



# Early Pre-clerkship Clinical Skills Assessments Predict Clerkship Performance

Lindsay C. Strowd<sup>1,2</sup> · Hong Gao<sup>2</sup> · Donna M. Williams<sup>2</sup> · Timothy R. Peters<sup>2</sup> · Jennifer Jackson<sup>2</sup>

Accepted: 2 February 2022 / Published online: 1 March 2022

© The Author(s) under exclusive licence to International Association of Medical Science Educators 2022

## Abstract

**Introduction** Assessment of medical students' clinical skills (CS) remains a challenge. Little is known about early predictors of future CS performance. This study examines the relationship between students' pre-clerkship clinical skills (PCCS) performance and year 3 clerkship performance measures.

**Methods** The authors performed a retrospective analysis of four medical student cohorts who matriculated between 2014 and 2017 and participated in a longitudinal pre-clerkship CS curriculum. A total of 440 students were included in the analyses. Students' clinical skills were assessed through a series of PCCS exams, each consisting of a single standardized patient encounter. First-year PCCS exams assessed history taking, physical examination, professionalism, and communication skills; second-year PCCS exams also assessed clinical documentation and clinical reasoning skills. Evaluators assigned a grade of "satisfactory," "borderline," or "unsatisfactory" for each skill set. Regression analyses compared year 3 performance outcomes between students with one or more "unsatisfactory" or "borderline" PCCS skill set grades and students assessed as "satisfactory" for all PCCS skill set assessments.

**Results** Thirty-two percent ( $n = 140$ ) of the 440 students had at least one borderline or unsatisfactory (US) PCCS skill set grade. These students performed significantly worse on year 3 National Board of Medical Examiner subject exams, workplace-based clinical performance evaluations, and overall year 3 performance compared to students who passed all PCCS exam components. In addition, a higher percentage of students with PCCS performance deficiencies failed the United States Medical Licensing Examination Step 2 CS exam on the first attempt versus students who passed all PCCS exam components.

**Conclusions** PCCS exam performance at our institution aligned with future student performance on multiple year 3 clerkship outcome measures. This pre-clerkship performance data can be used to identify at-risk students who would benefit from additional resources to achieve competency in the clerkship environment and future medical training.

**Keywords** Clinical skills assessment · Deliberate practice · Clerkship performance

## Introduction

Medical schools are tasked with ensuring students master knowledge and display competency in core clinical skills required for the practice of medicine. Such skills are encompassed by the American Association of Medical Colleges' (AAMC) Core Entrustable Professional Activities

(EPAs) for Entering Residency—a framework designed to guide curriculum development and competency-based assessment of medical students before their transition to residency [1]. Although most medical schools administer internally developed observed structured clinical examinations (OSCE) to assess 3<sup>rd</sup>- or 4<sup>th</sup>-year student's clinical skills [2], experts have expressed concerns about variability in passing standards among schools, the ongoing use of norm-referenced standards in which students are compared to one another rather than achievement of a competency standard, and reluctance among school leadership to fail under-performing students [3–7]. Beginning in 2004, the United States Medical Licensing Exam (USMLE) Step 2 Clinical Skills (CS) Exam was implemented as a national standardized exam to assess allopathic medical students'

✉ Lindsay C. Strowd  
lchaney@wakehealth.edu

<sup>1</sup> Department of Dermatology, Wake Forest University School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA

<sup>2</sup> Wake Forest University School of Medicine, Winston-Salem, NC, USA

clinical skills performance for licensing purposes. At many medical schools, a passing score on the USMLE Step 2 CS exam was a graduation requirement, and many residency training programs used performance on this exam as a selection criterion for applicants [8]. Following the discontinuation of the Step 2 CS exam in early 2021, individual medical schools have increased responsibility to reliably assess and attest to their students' clinical skills competence [9].

CS assessments that predict students' future performance provide validity for these assessments [10]. Direct observation of students' clinical skills in the clerkship years is inconsistent and often limited [11]. Those with preexisting CS performance deficiencies are at risk for ongoing performance difficulties. Therefore, early identification of these students and their specific area(s) of skill deficiency is important for remediation and skill improvement before clinical immersion and subsequent high-stakes testing. Studies evaluating the relationship between pre-clerkship clinical skills (PCCS) performance outcomes and student performance later in training are limited, and predicting which students will struggle in the clerkship years of medical school remains challenging, given the complexity of clinical competencies and their assessment [5, 12]. A few studies have demonstrated that students who underperform on pre-clerkship OSCEs also underperform on OSCEs later in their medical school curriculum [13–16]. Investigators at the Edward Hébert School of Medicine of the Uniformed Services University showed that students' PCCS performance predicted workplace-based clinical performance in the clerkships, clerkship National Board of Medical Examiner (NBME) subject exam scores [17], USMLE Step 2 CS performance [18, 19], and internship clinical performance [19, 20]. Beyond these studies, there is little literature examining correlation of students' PCCS performance with subsequent performance in training.

In this study, our goal was to address this research gap by investigating the relationship between medical students' performance on PCCS assessments and performance outcomes in the third-year clerkships of medical school, hypothesizing that students with PCCS performance deficiencies will demonstrate lower performance in the third year of medical school. Specifically, we sought to determine if students with one or more deficiencies on PCCS exams and students with no PCCS deficiencies differed with regard to year 3 NBME subject exam scores, workplace-based clinical performance evaluations, or Honors clerkship grade designation. As secondary outcomes, we also sought to characterize the frequency of PCCS exam performance deficiencies with respect to specific clinical skill sets and to determine if students with PCCS exam deficiencies were at higher risk of failing the USMLE Step 2 CS exam than students with no PCCS deficiencies.

