QUALITY OF NURSING DOCTORAL EDUCATION AND SCHOLARLY PERFORMANCE IN U.S. SCHOOLS OF NURSING: STRATEGIC AREAS FOR IMPROVEMENT

MI JA KIM, PhD, RN, FRCN, FAAN,* CHANG GI PARK, PhD,† SO HYUN PARK, MSN, RN,‡ SHAHEEN KHAN, MS, MBA, MPH,§ AND SHAKE KETEFIAN, EdD, RN, FAAN‖

Comprehensive evaluation of quality of nursing doctoral education (QNDE) in research-intensive universities has not been reported since 1980s. This study aimed to examine the QNDE from the perspectives of faculty and students/graduates and their relations to school characteristics, identify factors of the four domains of the QNDE that influence the QNDE, and analyze the relationship of QNDE to scholarly performance of nursing schools in the United States.

Seventy-two nursing schools offering research-focused nursing doctoral programs with National Institutes of Health (NIH) funding during 2004–2007 participated, and they responded to the questionnaire (see http://gknf.or.kr/research/). Twenty-nine deans/schools, 179 faculties, and 461 students/graduates responded. Both faculty and students/graduates groups rated quality positively. Schools in the top quartile group per NIH funding amounts showed significant differences in QNDE from the bottom quartile group. Program and faculty domains were identified as most important by the top quartile group, and items that were significantly associated with the quality were supportive environment for students’ learning, faculty mentorship, and assistance to students in understanding the value of programs of research and scholarship. Percentage of faculty member with research grants was significant predictors for all domains of QNDE, and time to degree was significant in explaining overall quality. (Index words: Nursing doctoral education; Quality evaluation; Strategic areas) J Prof Nurs 30:10–18, 2014. © 2014 Elsevier Inc. All rights reserved.

THE RAPID EXPANSION of doctor of philosophy (PhD) programs in nursing has been noted in the literature; this has ostensibly been intended to allay shortages in nurse faculty and in other scholarly ranks (Berlin & Sechrist, 2002). For example, research-intensive doctoral programs have expanded from 86 in 2001 to 133 in 2010 (American Association of Colleges of Nursing [AACN], 2011). This same trend has been noted worldwide (International Network for Doctoral Education in Nursing [INDEN], 2011; Ketefian, Davidson, Daly, Chang, & Srisuphan, 2005). At the same time, the profession has been concerned about doctoral program quality and the extent to which the resources, environments, experiences, and capacities of these programs and institutions are sufficient to prepare graduates to function at the high level that society expects from them (Minnick,
Norman, Donaghey, Fisher, & McKirgan, 2010). Thus, in order to assure the high quality of education and engage in ongoing assessments, appropriate measurement tools and evaluation models for PhD programs in nursing are necessary.

**Background**

Various evaluation frameworks have been proposed in the literature. Chioni (1985) and Jamann (1985) have proposed six criteria: faculty qualifications, program of study, resources, students, research, and evaluation. Since 1986, AACN has developed quality indicators around five criteria: students, faculty, program, resources, and evaluation (AACN, 2001). In an earlier edition of the AACN document, research was included as a separate category, which was later incorporated within the first four criteria of the current edition. For a number of years, doctoral programs in nursing and multiple disciplines made use of the graduate program self-assessment (GPSA), developed and sold by the Educational Testing Service; Barhyte and Holzemer (1980), Holzemer and Chambers (1986), and Holzemer (1987) used the GPSA for several rounds of doctoral program evaluation at the national level, with support from the Division of Nursing, U.S. Public Health Service.

Based on the literature about evaluation of doctoral programs in nursing, Kjellgren, Welin, and Danielson (2005) reported that more systematic evaluations on nursing education were necessary to guide the development of quality. Holzemer (1982) has suggested a systems view of quality assessment. This model has three components: context, environment, and product. Context (or input) deals with concrete elements that can be counted, such as faculty numbers and their qualifications, publications, research dollars garnered, number of books in the library, student grades, and scores on tests, and others. Product deals with the outcomes of various faculty responsibilities in education, research, and service. It focuses on the productivity of alumni, impact of research produced on society, and the impact of faculty, student, and alumni service to society and the professional community. Quality of environments relates to quality of teaching, faculty concern for students and the interactions of these groups, and the overall facilitation within the institution that is conducive to scholarly and professional productivity (Holzemer & Chambers, 1986). No valid and reliable tool has been designed specifically for the evaluation of research-focused nursing doctoral programs in a comprehensive manner within the United States and worldwide. Under the auspices of the INDEN (2004), an international task force, drawing on the AACN document of 2001 developed global quality standards, criteria, and indicators (QSCI) for doctoral programs to provide guidance to doctoral programs worldwide (Kim, McKenna, & Ketefian, 2006). The QSCI served as the basis for developing the current instrument, quality of nursing doctoral education (QNDE), which was used in this study.

Similarly, at the global level, there have been no reports involving evaluation of doctoral programs. Thus, for PhD programs nationally, using the QNDE (QNDE in this italic form indicates the questionnaire instrument we used for evaluation of QNDE) for the first time in the United States.

This study assessed the QNDE comprehensively by examining quality dimensions from the perspectives of students/graduates (recipients) and faculty (providers); at the same time, we examined the relationship of organizational features to quality. It is part of a seven-country study directed by the same principal investigator who directed this study in collaboration with national teams. The results of the studies in Japan and Korea have been published (Arimoto, Gregg, Nagata, Miki, & Murashima, 2012; Kim et al., 2012; Nagata et al., 2012). Findings from South Africa, Thailand, and the United Kingdom, as well as a comparative report across the participating countries, will be forthcoming in the international nursing literature. This article aims to examine the QNDE in the United States and the quality of the QNDE itself and determine the relationship of features of the QNDE and scholarly performance of doctoral programs.

