RESEARCH

Assessing the Relationship Between Didactic Performance and Standardized Examination Scores in Pharmacy Students

Kimberly K. Daugherty, PharmD\textsuperscript{a}, Daniel R. Malcom, PharmD\textsuperscript{a,b}

\textsuperscript{a} Sullivan University, College of Pharmacy and Health Sciences, Louisville, Kentucky
\textsuperscript{b} Associate Editor, \textit{American Journal of Pharmaceutical Education}, Arlington, Virginia

Submitted May 29, 2019; accepted December 23, 2019; published August 2020.

**Objective.** To examine the effect size of third professional (P3) year students’ grade point average (GPA) on Pharmacy Curriculum Outcomes Assessment (PCOA) scores and to summarize the effect size of PCOA scores on North American Pharmacist Licensure Examination (NAPLEX) scores.

**Methods.** To accomplish the objective, meta-analyses were conducted. For inclusion in the meta-analysis, studies were required to compare PCOA scores to and report data that permitted calculation of a numeric effect size for the chosen outcome variables. Multiple databases were searched, including PubMed, CINAHL, EMBASE, ProQuest Dissertations and Thesis (abstract limited), Academic Search Complete, and Google Scholar. Correlations were used as the effect size metric for all outcomes. All analyses used an inverse variance weighted random effects model. Study quality was reviewed for each study included in the meta-analyses.

**Results.** This study found that PCOA scores were moderately correlated with P3 GPAs, accounting for 14% to 48% of the variability in PCOA scores. The meta-analyses also showed that PCOA scores were moderately correlated with NAPLEX and accounted for 25% to 53% of the variability in NAPLEX scores. Both meta-analyses showed a high degree of heterogeneity and many studies included were of low quality.

**Conclusion.** This first set of meta-analyses to be conducted on the PCOA showed that third professional year GPA does correlate with PCOA results and that PCOA scores correlate with NAPLEX results. Though there are significant limitations to interpretation of the results, these results do help further elucidate the role of the PCOA as a benchmark of progress within the pharmacy curriculum.

**Keywords:** PCOA, meta-analysis, curricular outcomes, NAPLEX

INTRODUCTION

The Pharmacy Curriculum Outcomes Assessment (PCOA) is a comprehensive assessment tool developed by the National Association of Boards of Pharmacy (NABP) to evaluate student performance in the Doctor of Pharmacy (PharmD) curriculum.\cite{1} As a component of pharmacy school accreditation, Standards 2016 from the Accreditation Council for Pharmacy Education (ACPE) require that programs administer the PCOA annually to students nearing completion of the didactic portion of the curriculum and report the results.\cite{2,3,4} Prior to and after its incorporation as a requirement in Standards 2016, institutions reported using the PCOA for different purposes, including benchmarking against other programs, reviewing student performance in the curriculum, and assessing curricular quality.\cite{3,4} The purpose of the PCOA is to provide a comprehensive knowledge examination that could be used by all schools and colleges of pharmacy to assess student knowledge prior to entering the experiential part of the curriculum, despite programs having different lengths (0-6, 0-5, four-year, or three-year) and curricula.\cite{1} Programs can administer the PCOA in the manner that best fits their specific curricular model and timing of didactic and experiential coursework, including administering it multiple times as a cohort progresses through the program. However, as noted previously, ACPE accreditation requirements mandate reporting at minimum PCOA results for students “nearing completion of the didactic curriculum,” which would be in the third professional year (P3) prior to beginning advanced pharmacy practice experiences (APPEs) for most traditional four-year PharmD programs.

The PCOA is composed of 225 items. Two hundred of these items (88.9%) contribute to the score, while 25
(11.1%) are used as test items and do not count toward the overall score. The examination is divided into four main content areas: basic biomedical sciences (20 items, 10% of the examination), pharmaceutical sciences (66 items, 33% of the examination), social/behavioral/administrative sciences (44 items, 22% of the examination), and clinical sciences (70 items, 35% of the examination). Each of these four content areas is then further divided into 28 subtopic areas, examples of which include “Pharmacology and Toxicology” under Area 2.0: Pharmaceutical Sciences and “Clinical Pharmacogenomics” under Area 4.0: Clinical Sciences. Item types include traditional multiple-choice (for which only one response is correct), multiple-response (for which more than one item is correct), constructed-response (for which the test taker must supply a response), ordered response (for which the test taker must correctly sequence a list of items), and “hot spot” (for which the test taker must manually indicate the correct location within a picture or diagram). The most current blueprint for this examination was put into place in 2016.\textsuperscript{5-7}

