A Closer Look at Physician Assistants' Performance on the NCCPA Recertification Examinations

Brian Hess, PhD; Raja Subhiyah, PhD

Purpose: The purpose of this study was to investigate how physician assistants (PAs) in major focused areas of practice perform on each of the National Commission on Certification of Physician Assistants (NCCPA) recertification examinations. In addition, the relationship between certification—the Physician Assistant National Certifying Examination (PANCE) and recertification—the Physician Assistant National Recertifying Examination (PANRE) and Pathway II—scores was assessed. *Design*: The sample consisted of the scores of 3,402 PAs who took PANRE between 2000–2001 and 1,206 who took Pathway II between 1999–2001. Initial PANCE scores were also obtained for each member of the sample. Examinees' focused areas of practice included general/family practice (53%), surgical (25%), emergency medicine (11%), internal medicine (8%), and pediatrics (3%). *Results*: Results indicated consistency in both PANRE and Pathway II performance across different practice areas, even after controlling for length of time examinees had spent in their declared practice area. Results also showed no statistically significant correlation between the average amount of time spent seeing patients with a particular organ system condition and the number of questions answered correctly in that organ system of the recertification exam. Finally, PANCE scores significantly correlated with PANRE scores ($r^2 = .31$) and to a lesser extent with Pathway II scores ($r^2 = .07$) for both general/family practice and surgical PA populations. *Conclusions*: PAs across different major practice-focus areas of medicine perform similarly on the recertification examinations, regardless of how long they have spent in their practice area. PAs who on average spend more time seeing patients with a specific organ system condition do not do better in that area of the recertification exam. Additionally, PAs who had higher scores on PANCE also tended to have higher scores on the recertification exams, notably PANRE. The findings of this study are consistent with the training and practice model adopted by NCCPA to describe the philosophy behind the PA recertification examinations.

Introduction

The National Commission on Certification of Physician Assistants (NCCPA) recertification programs are designed to assess physician assistants' (PAs) general medical knowledge and skill

Brian Hess was the manager of Test Development & Research, National Commission on Certification of Physician Assistants, Norcross, Georgia. Raja Subhiyah is the associate vice president, Scoring Services, National Board of Medical Examiners, Philadelphia, Pennsylvania.

Correspondence should be addressed to:

Brian Hess, PhD
American Board of Internal Medicine
510 Walnut Street, Suite 1700
Philadelphia, PA 19106
E-Mail: bhess@abim.org

in primary care. However, data from the 2000 NCCPA stakeholders' study[1] and considerable anecdotal evidence indicate that there is a perceived need for some changes to the NCCPA recertification examination and/or process. The impetus for these changes was based on concerns of the growing number of PAs who are moving away from general practice into focused areas of practice. Their contention is that the NCCPA recertification examination and its test specifications are no longer appropriate for them because of its generalized content across areas of medicine. In response, the NCCPA Board of Directors began to explore additional or modified processes for recertification by taking into account focused areas of PA practice.

As the NCCPA Board of Directors examined the recertification process, it also initiated an empirical research program to support the validity of current examination scores. For instance, the impact of the practice-focus transition on NCCPA recertification exam performance was recently investigated by Hess and Subhiyah (in this issue).[2] Using factor analysis methods, they discovered that the current structural design and scoring model for the Physician Assistant National Recertifying Examination (PANRE) was consistent across general/family practice and surgical PAs. Only a single, general ability factor primarily explained variation in responses to questions for each content blueprint dimension. It was concluded

that recertifying surgical PA examinees do not appear to be at a disadvantage when taking PANRE, despite the fact that the exam measures PAs' general medical knowledge in primary care.

Whether or not the practice-focus transition across the PA profession has placed other major practice groups at a disadvantage when taking the NCCPA recertification exam has not been investigated. Thus, the purpose of the present study was to take a closer look at how PAs in major focused areas of practice perform on the current NCCPA recertification examinations. The relationship between examinees' PANCE scores and their recertification test (PANRE or Pathway II) scores was also investigated.

