 scores. However, in the adjusted linear regression model, gender was significantly associated with the IFOM scores. In the senior clerkship, for each extra mark in CLIPP 2 scores, the IFOM 2 score increased by 3.82 marks, with an additional 43.15 marks in male students’ scores when compared with female students’ scores (Table 4).

CLIPP survey
Most students were happy about the suitability of the setting and venue. They were also positive about the contents of the scenarios and questions, the authenticity of the questions, their clarity, the optimal use of software and the examination duration (Figure 1 and Table 5). However, many students did not feel that CLIPP prepared them well for the NBME or helped predict their performance in the NBME. They also felt that the feedback they received after the CLIPP assessment did not help them prepare for the NBME.

Discussion
This study confirms that both the summative NBME and IFOM scores increase in proportion to the formative CLIPP score, corroborating that both can be predicted by formative CLIP scores. This significant relationship between formative CLIPP scores and summative NBMEPSE scores, which is not surprising as both examinations assess knowledge in pediatrics, was also found between the CLIPP assessments – exclusively in pediatrics – and the final IFOM summative examinations, which evaluate knowledge in all branches of medicine. This strongly reinforces the strong predictive value of the formative CLIPP assessment regarding all summative examinations. However, this is not unexpected because these examinations assess students’ knowledge, also confirming the findings of previous studies demonstrating the positive correlation between the CLIPP examination and the summative NBME scores (Morrison et al., 2010). In addition, the use of its patient case simulations helps improving the teaching quality (Leong, Baldwin, & Adelman, 2003). Identifying poorly performing students in the formative CLIPP assessment and giving them appropriate feedback and remediation should, therefore, intuitively help the students improve their performance in the summative NBME and IFOM assessments.

Although this study employs CLIPP as a formative pediatric assessment, it is surprising that it also predicts the final IFOM examination that encompasses all branches of medicine.

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>All students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junior clerkship (Year 1)</td>
<td>0.72</td>
<td>0.72</td>
<td>0.72</td>
</tr>
<tr>
<td>Senior clerkship (Year 2)</td>
<td>0.73</td>
<td>0.44</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Source(s): Authors’ work

Table 3. Pearson’s r correlation between the NBME and CLIPP scores in each of the junior and senior clerkships.
<table>
<thead>
<tr>
<th>Students</th>
<th>Unadjusted models*</th>
<th>Adjusted model§</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient for CLIPP (95% ci)</td>
<td>Constant</td>
</tr>
<tr>
<td></td>
<td>Coefficient for CLIPP (95% ci)</td>
<td>Constant</td>
</tr>
<tr>
<td>NBME 1 (junior)</td>
<td>200</td>
<td>0.82 (0.71, 0.93)</td>
</tr>
<tr>
<td>NBME 2 (senior)</td>
<td>223</td>
<td>0.51 (0.38, 0.63)</td>
</tr>
<tr>
<td>IFOM 2 (senior)</td>
<td>151</td>
<td>3.71 (2.9, 4.5)</td>
</tr>
</tbody>
</table>

**Note(s):** * Simple linear univariate regression; § multivariate linear regression model adjusting for gender

Predicted NBME 1 score = (0.82 × CLIPP 1 score) + 20.4 ($p < 0.001$)

Predicted NBME 2 score = (0.54 × CLIPP 2 score) + 30.48 + 8.78 if male student ($p < 0.001$)

Predicted IFOM score = (3.82 × CLIPP 2 score) + 369.31 + 43.15 if male student ($p < 0.001$)

**Source(s):** Authors' work
It is not expected that improving students' knowledge in pediatrics, demonstrated in the CLIPP results, would enhance their knowledge in all other branches of medicine as tested in the IFOM. Therefore, we can speculate that the students’ scores in both CLIPP and IFOM
examinations reflect the students’ own inherent learning abilities across a wide range of domains instead of attributing the enhancement of their IFOM scores directly and solely on their pediatric CLIPP experience.

Surprisingly, the students did not perceive any value of the CLIPP feedback in helping them prepare for the NBME. This could be attributed to the fact that they currently receive only their CLIPP scores without any detailed feedback on their performance, thus making it impossible to identify their weaknesses and take the necessary measures to improve their future performance. The participants believe that the CLIPP examination is easier than the NBME examination as it focuses on a limited range of clear objective pediatrics conditions as opposed to the NBME examination, which covers a broad range of pediatric ailments and requires deeper and more advanced knowledge. This is perhaps the reason for which the students felt that the CLIPP examination had no significant role in improving their readiness for the NBME examination. However, they admitted that taking the CLIPP examination earlier in their clerkship helped them objectively benchmark their knowledge and identify their strengths and weaknesses. Consequently, this encouraged them to improve their knowledge.