Despite information from NABP and ACPE on the use of the PCOA, little reliable standardized national data are available on the manner in which programs are using the PCOA. Published reports have revealed programs using the examination as a part of a “capstone”-type approach to the pre-APPE coursework (with and without remediation), while others have used it as a benchmarking examination for multiple years of a curriculum by administering the examination to all students in the program each year.\textsuperscript{3,4} A consensus regarding incentive structures and/or remediation practices has not been reached. A national survey of 125 US pharmacy programs published in 2018 regarding use of the PCOA revealed that 59% of programs were using the PCOA total scaled score for curriculum assessment, with 66% using the total score percentile for that purpose as well.\textsuperscript{3} The same survey showed that 71% of programs desired to have more data prior to making curricular changes based on the PCOA, with 49% responding that benchmarking data would not be useful at the present time because programs administer the examination differently. Seventy-one percent of programs reported offering no incentives to students for completing the PCOA, but 23% reported that non-participation in the examination was considered a violation of professionalism standards or the student code of conduct.\textsuperscript{3}

Multiple studies to date have evaluated the PCOA as either an outcome of interest or a potential surrogate marker for student performance on other standardized examinations such as the North American Pharmacist Licensure Examination (NAPLEX) shown in Table 1. Other variables that have been studied as outcomes of interest include pharmacy student grade point averages (GPAs) as they progress through the program.\textsuperscript{8-25} The NAPLEX is a 250-item computerized examination administered by NABP and is required for licensure in the United States. The blueprint for the NAPLEX was last updated in 2016.\textsuperscript{26}

Despite significant discussion within the Academy regarding how to best utilize PCOA results and the examination’s continuous use for over a decade, no meta-analysis of published data regarding association of PCOA scores with metrics of interest has been published.\textsuperscript{27} Though only a relatively small number of studies have been published, a meta-analysis is timely to summarize current available data and to help illuminate future areas of research. According to the survey by Gortney and colleagues, 71% of programs desire more data prior to making curricular changes based on the PCOA.\textsuperscript{3} Additionally, with the requirement to administer the PCOA for all programs, a comprehensive analysis of its usefulness to date as a gauge of progress through the PharmD curriculum could be helpful for programs attempting to understand how their PCOA data can fit into their overall assessment plan. The objectives of this study were to summarize the available evidence on the effects of third professional year students’ GPAs (P3 GPAs) on their performance on the PCOA as well as the effects of their PCOA scores on their NAPLEX scores. The primary objectives of the meta-analyses conducted were to determine whether P3 GPA had a significant effect on PCOA scores and to summarize the effect that performance on the PCOA has on NAPLEX scores as reported in articles published to date.

**METHODS**

To be eligible for inclusion in this review, studies were required to compare the PCOA to an eligible outcome and report data that permitted calculation of a numeric effect size for the outcome variable(s) included in the study. Note that the eligibility criteria and coding scheme (described below) were identical for all outcomes included in this study.

Eligible outcomes included any outcome that was used to predict PCOA scores or that PCOA was used to predict. Outcomes were only excluded if there were fewer than 10 studies related to that specific outcome or if there were not enough data to calculate effect sizes. Examples of qualifying outcomes for the initial review included: pre-pharmacy GPA, PCAT scores, P1 GPA, P2 GPA, P3 GPA, NAPLEX results, motivation, age, sex, and degrees prior to matriculation. The list of final included outcomes

---

**Table 1: Study Eligibility Criteria**

- Outcome must be a valid measure of student performance or student attributes of interest.
- Outcome must be measured on a numeric scale.
- Outcome must be reported in a manner that permits effect size calculation.
- Study must include a control group or comparison group.
- Study must be published in a peer-reviewed journal.
- Study must include at least 10 eligible outcomes.
Table 1. Pre-Advance Pharmacy Practice Experience (Pre-APPE) Pharmacy Curriculum Outcomes Assessment (PCOA) Meta-Analyses Study Quality and Outcome Summary