Method

Sample

The data used in this study included responses of 3,402 candidates who took PANRE in 2000 and 2001 and 1,206 candidates who took Pathway II in 1999, 2000, and 2001. All candidates' records represented first attempts to pass each test in their last recertification cycle. In addition, PANCE scores (from PAs' first attempt to gain initial certification) were also used for each subject. In this study, the final standard scale scores were reported; the standard scale used converted raw scores so that the mean was 500 and the standard deviation was 100 (standard scale scores for the NCCPA examinations ranged from 200 to 800). These scores are equivalent across administrations and forms.[3] Pass rates or percentages were also reported. The passing score for all exams was 350 and was based on previous standard-setting procedures conducted by NCCPA and National Board of Medical Education (NBME) staff using a modified Angoff method and several separate expert panelists.

Determining PA Practice Focus

PAs in several different major focused areas of practice were represented in the sample. Each examinee's practice focus was determined according to the specialty reported category in longitudinal data obtained from the American Academy of Physician Assistants (AAPA)'s PA census survey from 1991 to 2001. For this study, the length of time examinees spent in a declared practice-focus area was defined as the number of past consecutive years reported in their practice focus at the time of the recertification exam. PAs who had spent less than two consecutive years in their practice focus at the time they took the recertification examination were eliminated from the analysis. Additionally, only a few major practice-focus groups were able to be determined from the data set and thus are represented in this study. The sample sizes of various other medical and surgical specialties/subspecialties were too small to extrapolate adequate statistical comparisons and inferences. Thus, the major practice-focus areas represented were general/family practice (53% of the sample), surgery (25%), emergency medicine (11%), internal medicine (8%), and pediatrics (3%). Finally, information from the NCCPA practice survey was used to determine whether PAs who, on average, spend most of their time seeing patients with certain organ system conditions obtain higher scores in that organ system area of the recertification exam.

Data Analysis

An analysis of covariance (ANCOVA) procedure was used to assess group differences in mean scores. Subsequently, a χ^2 test was used to compare pass rates across groups. The overall level of statistical significance was set at $P<.05$. However, because large sample sizes were used in this study, we reported and interpreted measures of effect size. Similarly, when we reported correlations, we also reported and interpreted the squared correlation coefficient (r^2) effect size, which represents the percentage of variance shared between the two variables.

Results

PANRE Performance

An ANCOVA was conducted with mean PANRE score as the dependent variable, type of practice focus as the independent variable, and length of time in a practice focus as a covariate. Results showed that the length of time in practice focus, while statistically significant, did not account for a meaningful amount of variance in PANRE scores. The effect size η^2 value of .002 indicates that only 0.2 % of the variance in PANRE scores was explained by length of time in practice. However, we retained this variable as a covariate in order to compare practice-focus mean scores, controlling for length of time spent in a practice focus.

Results indicated a statistically significant difference in mean scores between the major practice-focus groups, $F_{(4, 3397)}=72.81$, $P<.05$, $\eta^2=.07$. However, the effect size indicated that only 7% of the variance in PANRE scores is attributed to differences in type of practice focus, after controlling for length of time spent in that practice focus. While the data indicated that surgical PAs on average scored lower than the other practice-focus groups, the size of the difference was small. When comparing pass rates across the practice-focus areas, results of a χ^2 test indicated a statistically significant difference in pass rates, $\chi^2_{(4, N=3,402)}=36.23$, $P<.05$, $r_c=.09$. However, the contingency coefficient effect size (r_c) value of .09 indicated that only 9% of the variability in pass rates is attributed to differences in the type of practice focus. Table 1 presents the overall PANRE performance of PAs in major focused areas of clinical practice.

Table 1

PANRE Performance of PAs in Major Focused Areas of Practice

Practice area	N	M (SD)	Pass %
General/Family	1947	533 (81)	98.7
Surgical	760	481 (81)	95.4
Emerg Med	331	529 (79)	99.1
Internal Med	292	551 (81)	99.3
Pediatrics	72	528 (74)	98.6
Total	3402	522 (80)	98.1

Note: Group means are adjusted after controlling for length of time examinees spent in their practice focus.

In summary, there were very minor differences in PANRE scores across the different practice-focus areas, controlling for length of time spent in a practice focus. Pass rates for all groups were at or above 95%.