## Materials and Methods

### Subjects, Setting, and Study Design

This study was a single-institution, observational cohort study of four consecutive medical student classes who matriculated to Wake Forest School of Medicine (WFSM) between 2014 and 2017. We collected performance outcome data for these students during their first 3 years of medical school (i.e., 2014 to 2020). Students were excluded from data analysis only if they did not complete the entire third-year clerkship curriculum during the study period.

### Pre-clerkship Clinical Skills Course Structure

In addition to basic science coursework, all students at WFSM are required to participate in and pass a longitudinal PCCS course before transitioning to the third-year clerkships. This PCCS course is designed using deliberate practice principles [21], consisting of twenty-six class sessions distributed longitudinally throughout the 2-year pre-clerkship curriculum. Small groups of students are facilitated by 2 clinical faculty instructors. Class sessions consist of iterative, directly observed student practice of doctor-patient communication and interpersonal skills, history taking skills, physical examination skills, and professionalism tasks during standardized patient (SP) and hospitalized patient encounters. Students receive individualized formative performance feedback on patient encounter skills from their faculty instructors, SPs, and small-group peers; faculty also provide individualized feedback on students' clinical documentation skills and clinical reasoning skills.

### Pre-clerkship Clinical Skills Course Assessments

Students are assessed on these skill sets through a series of 6 to 7 clinical performance examinations distributed over the two pre-clerkship years; each exam consists of a single, 40-min SP encounter (for students who matriculated in 2017, the total number of second-year PCCS exams decreased from 3 to 2, for a total number of 6 PCCS exams, due to shortening of the pre-clerkship curriculum duration from 22 to 18 months). First-year PCCS exams assess students' performance of professionalism tasks, communication and interpersonal skills, history taking skills, and physical examination skills. In addition to these skill sets, second-year PCCS exams assess clinical reasoning and clinical documentation skills on a written history

and physical document. PCCS faculty instructors serve as the evaluators for the PCCS exams; except for the first PCCS exam of year 1 (which is formative), PCCS faculty instructors never evaluate their own small-group students.

The grading rubric for each PCCS exam consists of a list of directly observable tasks and behaviors for each skill set assessed (Appendix 1). Evaluators assess students' performance of each task as either "performed" or "did not perform" for communication/interpersonal skills, professionalism tasks, history taking skills, clinical documentation skills, and clinical reasoning skills; physical examination tasks are assessed as either "correctly and completely performed," "performed but incompletely or incorrectly," or "not performed." The grading rubric for each skill set also includes an item assessing the student's overall performance of that skill set, assessed by the evaluator as either "satisfactory" or "unsatisfactory."

### **Pre-clerkship Clinical Skills Course Grade Determinations**

Students are given a separate grade (satisfactory or US) for each of the 4 to 6 skill sets tested on each of the first- and second-year exams, respectively. For all students whose PCCS performance of any skill set is assessed by evaluators as unsatisfactory (US), the student's performance is independently reviewed by a member of the CS Review Team—a core group consisting of 3 experienced PCCS faculty and the 2 PCCS co-course directors. This team meets to discuss and establish, by consensus, satisfactory versus US performance criteria, a priori, for each skill set on each PCCS exam and applies those criteria to make standardized grade determinations for students under review (i.e., criterion-referenced standard setting). The PCCS co-course directors also use these criteria to screen the faculty evaluation forms for all students as an additional standardization measure; any students with performance concerns identified by this method are also reviewed by the CS Review Team. In addition, the CS Review Team identifies "borderline satisfactory performers"—students whose performance is low for stage of training but meets criteria for satisfactory performance.

Remediation is required for all students with US performances on one or more skill sets and is encouraged (but not required) for students with borderline performances. Reassessment of students' skills occurs on the next scheduled PCCS exam. Students must demonstrate satisfactory performance in all skill sets by the end of each academic year in order to pass the PCCS course, and passing the course is required to progress to the next training year. Individual students' PCCS performance data is not provided to year 3 clerkship administrators or faculty.

### **Year 3 Clinical Clerkship Structure and Assessment**

All WFSM students participate in the same core clinical clerkships during their third year of the curriculum: family medicine, internal medicine, emergency medicine, obstetrics and gynecology, pediatrics, psychiatry, neurology, and surgery. Students receive a clinical evaluation score for each clerkship comprised of workplace-based evaluations assessing multiple domains of students' clinical skills performance, including history taking and physical examination skills, communication skills, clinical reasoning skills, oral presentation skills, clinical teamwork skills, and professionalism. Final third-year clerkship grades are calculated based on students' clinical evaluation scores, National Board of Medical Examiners (NBME) subject exam score, and other clerkship-specific performance measures. Of note, our students take the NBME subject exam for Emergency Medicine (EM) as part of their EM clerkship grade, though this exam was not specifically designed for third-year medical students. Students are given clerkship grades of Honors, High Pass, Pass, Low Pass, or Fail.