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year Published</th>
<th>When Given</th>
<th>Study Time Provided (Y/N)</th>
<th>Incentives Provided (Y/N)</th>
<th>How GPA Calculated</th>
<th>Sample Size</th>
<th>Outcomes Analyzed</th>
<th>Correlation Results (r) by Outcome of Interest (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Campbell, et al&lt;sup&gt;8&lt;/sup&gt;</td>
<td>2018</td>
<td>NR</td>
<td>Low</td>
<td>NR</td>
<td>NR</td>
<td>246</td>
<td>NAPLEX</td>
<td>NAPLEX r = .72</td>
</tr>
<tr>
<td>Ferrence &amp; Welch&lt;sup&gt;9&lt;/sup&gt;</td>
<td>2017</td>
<td>NR</td>
<td>Mod</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td>NAPLEX</td>
<td>NAPLEX gr 1 (255) r = .571, NAPLEX gr 2 (194) r = .677</td>
</tr>
<tr>
<td>Garavalia, et al&lt;sup&gt;10&lt;/sup&gt;</td>
<td>2017</td>
<td>Spring</td>
<td>Low</td>
<td>Y</td>
<td>NR</td>
<td>215</td>
<td>P3GPA</td>
<td>NAPLEX r = .56, P3GPA r = .51</td>
</tr>
<tr>
<td>Gillette, et al&lt;sup&gt;11&lt;/sup&gt;</td>
<td>2017</td>
<td>Winter</td>
<td>Low</td>
<td>Y</td>
<td>Cum to Fall P3</td>
<td>138</td>
<td>P3GPA</td>
<td>P3GPA r = .62</td>
</tr>
<tr>
<td>Hein, et al&lt;sup&gt;15&lt;/sup&gt;</td>
<td>2017</td>
<td>Winter</td>
<td>Low</td>
<td>N</td>
<td>NR</td>
<td>384</td>
<td>P3GPA</td>
<td>P3GPA r = .5, NAPLEX r = .64</td>
</tr>
<tr>
<td>McDonough, et al&lt;sup&gt;16&lt;/sup&gt;</td>
<td>2017</td>
<td>Fall</td>
<td>Low</td>
<td>N</td>
<td>Cum to Fall P3</td>
<td>159</td>
<td>P3GPA</td>
<td>P3GPA r = .5</td>
</tr>
<tr>
<td>Naughton &amp; Friesner&lt;sup&gt;18&lt;/sup&gt;</td>
<td>2014</td>
<td>Spring</td>
<td>Low</td>
<td>NA</td>
<td>Y</td>
<td>108</td>
<td>NAPLEX</td>
<td>NAPLEX r = .59</td>
</tr>
<tr>
<td>Rudolph, et al&lt;sup&gt;19&lt;/sup&gt;</td>
<td>2019</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>1454</td>
<td>NAPLEX</td>
<td>NAPLEX r = .54</td>
</tr>
<tr>
<td>Scott, et al&lt;sup&gt;20&lt;/sup&gt;</td>
<td>2010</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>75 (each cohort)</td>
<td>P3GPA</td>
<td>P3GPA gr 1 r = .71, P3GPA gr 2 r = .46, P3GPA gr 3 r = .26</td>
</tr>
<tr>
<td>Souza, et al&lt;sup&gt;21&lt;/sup&gt;</td>
<td>2014</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>49</td>
<td>NAPLEX</td>
<td>NAPLEX r = .62</td>
</tr>
<tr>
<td>Verdell, et al&lt;sup&gt;23&lt;/sup&gt;</td>
<td>2018</td>
<td>Spring</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>150</td>
<td>P3GPA</td>
<td>P3GPA r = .64</td>
</tr>
<tr>
<td>Waskiewicz&lt;sup&gt;24&lt;/sup&gt;</td>
<td>2011</td>
<td>NR</td>
<td>Low</td>
<td>N</td>
<td>NR</td>
<td>33</td>
<td>P3GPA</td>
<td>P3GPA r = .5</td>
</tr>
<tr>
<td>Welch &amp; Karpen&lt;sup&gt;25&lt;/sup&gt;</td>
<td>2018</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td>P3GPA</td>
<td>P3GPA gr 1 (77) r = .46, P3GPA gr 2 (146) r = .57, NAPLEX gr 1 (66) r = .64, NAPLEX gr 2 (135) r = .66</td>
</tr>
</tbody>
</table>

Abbreviations: NR = not reported by study authors, NAPLEX = North American Pharmacist Licensure Exam, GPA = grade point average, PCOA = Pharmacy Curriculum Outcomes Assessment
was narrowed down per the exclusion criteria above of needing at least 10 or more studies comparing the outcome of interest. For this study, “PCOA scores” refers to the pre-APPE total scaled score provided by NABP using the scoring model described above. Any study that used pre-APPE PCOA scores was eligible for inclusion. Studies involving pre-APPE PCOA scores were eligible for participation as were studies involving single and/or multiple schools and colleges of pharmacy. Data regarding the number of schools included in each study will be discussed in the results section.