Pathway II Performance

Results of an ANCOVA similar to the one conducted for PANRE indicated no statistically significant differences in mean scores among the different practice-focus groups after controlling for length of time spent in a practice focus ($P>.05$). When comparing pass rates, results of a χ^2 test indicated no statistically significant difference in pass rates ($P>.05$) among PAs in different practice-focus groups. In summary, Pathway II performance was consistent across the different practice-focus areas after controlling for length of time spent in a practice-focus area. Table 2 presents the overall Pathway II performance of PAs in focused areas of clinical practice.

Relationship Between PANCE and Recertification Exam Scores

Table 3 shows the correlation of PANCE scores with PANRE and Pathway II scores. As expected, PANCE scores (first attempt) predicted performance on PANRE better ($r=.56$) than for Pathway II ($r=.27$). Both PANCE and PANRE are proctored tests, while Pathway II is a take-at-home test that, by

Table 3				
Correlation Between PANCE and NCCPA Recertification Examination Scores				
	PANCE w/PANRE		PANCE w/Pathway II	
Group	N	r	N	r
General	1947	.56	438	.31
Surgical	760	.56	374	.21
Total	2707	.56	812	.27
Note: Data is from examinees' first attempt to pass each exam.				

design, measures different skills and concepts. Squared correlations (r^2), representing the amount of variability in PANRE scores explained by PANCE scores, was 31%. In anticipated contrast, the Pathway II score variance explained by PANCE scores was only 7%. Correlations of PANCE scores with PANRE scores were the same for general/family practice PA and surgical PA examinees.

Interestingly, surgical PAs who took PANRE also tended to score slightly lower on their first attempt at PANCE (for initial certification), as shown in Table 4. However, as with PANRE, this difference was not meaningful.

Organ System Performance

Correlations between the average amount of time spent seeing patients with a particular organ system condition and number of questions answered correctly in that organ system of the exam were calculated. Results showed no statistically significant correlations for any of the 13 organ systems (correlations were all less than .04 for all the organ systems). These results were the same for PANRE and Pathway II examinees.

Discussion

The purpose of the present study was to take a closer look at how PAs in different focused areas of practice performed on each of the current NCCPA recertification examinations. The relationship between PANCE and recertification scores was also investigated.

Very small differences were found in PANRE scores across major practice-focus areas; however, these differences were not meaningful, as they accounted

for only a small portion of the variance in PANRE scores. No significant differences were found across these groups on Pathway II. No differences were found across length of time spent in their declared practice area in either PANRE or Pathway II. Finally, pass rates across all groups represented in this study were well above 90%. This suggests that PAs across major practice-focus areas of medicine generally perform well on the NCCPA recertification examinations—regardless of how long PAs spend in their practice-focus area.

While concern has been raised by some PAs in specialty areas about their continued ability to pass a generalist examination, it does not appear that the major practice-focus groups represented are at risk for lower performance on the NCCPA recertification examinations. Further, these results are consistent with Hess and Subhiyah[2] in that the performance of surgical PAs on PANRE does not indicate that they are at a disadvantage when taking the PANRE. Moreover, our data are consistent with the training and practice model adopted by NCCPA to describe the philosophy behind the PA recertification exams.

Table 2			
Pathway II Performance of PAs in Major Focused Areas of Practice			
Practice area	N	M (SD)	Pass %
General/Family	475	531 (85)	97.7
Surgical	406	524 (90)	94.4
Emerg Med	168	532 (92)	95.2
Internal Med	84	540 (99)	97.6
Pediatrics	73	518 (93)	94.5
Total	1206	529 (89)	96.0
Note: Pathway II data is aggregated across the 1999 through 2001 administrations. Group means are adjusted after controlling for length of time examinees spent in their practice focus.			

Table 4			
PANCE Performance of General/Family Practice and Surgical PA Examinees			
	PANCE		
Specialty	N	M (SD)	Pass %
General	1947	527 (75)	92.9
Surgical	760	515 (76)	90.1
Total	2707	523 (76)	92.1
Note: Data are from PANRE examinees only.			