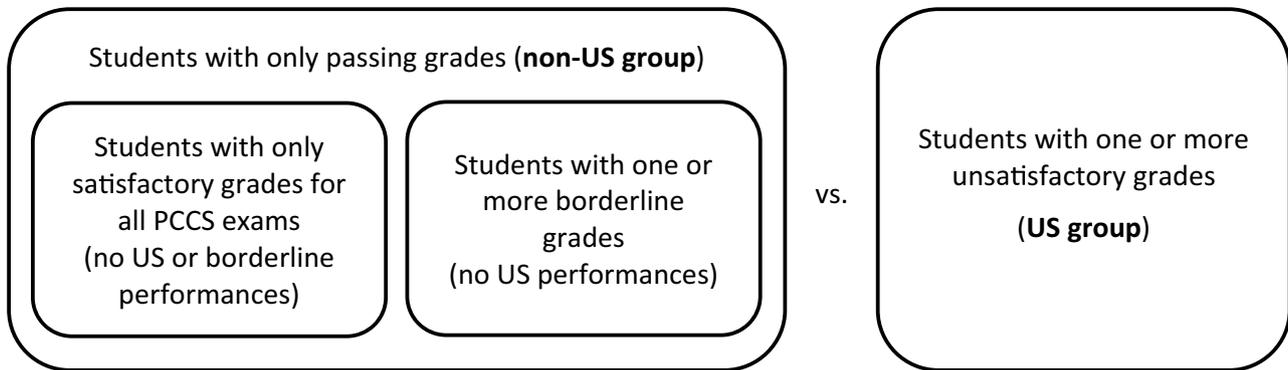
### **Data Analysis**

For the purposes of this analysis, a CS performance deficiency was defined as an US or borderline grade on any skill set for any PCCS exam in the first or second year of the PCCS course. Students with at least one US performance on any PCCS exam were designated as the US group, and students without any US performances on PCCS exams were designated as the non-US group (note: this group included students with "borderline" performances). Students with at least one CS performance deficiency of either type (i.e., US or borderline) on any PCCS exam were designated as the US/borderline (US/B) group, and students without any CS performance deficiencies on any PCCS exam were designated as the all-passing group (Fig. 1).

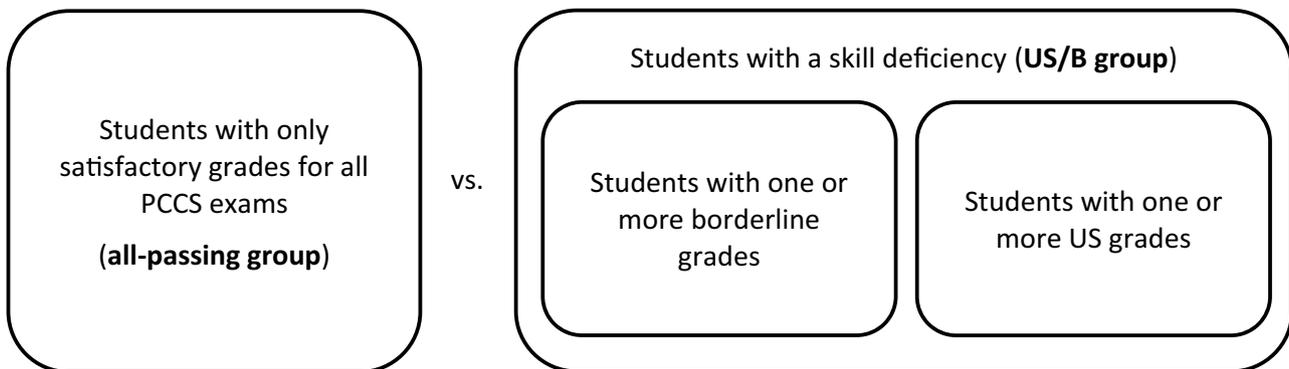
We characterized the number and distribution of CS performance deficiencies among students in the US/B group by calculating the total number and percentage of students with CS performance deficiencies and the total number of CS performance deficiency instances using descriptive statistics. We categorized students with CS performance deficiencies by clinical skill set and by number of total deficiency instances per student and calculated the total number and percentage of students in each category. We also calculated the total number and percentage of CS performance deficiency instances among students in each of these categories to examine their distribution.

We performed analyses comparing the US student group to the non-US group to examine the relationship between PCCS performance and performance outcomes within each third-year clerkship, including students' clinical

**Analysis 1:** students with only passing grades vs. students with one or more unsatisfactory (US) grades:



**Analysis 2:** students with no skill deficiencies vs. students with one or more skill deficiencies:



**Fig. 1** Groupings of students used for analyses, according to PCCS grades assigned

evaluation score and NBME subject examination score. For these analyses, students' raw third-year clinical evaluation scores and NBME subject exam scores were converted to *Z* scores for easier comparison between scores (for NBME subject exam scores, we used a mean of 70 and standard deviation of 8). We assessed for significant group differences (US vs. non-US) in achieving Honors for the clerkship grade and for performance on the USMLE Step 2 CS exam. The same analyses were performed to compare the US/B group to the all-passing student group. Finally, we performed an additional analysis on the subset of students who received a PCCS exam performance deficiency on one of the first-year PCCS exams (defined as the first-year US/B group). All data analyses were performed using SAS Enterprise Guide 7.1 for Windows (Cary, NC).

Institutional IRB approval was obtained for this study (IRB00043836), and students were provided opt-out consent for participation. Identifying information was

removed from records, and students were assigned a unique study number to protect individual identity.