All study designs were eligible for inclusion in the review; however, if an intervention versus control group design had been used for the study, only the control group data were used in the analysis. This action was taken so that all results only contained data for the PCOA versus the outcome in question because of the limited number of active intervention studies regarding this topic. Specific studies that contained an intervention group are noted in the results section. The publication date of the study was not limited to ensure the largest dataset possible for each outcome. Study data had to be available in the form of a poster, abstract, journal article, or communication with the authors.

An attempt was made to identify and retrieve the entire population of published and unpublished literature that met the inclusion criteria for this study. The following electronic databases were searched from the start of the database through February 2019: PubMed, CINAHL, EMBASE, ProQuest Dissertations and Thesis (abstract limited), Academic Search Complete, and Google Scholar. Databases for the following pharmacy journals were also searched: Innovations in Pharmacy, American Journal of Health-System Pharmacy, Pharmacotherapy, Currents in Pharmacy Teaching and Learning, and the American Journal of Pharmaceutical Education. Reference lists of all included articles were also reviewed for inclusion. In general, search terms were Pharmacy, Curriculum, Outcomes, Assessment, OR PCOA using the truncation and nearness terminology based on the rules for the specific database being searched.

Abstracts were retrieved for all search hits from all the databases and reference lists. The two researchers screened all abstracts to first eliminate any clearly irrelevant study reports (ie, ones not discussing the PCOA). Reviewers were not blinded to study location. Full-text versions of papers, posters, and abstracts were retrieved for the remaining articles that were not excluded in the initial round. If there was ambiguity as to whether the article met inclusion criteria, it was included in the initial review. Article screening was completed using Rayyan QCRI (Qatar Computing Research Institute, rayyan.qcri.org) abstract review system. The two researchers then reviewed all retrieved articles to determine which ones would be included in the study. The two researchers had reached 95% agreement during both review sessions. Disagreements were discussed until a consensus was reached. Once articles were included, the two researchers then met to classify the articles into outcome categories to determine the final outcomes and articles for inclusion in the study. The final outcomes with more than 10 articles per outcome were P3 GPA and NAPLEX scores.

All coding on studies that were included after the initial screening (outcomes with 10 or more articles) was done by the two researchers, who had written and reviewed the initial coding document. Eligible studies were coded on multiple variables including the authors and year of publication, outcome of the study (PCOA vs outcome), number of programs involved in the study, type of institution (public/private), length of the PharmD program (four-year, three-year, 0-6/0-5), which PCOA blueprint was used (prior to 2016/after), which NAPLEX blueprint was used (prior to 2016/after), semester or quarter system, and time of year the examination was given. Data were also collected (where available) on stakes for the examination (low, medium, or high), whether study time was given, and whether incentives were provided to students, as well as how student GPAs were calculated. A low stakes examination was defined as one not used for progression or grades, whereas a “moderate stakes” examination used grades or required remediation but progression was not held back. A high stakes examination was one where progression was tied to the student’s score. Statistical data were then collected related to the outcome variables (means, standard deviations, sample sizes, correlation values). All fields on the form had a “not known or not listed” option.

A pilot coding was done with two articles to determine agreement on the coding document and to ensure all relevant data were collected from the articles. The researchers discussed any questions as they arose during the process, and a consensus was reached on coding for each study. Initial coding was done individually on paper by each coder, and then the consensus coding was transposed to an electronic spreadsheet for analysis. The coders had about a 95% agreement on the coding for each study on initial review. A 100% agreement was reached prior to spreadsheet creation. Effect size calculators were built into the electronic spreadsheet.

Data Analysis

Correlations were used as the effect size metric for all outcomes in the two meta-analyses done in this report. All the studies included in these meta-analyses used
correlations to compare the PCOA to the outcome being studied. All analyses were done in R-Studio, Version 1.1 (RStudio, Inc) and used an inverse variance weighted random effects model that included both the sampling variance and the between studies variance components in the study weights. A random effects model was chosen over fixed effects because the researchers theorized that there were various educational factors that could lead to there being no common effect sizes across all studies. These educational factors could include differences in program curricula, when the examination is given in the curriculum, and student motivation, among others. Mean effect sizes for the random weights were calculated using 95% confidence intervals for all studies. Estimates of Cochran’s Q and I2 were used to test for heterogeneity in the effect sizes. Even though the same sample of studies was used for different outcomes, the outcomes were treated as independent entities. This allows for the assumption of independence for the effect size estimates for each outcome.