Similarly, when PAs did well overall on the recertification exams, they performed well across all organ system areas of the exam; when PAs did not perform well overall, they did not perform well across all organ systems. In fact, PAs who reported (at the time of recertification) spending on average more time seeing patients with a particular organ system condition did not tend to correctly answer more questions in that organ system area on the exam. This may suggest that the NCCPA recertification (or indeed, the certification) exams are not specialized enough to pick up on the differences between PAs who work most often in one or another area. In other words, the NCCPA examinations measure core knowledge that is shared by these areas. This is consistent with the NCCPA's philosophy that these exams are constructed for the general practitioner. They are not targeted to the eclectic specialist. Thus, this finding is consistent with the PA "generalist" training and practice model and lends support to the validity of the NCCPA recertification examination scores.

Finally, the results of this study indicated that PAs who scored higher on PANCE also tended to score higher on the recertification exams, particularly PANRE. These correlations were consistent across general/family practice and surgical PAs. This suggests that PANCE scores have respectable predictive validity vis-à-vis PANRE scores. The amount of shared variance between PANCE and PANRE scores (31%) may be attributed to the similarity of the two tests in format, method, and content. This is expected as most certification exam scores across the professions typically correlate fairly highly with recertification scores when the same methods are used. Conversely, because the Pathway II is a take-at-home test, it measures (by design) a largely different set of skills in a largely different operating environ-

ment. Hence, weaker correlations are expected and duly obtained.

The predictive potential of PANCE of recertification exam scores may partially explain why surgical PAs tended to score on average a bit lower on *both* the NCCPA certification and recertification exams. While these differences were very small, this pattern suggests that something other than the structure of the NCCPA examinations themselves may be contributing to their slightly lower mean scores.[2]

It may be of value to investigate how PAs in different practice-focus areas prepare for the NCCPA recertification exams. For example, the leadership of the AAPA Surgical Congress has previously expressed the need for a substantial review of generalist content by their constituents in order to prepare for the NCCPA examinations. They suggested that continuing medical education (CME) should be directed at this goal (in addition to activities that target their focused area of practice, which typically result in the accumulation of an excess of the hours needed for certification maintenance). This type of preparation may explain the consistency in performance between groups seen in the present study. Furthermore, future studies could also examine the type and amount of CME undertaken to delineate the impact on recertification performance. Other preparation characteristics, such as self-study, might be examined as well.

Limitations

Not every medical specialty/subspecialty was uniquely represented. For example, surgical PA examinees in this study were identified via AAPA survey data as any PA practicing in any one or more surgical specialty or subspecialty. PAs were not identified as belonging to a very specific surgical area such as orthopedics, cardiovascular surgery, etc. This may pose a limitation in that uniqueness

of any one surgical specialty or subspecialty could not be investigated. However, because these specialized areas of surgery share a great deal in common, it made sense for the purpose of this study to identify and include one overall surgical category. Finally, we defined length of time spent in a practice-focus area as the consecutive number of years PAs worked in their declared practice area at the time of recertification. This may be confounded by the mobility within the PA profession (ie, moving back to one practice focus after several changes may not accurately reflect the experience accumulated over one's professional career).

Conclusion

PAs across major practice-focus areas of medicine perform consistently well on the NCCPA recertification examinations, regardless of the length of time they spent in their practice focus. PAs who on average spend more time seeing patients with a specific organ system condition do not perform better in that area of the NCCPA recertification tests. Additionally, PAs who perform higher on PANCE also tend to score higher on the PANRE. The findings obtained in this study are consistent with the training and practice model adopted by NCCPA to describe the philosophy behind the PA recertification examinations.

References

1. National Commission on Certification of Physician Assistants. *Assessment of the Current and Future Uses and Value of NCCPA Certification*. NCCPA: Norcross, Ga; 2000.
2. Hess B, Subhiyah R. Confirmatory factor analysis of the NCCPA Physician Assistant National Recertifying Examination. *Perspective on Physician Assistant Education*. 2004;15:55-61.
3. Subhiyah R, Hess B, Arbet S. Building a certification examination program, Part II: Scoring and standard setting. *Perspective on Physician Assistant Education*. 2004;15:21-24.