## Results

### Student PCCS Exam Performance

A total of 440 students were included in the data analysis. Three hundred students (68%) passed all skill set components of all PCCS exams (i.e., the all-passing group). One hundred and forty students (32%) had at least one CS performance deficiency on a PCCS exam, defined as having either an unsatisfactory or a borderline performance in one section of the exam (i.e., the US/borderline group, Table 1). Within the US/B group, 73 students (52%) had only 1 skill deficiency, while 67 students (48%) had 2 or more total CS

**Table 1** Distribution of pre-clerkship clinical skills exam grades among medical students at Wake Forest School of Medicine, 2014–2018 (*n* = 440)

| Year in training                     | Students with only satisfactory grades for all PCCS exams, <i>n</i> (%) <sup>a</sup> | Number of students with 1 or more deficiencies on PCCS exams, <i>n</i> (%) <sup>b</sup> |
|--------------------------------------|--|---|
| First-year PCCS exams                | 385 (87%)  | 55 (13%)  |
| Second-year PCCS exams               | 333 (76%)  | 107 (24%)   |
| Total across both years <sup>c</sup> | 300 (68%)  | 140 (32%)   |

Abbreviations: PCCS, pre-clerkship clinical skills

<sup>a</sup>This represents the all-passing group (i.e., no borderline or unsatisfactory performances)

<sup>b</sup>This represents the US/borderline group

<sup>c</sup>The total number of unique students with an US or borderline grade was 140. Some students in the US/borderline group had skill deficiencies in both the first and second training years of the PCCS course, thus leading to the above total numbers of US/borderline students in each year of training

performance deficiencies, comprising 15% of the total student cohort.

### Distribution of CS Deficiencies by Skill Set

Among skill sets, deficiencies in history taking skills occurred with the highest frequency, accounting for 38% of all US/B instances, followed by physical examination skills (27%), clinical reasoning skills (23%), clinical documentation skills (7%), communication/interpersonal skills (5%), and professionalism (0.4%). The breakdown of CS deficiencies by skill set and curriculum year is presented in Table 2.

### PCCS Performance and Year 3 Performance

#### Comparison Between US and Non-US Groups (Analysis 1)

Data showed a consistent trend of students in the US group achieving lower clinical evaluation scores than in the non-US group across clerkships, and this group performance difference achieved statistical significance for all clerkships except for Psychiatry. For NBME subject exam performance,

we noted a similar pattern: the US group scored significantly worse than the non-US group across all clerkships. Odds of achieving Honors were significantly lower in the US group in all clerkships except for Emergency Medicine. All of the related analysis results are presented in Table 3.

#### Comparison Between US/B and All-passing Groups (Analysis 2)

When students with borderline performances were added to the US student cohort (i.e., US/B group), the analysis showed significant differences between the US/B group and the all-passing group in clinical evaluation scores (*P* < 0.05) for all clerkships except Psychiatry. NBME subject exam performance and Honors designation were significantly different between the US/B and all-passing groups across all clerkships (Table 4).

#### First-Year PCCS Performance and Year 3 Performance

When comparing only those students who had a PCCS exam performance deficiency within the first year of the PCCS course (i.e., first-year US/B group) with students

**Table 2** Distribution of clinical skill set deficiencies among pre-clerkship medical students with one or more performance deficiencies on first- and second-year pre-clerkship clinical skills exams (*n* = 140), Wake Forest School of Medicine, 2014–2018<sup>a</sup>

| Number of students with CS deficiencies, by skill set and year of training: | First-year PCCS exams |                             |   | Second-year PCCS exams |                             |  | Total across both years of training |
|---|-----------------------|-----------------------------|---|------------------------|-----------------------------|--|-------------------------------------|
|   | US ( <i>N</i> = 31)   | Borderline ( <i>N</i> = 32) | 1 <sup>st</sup> year total ( <i>N</i> = 55) | US ( <i>N</i> = 70)    | Borderline ( <i>N</i> = 54) | 2 <sup>nd</sup> year total ( <i>N</i> = 107) |                                     |
| Communication and interpersonal skills                                      | 0                     | 5                           | 5   | 3                      | 4                           | 7  | 12                                  |
| Professionalism tasks   | 0                     | 1                           | 1   | 0                      | 0                           | 0  | 1                                   |
| History taking skills   | 28                    | 24                          | 52  | 16                     | 31                          | 47   | 99                                  |
| Physical examination skills   | 12                    | 9                           | 21  | 24                     | 26                          | 50   | 71                                  |
| Clinical documentation skills <sup>b</sup>                                  | N/A                   | N/A                         | N/A   | 13                     | 6                           | 19   | 19                                  |
| Clinical reasoning skills <sup>b</sup>                                      | N/A                   | N/A                         | N/A   | 47                     | 13                          | 60   | 60                                  |

Abbreviations: CS, clinical skills; PCCS, pre-clerkship clinical skills; US, unsatisfactory performers; N/A, not applicable

<sup>a</sup>The total number of unique students with an US or borderline grade was 140. Some students in the US/borderline group had skill deficiencies in multiple skill domains, thus leading to the above total numbers of US/borderline students for each year of the PCCS course

<sup>b</sup>Skill set summatively assessed only in the second training year of the PCCS curriculum