A small number of studies were missing sample sizes and correlations and others were just missing sample sizes. For studies missing both sample sizes and correlations, the authors of the studies were contacted and provided the missing information. For the studies missing only sample sizes, sample sizes were obtained using data from the American Association of Colleges of Pharmacy, which publishes data on student enrollments annually. Given the small number of studies for each outcome, no moderator or sensitivity analysis was conducted.

Study quality was reviewed for each study included in the meta-analyses. No specific study quality tool was used for review. Studies were considered high quality if they documented information on when the examination was given, what stakes were provided for the examination, whether study time was provided to the students, whether incentives provided for the examination were mentioned, and whether they documented how student GPAs were calculated (P3 GPA meta-analysis only). These study qualities were chosen as these are items that have been discussed in the PCOA literature as having effects on examination performance. Studies that contained four or five qualities were considered “high quality.” Studies with three qualities were considered “moderate quality,” and those with two or fewer qualities were considered “low quality.”

The Egger regression test and the Henmi and Copas (hc) estimate of true average effect size were used to assess publication bias for each outcome. Given the small number of studies for each outcome, multiple publication bias tests were done for triangulation of the results. Results of these tests were used to determine the potential impact any publication bias might have on the meta-analysis results.

RESULTS

A PRISMA diagram of all identified studies for the review is presented in Figure 1. A total of 1669 articles were identified during the initial database search, plus one additional article from Google Scholar. After removal of duplicates, 1232 articles were added to the Rayyan QCRI abstract reviewer for initial review by the researchers. The initial screen using the inclusion/exclusion criteria led to the selection of 35 articles for full study review. Upon review of these articles, the researchers found that 10 of the articles contained multiple abstracts that had been presented at national meetings. Review of these articles showed 23 individual abstracts that met initial screen criteria for potential inclusion. The total number of full studies, abstracts, and posters that were reviewed for final potential inclusion in the meta-analyses was 48. The large number of exclusions at this stage resulted from search terms used in the electronic database search that were very liberal to ensure that all relevant studies were located. Of the 48 full articles, abstracts, and posters that were reviewed, 29 studies met inclusion/exclusion criteria to be considered for grouping by outcomes. Article exclusions are described in Figure 1.

Of the 29 studies that were considered for grouping, 13 unique articles met the criteria of having a group of 10 or more articles related to a particular outcome. The remaining 16 studies were too specific in scope to put into groupings as there was only one article with the outcome, did not contain the needed data for the meta-analysis, and the authors did not respond to a request for data, or the databases were duplicates of another article already included in the study (eg, a conference abstract describing data that was later included in a published manuscript). The 13 unique articles grouped into the following outcomes: NAPLEX=eight, and P3 GPA=eight. Table 1 contains a summary of study quality dimensions for each study included in the meta-analyses as well as a summary of the outcomes for each study and the number of students included in the sample.

Each of the outcomes (P3 GPA and NAPLEX) was reviewed in a separate meta-analysis. For each meta-analysis, data presented include $\tau^2$, tests for heterogeneity, and the model’s effect size, standard error, 95% confidence interval, and p value. Each model’s forest plot based on the respective effect sizes, shown in Figures 2 (for P3 GPA) and 3 (for NAPLEX) are provided for visual review of the data. Results were converted back to correlations for more consistent and understandable interpretation in the description below.
Eight different studies (Table 1) were used in the P3 GPA meta-analysis. Between these eight studies, 11 unique samples were available for analysis. \(\text{effect size} = .62, \text{SE} = .04, 95\% \text{CI} = .54-.71, \ p < .0001\). The \(\tau^2\) was .012, and the \(I^2\) showed moderate heterogeneity (61.32\%). The Q-test for heterogeneity was significant (Q(10) = 23.38, \(p = .009\)). The P3 GPA showed a moderate and significant correlation with PCOA scores (\(r = .55, 95\% \text{ confidence interval} = .54-.71\), with the P3 GPA accounting for 29\% to 50\% of the variability of PCOA scores. Based on the amount of heterogeneity found, the true effect size for the P3 GPA would widen but remain significant (effect size = .52-.92; \(r = .48-.73\)).