In contrast to the plethora of reports in the literature evaluating the NBME and the United States Medical Licensing Examination (USMLE) (Boscardin, Earnest, & Hauer, 2020; Torre et al., 2020; Rosenthal et al., 2019), the paucity of publications regarding the relationship between the CLIPP assessment and other examinations in pediatric clerkships does not allow researchers to make any comparisons (Schifferdecker et al., 2012; Sox et al., 2018). This is probably due to the fact that CLIPP is an exclusive pediatric assessment as opposed to the NBMEPSE examinations, which cover all medical disciplines.

This study has some limitations. In addition to the small sample size of students, not all the students took the two formative CLIPP assessments, the two summative NBMEs or even the final IFOM assessment. In addition, the effects of the remediation provided to some students after their CLIPP assessment reports and just before the summative assessments were not analyzed because no systematic documentation existed. Finally, because the CLIPP examination was an optional and voluntary formative assessment, its outcomes could not be assessed in all students. However, the findings of this study showed no significant differences in the NBME and IFOM scores between students who voluntarily took the CLIPP assessment and those who did not. This implies that the proportionality between the CLIPP assessment scores and both summative assessments scores could be attributed to the students’own intellectual capabilities as opposed to the feedback received from the summative assessments.

Future studies should consider using a larger sample size and conducting a formal analysis of the role of remediation measures implemented after the CLIPP on the summative assessment scores. Including the students’ NBME performance in other medical disciplines (which do not have CLIPP assessments) in the analysis of the final IFOM scores would also help better delineate the specific contribution of the CLIPP and NBMEPSE pediatric assessments to the IFOM.

**Conclusion**

This study reveals that the CLIPP examination is a good predictor of students’ performance in both the NBMEPSE assessments during pediatrics clerkships and the IFOM assessment prior to graduation. Therefore, giving students detailed feedback on their performance in the CLIPP examination is essential and should be an integral part of their educational experience. Appraising students assessment tools provides robust evidence-based information for improving their performance. Future studies should include larger sample size to evaluate the role of pre-assessment preparation, post-assessment feedback and appraisal of remedial
measures to improve future performance. Introducing and examining a new assessment tool matched with appropriate elements of instructional design is imperative. These may include gamification, interface design, density of the content presented, in addition to the learning and cognitive styles of the learners.

References


CLIPP scores predict NBMEPSE and IFOM scores


Supplemental Digital Appendix 1

Study survey
Please answer all items as much as you can, even if some do not seem to apply to the study. There are six-point scoring scale for each item

<table>
<thead>
<tr>
<th>N/A</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>not applicable</td>
<td>strongly disagree</td>
<td>disagree</td>
<td>neutral</td>
<td>agree</td>
</tr>
</tbody>
</table>

1- **Content**
1- The CLIPP exam assesses the pediatrics clerkships’ learning objectives.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
2- The CLIPP exam is at your academic-year student level.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
3- The content of the CLIPP exam, including the scenarios and answer options, is clear, well written, and up to date as NBME.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
4- The CLIPP exam cases are authentic.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
5- The CLIPP exam provides appropriate preparation for pediatrics NBME.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
6- The CLIPP exam is a useful experience of pediatrics clerkships.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
7- The CLIPP exam results feedback helps you predict your future performance in Pediatrics NBME.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
8- The CLIPP exam results feedback guides your learning objectives plan.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
9- The CLIPP exam results feedback is a useful tool in preparing for pediatrics NBME.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
10- The CLIPP exam and results feedback are important to be included in the pediatric clerkship curriculum.
    - N/A
    - 1
    - 2
    - 3
    - 4
    - 5

2- **CLIPP exam timing and setting**
11- The CLIPP exam timing is appropriate to support your preparation for the end-of-rotation pediatrics NBME exam.
   - N/A
   - 1
   - 2
   - 3
   - 4
   - 5
12- The CLIPP exam duration is sufficient to answer all questions.
    - N/A
    - 1
    - 2
    - 3
    - 4
    - 5
13- The CLIPP exam software is used optimally.
    - N/A
    - 1
    - 2
    - 3
    - 4
    - 5
14- The CLIPP exam scenarios, questions, and answer options are visually clear and suitable.
    - N/A
    - 1
    - 2
    - 3
    - 4
    - 5
15- The CLIPP exam setting/venue is appropriate.
    - N/A
    - 1
    - 2
    - 3
    - 4
    - 5

What concerns you most about the CLIPP or NBME examination?

Please describe the best things about the CLIPP and NBME examination?

---

**Corresponding author**
Lolowa Almekhaini can be contacted at: lolwa.mukini@uaeu.ac.ae

---

For instructions on how to order reprints of this article, please visit our website: [www.emeraldgrouppublishing.com/licensing/reprints.htm](http://www.emeraldgrouppublishing.com/licensing/reprints.htm)
Or contact us for further details: permissions@emeraldinsight.com