**Table 3** Comparison of third-year performance outcomes between medical students with one or more unsatisfactory performances on pre-clerkship clinical skills exams (US group,  $n=93$ ) and students with no unsatisfactory performances (non-US group,  $n=347$ ), Wake Forest School of Medicine, 2014–2020

| Clerkship                 | Clinical evaluation score (Z score, SD) |              |         | NBME subject exam score (Z score, SD) |              |         | Odds of achieving Honors clerkship grade (Wald chi-square statistic) |              |         |
|---------------------------|---|--------------|---------|---------------------------------------|--------------|---------|--|--------------|---------|
|                           | US group                                | Non-US group | P value | US group                              | Non-US group | P value | US group   | Non-US group | P value |
| Emergency Medicine        | -0.05 (1.14)                            | 0.23 (0.99)  | 0.02    | -0.34 (1.00)                          | 0.16 (1.03)  | <0.0001 | 8.99%  | 14.08%       | 0.21    |
| Family Medicine           | -0.06 (0.93)                            | 0.20 (0.97)  | 0.02    | -0.16 (1.05)                          | 0.48 (1.07)  | <0.0001 | 18.28%   | 38.29%       | 0.0004  |
| Internal Medicine         | -0.24 (0.91)                            | 0.22 (0.92)  | <0.0001 | 0.26 (1.34)                           | 0.80 (1.19)  | 0.0002  | 8.7%   | 19.48%       | 0.02    |
| Pediatrics                | -0.28 (0.94)                            | 0.06 (1.10)  | 0.006   | 0.38 (1.29)                           | 0.79 (1.22)  | 0.004   | 17.20%   | 33.14%       | 0.003   |
| Obstetrics and Gynecology | -0.17 (0.90)                            | 0.07 (0.99)  | 0.04    | 0.004 (1.10)                          | 0.59 (1.07)  | <0.0001 | 14.89%   | 34.10%       | 0.0005  |
| Neurology                 | -0.21 (1.09)                            | 0.05 (1.00)  | 0.03    | 0.02 (1.21)                           | 0.73 (1.06)  | <0.0001 | 12.77%   | 31.43%       | 0.0005  |
| Psychiatry                | -0.10 (1.03)                            | -0.01 (1.16) | 0.49    | 0.44 (1.54)                           | 1.14 (1.41)  | 0.0003  | 29.79%   | 54%          | <0.0001 |
| Surgery                   | -0.25 (0.99)                            | 0.10 (0.99)  | 0.003   | 0.04 (1.11)                           | 0.54 (1.10)  | <0.0001 | 8.51%  | 19.48%       | 0.015   |

Abbreviations: *SD*, standard deviation; *NBME*, National Board of Medical Examiners; *US group*, students with 1 or more unsatisfactory performances on pre-clerkship clinical skills exams; *non-US group*, students with no unsatisfactory performances on pre-clerkship clinical skills exams

who did not have PCCS exam performance deficiencies in their first year of the PCCS course, clinical evaluation scores were different between these two groups for six of the eight clerkships, while NBME performance was different between these two groups across all clerkships except for Emergency Medicine. Odds of achieving Honors designation were significantly lower in the first-year US/B group for five of the eight clerkships (Table 5).

For the analysis examining group differences in PCCS exam performance and USMLE Step 2 CS exam performance, due to the timing and impact of the novel coronavirus (COVID-19) pandemic on the USMLE Step 2 CS exam, we only had USMLE Step 2 CS performance data for 3 of the 4 student classes ( $N=311$ , 70%). A higher percentage of students in the US/B group failed the USMLE Step 2 CS exam (11%) compared to the PCCS passing student group

**Table 4** Comparison of third-year performance outcomes among students with one or more pre-clerkship clinical skill deficiencies. Unsatisfactory/borderline group of students from first and second years ( $n=140$ ) versus students with no pre-clerkship clinical skills deficiencies (all-passing group,  $n=300$  students), Wake Forest School of Medicine, 2014–2020. Analysis of data included students' performance on all pre-clerkship clinical skills exams throughout the preclinical curriculum

| Clerkship               | Clinical evaluation score (Z score, SD) |                   |         | NBME subject exam score (Z score, SD) |                   |         | Odds of achieving Honors clerkship grade (Wald chi-square statistic) |                       |         |
|-------------------------|---|-------------------|---------|---------------------------------------|-------------------|---------|--|-----------------------|---------|
|                         | US/B group                              | All-passing group | P value | US/B group                            | All-passing group | P value | US/B group (%)   | All-passing group (%) | P value |
| Emergency Medicine      | -0.05 (1.00)                            | 0.28 (0.98)       | 0.001   | -0.10 (1.06)                          | 0.30 (0.98)       | 0.0002  | 8.09%  | 15.31%                | 0.04    |
| Family Medicine         | -0.05 (0.97)                            | 0.24 (0.94)       | 0.003   | -0.03 (1.01)                          | 0.52 (1.09)       | <0.0001 | 22.54%   | 39.53%                | 0.0005  |
| Internal Medicine       | -0.23 (0.92)                            | 0.28 (0.90)       | <0.0001 | 0.25 (1.22)                           | 0.89 (1.20)       | <0.0001 | 7.80%  | 21.67%                | 0.0006  |
| Pediatrics              | -0.29 (1.21)                            | 0.12 (0.98)       | 0.0001  | 0.39 (1.16)                           | 0.86 (1.26)       | 0.0002  | 22.54%   | 33.22%                | 0.02    |
| Obstetrics & Gynecology | -0.17 (0.90)                            | 0.11 (1.00)       | 0.005   | 0.13 (1.05)                           | 0.63 (1.08)       | <0.0001 | 21.83%   | 33.89%                | 0.01    |
| Neurology               | -0.20 (1.03)                            | 0.08 (0.99)       | 0.006   | 0.07 (1.10)                           | 0.81 (1.07)       | <0.0001 | 18.18%   | 31.89%                | 0.003   |
| Psychiatry              | -0.10 (1.16)                            | -0.01 (1.16)      | 0.49    | 0.44 (1.54)                           | 1.14 (1.40)       | <0.0001 | 29.79%   | 54%                   | <0.0001 |
| Surgery                 | -0.25 (1.00)                            | 0.10 (0.99)       | 0.003   | 0.04 (1.11)                           | 0.54 (1.10)       | <0.0001 | 8.51%  | 19.48%                | 0.002   |