Eight different studies (Table 1) were used in the NAPLEX meta-analysis. Between these eight studies, 10 unique samples were available for analysis. \(\text{effect size} = .72, \text{SE} = .04, 95\% \text{CI} = .64-.80, \ p < .0001\). The \(\tau^2\) was .011 and the \(I^2\) showed moderate heterogeneity (69.5\%). The Q-test for heterogeneity was significant (Q(9) = 32.74, \(p = .0001\)). The PCOA showed a moderate and significant correlation with NAPLEX scores (\(r = .62, 95\% \text{ confidence interval} = .57-.66\), with the PCOA accounting for 32\%-44\% of the variability of NAPLEX scores. Based on the amount of heterogeneity found, the true effect size for the NAPLEX would widen but remain significant (effect size = .52-.92; \(r = .48-.73\)).

**Study Quality and Publication Bias Analysis**

Study quality was reviewed for all papers included in this study. Publication bias was reviewed for each outcome of the meta-analyses to determine whether effect sizes would be changed because of any bias found. Of the 13 studies that were included in the different meta-analyses done in this study, none included all four (or all five for P3 GPA studies) components of study quality mentioned in the methods section of this paper (when the
examination was given, stakes for the examination, study time provided, incentives provided, and how GPAs were calculated [if applicable]. One study had high study quality, three had moderate study quality, and nine had low study quality. Nine studies addressed stakes for the examination with eight of the nine using low stakes (not used for progression or grading) and one using moderate stakes, meaning it used the examination to determine a grade or require remediation but progression was not halted. Only three studies discussed whether their schools provided study time for students prior to the examination, with none of them providing time. Seven of the studies discussed incentives, with four indicating that some type of incentive was provided to students and three explicitly reporting that no incentives were provided to students. Only two of the studies described how GPAs were calculated in the study. Table 1 describes the breakdown of each of these quality indicators by study.

Using the Egger’s regression test, P3 GPA and NAPLEX all showed a nonsignificant $p$ value, which supports little evidence of publication bias. The results of the Egger’s test were confirmed using the Henmi and Copas Test with P3 GPA but not the NAPLEX. The effect sizes calculated by this test for the P3 GPA model (hc = .64 and this meta-analysis = .62) showed an estimated true effect size (hc) to be identical to the one found when the meta-analyses were run in this study. The Henmi and Copas (hc) test did show a trend towards a slight publication bias for the NAPLEX data (hc = .68 and this meta-analysis = .72), though the results were still significant. The bias shown does not appear to account for the total effect size, and the magnitude of the effect sizes would not be interpreted differently despite the potential

---

**Table 1**

<table>
<thead>
<tr>
<th>Study</th>
<th>Stake Description</th>
<th>GPA Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scott, 2010c</td>
<td>Low stakes</td>
<td>0.27 [0.03, 0.50]</td>
</tr>
<tr>
<td>Scott, 2010b</td>
<td>Low stakes</td>
<td>0.50 [0.27, 0.73]</td>
</tr>
<tr>
<td>Weich, 2018a</td>
<td>Low stakes</td>
<td>0.50 [0.27, 0.73]</td>
</tr>
<tr>
<td>McDonough, 2017</td>
<td>Low stakes</td>
<td>0.54 [0.30, 0.70]</td>
</tr>
<tr>
<td>Waskiewicz, 2011</td>
<td>Low stakes</td>
<td>0.55 [0.19, 0.91]</td>
</tr>
<tr>
<td>Garavalia, 2017</td>
<td>Low stakes</td>
<td>0.63 [0.50, 0.77]</td>
</tr>
<tr>
<td>Weich, 2018b</td>
<td>Low stakes</td>
<td>0.65 [0.49, 0.82]</td>
</tr>
<tr>
<td>Hein, 2017</td>
<td>Low stakes</td>
<td>0.69 [0.59, 0.79]</td>
</tr>
<tr>
<td>Gillette, 2017</td>
<td>Low stakes</td>
<td>0.73 [0.56, 0.90]</td>
</tr>
<tr>
<td>Verdel, 2018</td>
<td>Low stakes</td>
<td>0.76 [0.60, 0.92]</td>
</tr>
<tr>
<td>Scott, 2010a</td>
<td>Low stakes</td>
<td>0.89 [0.66, 1.12]</td>
</tr>
<tr>
<td>RE Model</td>
<td></td>
<td>0.62 [0.54, 0.71]</td>
</tr>
</tbody>
</table>