Abbreviations: *SD*, standard deviation; *NBME*, National Board of Medical Examiners; *US/B group*, students with 1 or more unsatisfactory or borderline performances on pre-clerkship clinical skills exams

**Table 5** Comparison of third-year performance with first-year PCCS performance only. Outcomes among students with one or more first-year clinical skills exam deficiencies (first-year unsatisfactory/borderline group,  $n=55$ ) versus students with no first-year clinical skills

exam deficiencies (all-passing group,  $n=385$  students), Wake Forest School of Medicine, 2014–2020. Analysis of data included students’ performance on clinical skills exams from the first year of preclinical instruction/assessment only

| Clerkship               | Clinical evaluation score (Z score, SD) |                   |                | NBME subject exam score (Z score, SD) |                   |                | Odds of achieving Honors clerkship grade (Wald chi-square statistic) |                       |                |
|-------------------------|---|-------------------|----------------|---------------------------------------|-------------------|----------------|--|-----------------------|----------------|
|                         | Year 1 US/B group                       | All-passing group | <i>P</i> value | Year 1 US/B group                     | All-passing group | <i>P</i> value | Year 1 US/B group (%)  | All-passing group (%) | <i>P</i> value |
| Emergency Medicine      | −0.31 (1.25)                            | 0.23 (0.97)       | 0.0003         | 0.03 (1.03)                           | 0.06 (1.04)       | 0.83           | 1.96%  | 14.51%                | 0.04           |
| Family Medicine         | −0.20 (0.95)                            | 0.20 (0.95)       | 0.003          | −0.42 (0.94)                          | 0.46 (1.07)       | <0.0001        | 21.43%   | 35.92%                | 0.04           |
| Internal Medicine       | −0.39 (0.98)                            | 0.19 (0.91)       | <0.0001        | −0.28 (1.07)                          | 0.83 (1.20)       | <0.0001        | 5.45%  | 18.91%                | 0.02           |
| Pediatrics              | −0.46 (1.51)                            | 0.05 (0.99)       | 0.0008         | −0.22 (1.09)                          | 0.84 (1.21)       | <0.0001        | 25%  | 30.49%                | 0.40           |
| Obstetrics & Gynecology | −0.14 (0.89)                            | 0.04 (0.99)       | 0.20           | −0.29 (0.98)                          | 0.58 (1.07)       | <0.0001        | 25%  | 30.75%                | 0.38           |
| Neurology               | −0.32 (1.09)                            | 0.04 (0.99)       | 0.011          | −0.46 (1.07)                          | 0.73 (1.06)       | <0.0001        | 17.54%   | 28.94%                | 0.08           |
| Psychiatry              | −0.10 (1.01)                            | −0.02 (1.15)      | 0.62           | −0.06 (1.38)                          | 1.16 (1.41)       | <0.0001        | 28.07%   | 51.94%                | 0.001          |
| Surgery                 | −0.27 (0.96)                            | 0.07 (1.00)       | 0.04           | −0.19 (1.01)                          | 0.53 (1.11)       | <0.0001        | 12.50%   | 17.83%                | 0.37           |

Abbreviations: *SD*, standard deviation; *NBME*, National Board of Medical Examiners; *US/B group*, students with 1 or more unsatisfactory or borderline performances on pre-clerkship clinical skills exams

(3%), and this difference was statistically significant (Wald chi-square statistic 6.71,  $P=0.0096$ ).

### Discussion

PCCS exam performance at our institution aligned with students’ future performance on multiple performance measures in the third-year clinical clerkships, with low PCCS exam performers demonstrating lower clerkship workplace-based clinical evaluation scores, lower NBME scores, lower odds for achieving Honors for overall clerkship grade, and a higher risk of failing USMLE Step 2 CS. Most low-performing students had only 1 CS deficiency over the course of the PCCS course, with a significant percentage of the total number of CS deficiencies accounted for by a subset of students. Over one-third of low performers were identified within the first year of the course. The results of our study suggest that a longitudinal PCCS course with robust assessment methods can predict which students are likely to underperform in their clerkships based on early PCCS exam performances. We found student performance in a rigorous clinical skills assessment program correlates strongly with future performance in both real and simulated patient encounters. This supports the development of clinical skills assessments that reliably predict performance in more advanced clinical practice environments as an achievable goal. Development of a robust local CS assessment program is particularly

important given the recent elimination of the USMLE Step 2 CS exam as a national, standardized assessment students’ clinical skills competencies [9].

CS assessments that predict students’ future performance provide validity for these assessments [10]. Our findings are consistent with previous studies assessing the relationship between PCCS exam performance outcomes and future clerkship performance, including the work of LaRoche and colleagues who found PCCS exam performance predicted third-year clerkship NBME exam performance and clerkship final grades [17]. Our findings are also consistent with previous studies showing PCCS exam performance predicts later high-stakes OSCE performance [13–16, 18, 19]. In contrast to these studies, in which PCCS exams involved multiple SP encounters, we were able to identify predictors of future clerkship performance from PCCS exams involving a single SP encounter. Although previous research has shown improved reliability of CS assessments involving multiple patient encounters [5, 22], our findings suggest that it is possible to identify low-performing students with robust assessments involving only a single SP encounter. This has potential implications for decreased cost and resource utilization, which can be substantial for such exams [23].

Another important difference between our study and previous studies is the timing of PCCS performance data collection. In previous studies, PCCS exams were administered at the end of the pre-clerkship period. In contrast, we measured PCCS exam performance outcomes serially throughout the pre-clerkship period, beginning a few months into the

first year of the PCCS course, and found that even early PCCS exam performance data could predict later student performance. The ability to identify struggling learners early in their training has significant implications for students whose skill deficiencies might otherwise go undetected. Low-performing medical students have lower levels of self-efficacy and higher levels of anxiety and frustration, compared to higher-performing peers [24]. Attending to students' self-efficacy early in their training is important, given the impact of self-efficacy on academic achievement in medical school [25, 26]. The later in training a struggling learner is identified, the higher the likelihood that they will develop exhaustion and discouragement, requiring more extensive make-up and negatively impacting the student's ability to modify or make changes to their self-regulated learning behaviors [24]. Clinical skills assessment efforts should attempt to identify individual students' specific deficiencies early in training, to optimize remediation efforts and future clinical skills performance. A PCCS assessment program that can identify students early who are likely to underperform in the clinical clerkships allows for development of effective, targeted instructional programs to enhance later performance.

While our study provides important findings related to early predictors of student clerkship performance, it did not evaluate the underlying etiology of student performance deficiencies on PCCS exams. We also did not evaluate the relationship between PCCS outcomes and clinical performance following undergraduate medical training. Additional studies are needed to further explore these aspects of medical student training.

## Limitations

This was a single-institution study, and our results may not be generalizable to other institutions where the PCCS course structure, content, instructional methods, and assessment methods may differ. Despite intensive faculty development and the use of standardized grading rubrics, faculty evaluators' identification of US or borderline performances during PCCS exams involves some degree of subjectivity. Student performance evaluations during the third-year clerkships also involve subjective assessments of student performance. Each clerkship at our institution employs slightly different criteria to calculate students' final clerkship grades, and the NBME subject exam score contribution to the overall clerkship grade varies among clerkships. The number of third-year clinical evaluations completed for a given student varies among students and is comprised of resident, fellow, and faculty evaluators. We attempted to account for this variability by performing and presenting analysis of multiple, separate measures of clerkship performance, including NBME subject exam scores, clinical evaluation scores, and Honors designation.

## Conclusions

Robust clinical skills assessments administered early in the pre-clinical curriculum can predict future clinical performance of medical students. Educators should use these assessment tools to identify students at risk for later performance deficiencies early in the curriculum when there is more opportunity to provide additional resources and remediation to these students.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40670-022-01519-8>.

**Author Contribution** All authors contributed to the study design, data analysis and interpretation, and drafting and revision of the manuscript.

**Availability of Data and Material** Not applicable.

**Code Availability** Not applicable.

## Declarations

**Ethics Approval** This study was approved by the Wake Forest School of Medicine IRB, IRB00043836.

**Conflict of Interest** The authors declare no competing interests.

**Previous Presentations** A portion of this data was submitted and accepted for poster presentation at the Association of American Medical Colleges Lean Serve Lead 2020 Virtual Conference.

## References

1. Obeso V, Brown D, Aiyer M, Barron B, Bull J, et al. Core EPAs for Entering Residency Pilot Program. Toolkits for the 13 core entrustable professional activities for entering residency. Washington, DC: Association of American Medical Colleges; 2017. Available from: <https://www.aamc.org/what-we-do/mission-areas/medical-education/cbme/core-epas/publications>. Accessed 26 February 2021.
2. Association of American Medical Colleges. Number of medical schools requiring final SP/OSCE examination: 2011 - 2012 through 2015 - 2016. <https://www.aamc.org/data-reports/curriculum-reports/interactive-data/sp/osce-final-examinations-us-medical-schools>. Accessed 24 February 2021.
3. Ecker DJ, Milan FB, Cassese T, Farnan JM, Madigosky WS, Massie FS Jr, Mendez P, Obadia S, Ovitsh RK, Silvestri R, Uchida T, Daniel M. Step up-not on-the step 2 clinical skills exam: directors of Clinical Skills Courses (DOCS) oppose ending step 2 CS. *Acad Med*. 2018;93(5):693–8. <https://doi.org/10.1097/ACM.0000000000001874>. PMID: 28834843.
4. Setyonugroho W, Kennedy KM, Kropmans TJ. Reliability and validity of OSCE checklists used to assess the communication skills of undergraduate medical students: a systematic review [published online ahead of print June 27, 2015]. *Patient Educ Couns*. <https://doi.org/10.1016/j.pec.2015.06.004>.
5. Brannick MT, Erol-Korkmaz HT, Prewett M. A systematic review of the reliability of objective structured clinical examination scores. *Med Educ*. 2011;45(12):1181–9. <https://doi.org/10.1111/j.1365-2923.2011.04075.x>. Epub 2011 Oct 11 PMID: 21988659.