---

**Figure 2. Forest Plot of Effect Sizes Comparing Professional Year 3 Grade Point Average to Scores on the Pharmacy Curriculum Outcomes Assessment (PCOA)**

**Figure 3. Forest Plot of Effect Sizes Comparing Scores on the National Pharmacist Licensure Exam to Scores on the Pharmacy Curriculum Outcomes Assessment (PCOA)**
DISCUSSION

Results from this study show PCOA scores were moderately correlated with P3 GPAs accounting for 14% to 48% of the variability in PCOA scores after taking study heterogeneity into account. This analysis also showed that PCOA scores were moderately correlated with NAPLEX scores and accounted for 23% to 53% of the variability in NAPLEX scores after taking heterogeneity of the studies into account.

The results found in the meta-analyses done in this study are similar to the results found in some of the larger studies on the PCOA published to date. Giuliano and colleagues reviewed 142 second professional year students to determine predictors of student performance on the PCOA using correlations and linear regression. This study found that significant predictors of success on the PCOA were P1 GPA, institution, PCAT reading, being an “accommodator” according to the Health Professionals’ Inventory of Learning Styles (H-PILS), and not favoring “reading” as a learning preference. Higher P1 GPA and PCAT reading score were associated with higher PCOA scores, while the other indicators were associated with lower PCOA performance.

In 2017, Gillette and colleagues published a linear regression study using from 133 to 293 students per cohort (P1, P2, and P3) to determine predictors of performance on the PCOA. This study found that PCAT scores, Health Science Reasoning Test (HSRT) scores, and professional year cumulative GPA were the only consistent predictors of higher PCOA scores. Rudolph and colleagues conducted a large, multicenter study with six PharmD programs (N=1460) to determine the relationship between PCOA and NAPLEX scores. This study found that PCOA and NAPLEX scores were moderately correlated (r=.54), with PCOA total scores accounting for 30% to 33% of the variance in NAPLEX total scores.

Results from these three studies align with the results from our meta-analyses as we also showed that third professional year GPA was predictive of PCOA scores. Our analyses also showed that PCOA scores correlate with NAPLEX scores with similar variability (38%) as seen in the study by Rudolph and colleagues (30% to 33%), which was the largest sample of the studies analyzed (N=1454). Results from these meta-analyses should be interpreted with caution, however, because of the high amount of heterogeneity seen in most of the models. While the results are similar to the ones seen in the individual studies, larger studies with reliable methods are needed to support or refute these results.

While the results of our study are significant, there are some important limitations to note. First, our analyses relied only on published data rather than raw data. Of note, NABP does not currently provide any raw assessment or benchmarking data for PharmD programs to use for the PCOA or NAPLEX. As of this writing, only data made available from individual programs through publications and poster presentations can be used to assess the overall impact of variables on PCOA or PCOA on other variables of interest. Additionally, despite the multiple studies identified seeking to associate PCOA scores to outcomes of interest, there is a lack of clarity from NABP and ACPE on how to best use PCOA data across programs (or even a consensus that the PCOA should be used in this manner). Second, because of the inconsistencies in how the PCOA is administered across programs, we were only able to put a small number of studies into each category for analysis. This small number limited the number of results available for each category and may have introduced potential publication bias. In addition, because of the observational nature of the reports and reliance on data collected by individual study designers, our analysis included studies of relatively poor quality. This limited the ability to perform moderator studies to ascertain what variables affected the results. Finally, most studies were conducted within a single pharmacy program, which may limit the applicability of the results outside of those individual programs.

Despite the consistent results seen in our study, more studies are needed to determine whether factors such as incentives, examination timing, and study time effect the results seen as there were not enough studies to perform moderator analyses using these areas. More studies are also needed to review some of the other factors that have been found to be predictive in other studies such as student motivation, student learning styles, and the Health Sciences Reasoning Test.

CONCLUSION

Results from this first set of meta-analyses to be conducted on the PCOA data published to date does add to the literature currently available on this topic. Our study showed that third professional year GPA does correlate with PCOA results and that PCOA scores correlate with NAPLEX results. While there are significant limitations to interpreting the results of this analysis, we hope that programs are able to use the results to better understand
the role of the PCOA as a benchmark for progress through the PharmD curriculum.

REFERENCES