6. Guerrasio J, Furfari KA, Rosenthal LD, Nogar CL, Wray KW, Aagaard EM. Failure to fail: the institutional perspective. *Med Teach*. 2014;36(9):799–803. <https://doi.org/10.3109/0142159X.2014.910295>. Epub 2014 May 20 PMID: 24845780.
7. Hauer KE, Teherani A, Kerr KM, Irby DM, O'Sullivan PS. Consequences within medical schools for students with poor performance on a medical school standardized patient comprehensive assessment. *Acad Med*. 2009;84(5):663–8. <https://doi.org/10.1097/ACM.0b013e31819f9092>. PMID: 19704205.
8. Green M, Jones P, Thomas JX Jr. Selection criteria for residency: results of a national program directors survey. *Acad Med*. 2009;84(3):362–7. <https://doi.org/10.1097/ACM.0b013e3181970c6b>. PMID: 19240447.
9. United States Medical Licensing Examination. Work to relaunch USMLE Step 2 CS discontinued. Available from: <https://www.usmle.org/announcements>. Accessed 3 February 2021.
10. Downing SM. Validity: on meaningful interpretation of assessment data. *Med Educ*. 2003;37(9):830–7. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>. PMID: 14506816.
11. Howley LD, Wilson WG. Direct observation of students during clerkship rotations: a multiyear descriptive study. *Acad Med*. 2004;79(3):276–80. <https://doi.org/10.1097/00001888-200403000-00017>. PMID: 14985204.
12. Chang A, Boscardin C, Chou CL, Loeser H, Hauer KE. Predicting failing performance on a standardized patient clinical performance examination: the importance of communication and professionalism skills deficits. *Acad Med*. 2009;84(10 Suppl):S101–4. <https://doi.org/10.1097/ACM.0b013e3181b36f8b>. PMID: 19907367.
13. Klamer DL, Borgia PT. Can students' scores on preclerkship clinical performance examinations predict that they will fail a senior clinical performance examination? *Acad Med*. 2011;86(4):516–20.
14. Cleland JA, Milne A, Sinclair H, Lee AJ. Cohort study on predicting grades: is performance on early MBChB assessments predictive of later undergraduate grades? *Med Educ*. 2008;42(7):676–83. <https://doi.org/10.1111/j.1365-2923.2008.03037.x>. PMID: 18588648.
15. Martin IG, Jolly B. Predictive validity and estimated cut score of an objective structured clinical examination (OSCE) used as an assessment of clinical skills at the end of the first clinical year. *Med Educ*. 2002;36(5):418–25. <https://doi.org/10.1046/j.1365-2923.2002.01207.x>. PMID: 12028391.
16. Chima M, Dallaghan GB. Does student performance on pre-clinical OSCEs relate to clerkship grades? *Med Educ Online*. 2016;22(21):31724. <https://doi.org/10.3402/meo.v21.31724>. PMID: 27340087; PMCID: PMC4919367.
17. LaRochelle JS, Dong T, Durning SJ. Pre-clerkship clinical skills and clinical reasoning course performance: explaining the variance in clerkship performance: pre-clerkship predictors of clerkship variance. *Perspect Med Educ*. 2016;5(4):215–21. <https://doi.org/10.1007/s40037-016-0287-z>. PMID: 27432368; PMCID: PMC4978640.
18. Dong T, Swygert KA, Durning SJ, et al. Validity evidence for medical school OSCEs: associations with USMLE® step assessments. *Teach Learn Med*. 2014;26(4):379–86. <https://doi.org/10.1080/10401334.2014.960294>.
19. Dong T, LaRochelle JS, Durning SJ, Sagui A, Swygert K, Artino AR Jr. Longitudinal effects of medical students' communication skills on future performance. *Mil Med*. 2015;180(4 Suppl):24–30. <https://doi.org/10.7205/MILMED-D-14-00565>. PMID: 25850123.
20. LaRochelle JS, Dong T, Durning SJ. Preclerkship assessment of clinical skills and clinical reasoning: the longitudinal impact on student performance. *Mil Med*. 2015;180(4 Suppl):43–6.
21. Ericsson KA. Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Acad Med*. 2004;79(suppl 1):S70–81.
22. Crossley J, Davies H, Humphris G, Jolly B. Generalisability: a key to unlock professional assessment. *Med Educ*. 2002;36(10):972–8. <https://doi.org/10.1046/j.1365-2923.2002.01320.x>. Erratum in: *Med Educ*. 2003 Jun;37(6):574. PMID: 12390466.
23. Brown C, Ross S, Cleland J, Walsh K. Money makes the (medical assessment) world go round: the cost of components of a summative final year Objective Structured Clinical Examination (OSCE). *Med Teach*. 2015;37(7):653–9. <https://doi.org/10.3109/0142159X.2015.1033389>. Epub 2015 Apr 29 PMID: 25923233.
24. Artino AR Jr, Hemmer PA, Durning SJ. Using self-regulated learning theory to understand the beliefs, emotions, and behaviors of struggling medical students. *Acad Med*. 2011;86(10 Suppl):S35–8. <https://doi.org/10.1097/ACM.0b013e31822a603d>. PMID: 21955765.
25. Guntern S, Korpershoek H, van der Werf G. Benefits of personality characteristics and self-efficacy in the perceived academic achievement of medical students. *Educ Psychol*. 2017;37(6):733–44. <https://doi.org/10.1080/01443410.2016.1223277>.
26. Stegers-Jager KM, Cohen-Schotanus J, Themmen AP. Motivation, learning strategies, participation and medical school performance. *Med Educ*. 2012;46(7):678–88. <https://doi.org/10.1111/j.1365-2923.2012.04284.x>. PMID: 22691147.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.