**Aims of the Study**

Specific aims (SA) of this study were to (a) describe and compare the QNDE with the perspectives of faculty and students/graduates, (b) identify the domains of the QNDE that account for the quality of doctoral education, and (c) examine the relationship of scholarly performance of nursing schools to the four domains of the QNDE.

**Methods**

A descriptive, cross-sectional, on-line survey design was used to evaluate selected domains of nursing doctoral education quality from the perspective of nursing faculty and nursing doctoral students and graduates and examine the relationship between the QNDE and school-level scholarly performance.

Seventy-two schools of nursing in the United States were selected if they (a) offered research-focused nursing doctoral programs and (b) received National Institutes of Health (NIH) funding during the years 2004–2007, including those with 1-year funding. Four-year NIH funding amounts were used to capture the broader-spectrum funding history of schools; 2007 data were the latest available at the beginning of this study. Rankings of nursing schools for the year of 2004 were obtained from the linkdsa.pub2004c (and for the years 2005 through 2007 were from http://report.nih.gov/award/award.cfm). When the rankings were not provided by the NIH, we ranked them according to the sum of total funding of each school.

**Instruments**

**QNDE Questionnaires**. The QNDE questionnaire survey tool was based on international QSCI developed by an
international committee of the INDEN (Kim et al., 2006). The QSCI was modified from the AACN indicators of quality in research-focused doctoral education in nursing (AACN, 2001). Four domains of the QNDE (program, faculty, resources, and evaluation) were evaluated in this study from the perspectives of faculty and students/graduates. The items in the QNDE were designed for online questionnaire format (http://gknf.or.kr/research/); a new section was added for deans/department heads to obtain information on school characteristics and demographic data. Validity and reliability of the QNDE needed to be examined because this was the first time the QSCI-based instrument was used to evaluate the QNDE in the United States; the procedure we used to do so is described below.

For Faculty, Students, and Graduates. The 43-item questionnaire for faculty and graduates/students consisted of four domains: the quality of program/curriculum (17 items), faculty (12 items), resources (9 items), and evaluation (5 items). Some sample items are presented here for each domain: all students receive formal training in ethics and the protection of human participants and/or animal subjects in research (program); faculty members mentor and assist students to understand the value of programs of research and scholarship (faculty); library has sufficient holdings, search engines, and databases (resources); program evaluation is systematic, ongoing, comprehensive, and it focuses on the university's and the program's specific mission (evaluation). Each item was rated on a 4-point scale (1 = strongly disagree to 4 = strongly agree). Appropriateness, feasibility, and user-friendliness of the questionnaire items were evaluated by faculty members with experience in doctoral education in the United States, and the on-line survey was then pilot tested with several faculty members (n = 4) and PhD students/graduates (n = 5) before it was used for the main study. Their feedback on the wording and format of the questionnaire for deans was incorporated in the questionnaires. Their suggestions about ease of administration and timeliness of accessing the survey Web site were addressed by the Web site coordinator.

Questionnaire for Deans. The questionnaire for deans included items regarding school characteristic, such as consistency of purpose of nursing doctoral education with that of parent institution, number of enrolled doctoral students per year, number of graduates per year, and number of employed students after graduation per year.

Validity and Reliability of the QNDE Questionnaire. Content validity was established by expert opinions of four professors in the United States who had experience in nursing doctoral education (Rossiter, 2002). Because the items in the questionnaire in this study were formative constructs, construct reliability and validity of the QNDE questionnaire were examined using an established method for doing so (Diamantopoulos & Winklhofer, 2001). The construct reliability can be established with the absence of multicollinearity by using multiple linear regressions with the summed scores of each domain as dependent variables (i.e., program, faculty, resources, and evaluation) and items included in the quality indicators as independent variables (Diamantopoulos & Siguaw, 2006). Analysis of our data showed that the range of variation inflation factor (VIF) values of the program domain was 1.67–2.99; faculty domain, 1.94–3.27; resource domain, 1.55–2.71; and evaluation domain, 2.37–3.27. All items as predictors of multiple regressions of overall perceived quality showed lower VIF values. Therefore, the indicator reliability requirement was satisfied (Diamantopoulos & Siguaw, 2006). Analysis of data from a study in Japanese (Arimoto et al., 2012; Nagata et al., 2012) that is one of our international collaborative research projects, using the same QNDE questionnaire, same data analysis methods, and same principal investigator as this study, showed similar validity and reliability. Formative construct validity was confirmed with all statistically significant indicator weights for the four domains. The indicator weights were estimated with partial least squares path modeling for formative construct analysis (Chin, 1998). External validity of the four domains was confirmed with the significant associations among four domain scores and overall rate of quality score (P < .01; Diamantopoulos & Winklhofer, 2001).

Scholarly Performance. Scholarly performance was added to examine its relationship with the QNDE. It was measured by annual number of publications per faculty member, number of citations, percent of faculty with grants, and time to degree of students in each school. These data were obtained from the Chronicle of Higher Education (CHE, 2011) because our data collection instruments did not include these except for time to degree. Unfortunately, the CHE data were available for only 18 nursing schools out of the 72 participating schools, and data included were only those for the 2005–2006 academic year, which is within the time frame in this study (2004–2007), although not identical to it.

Data Collection Procedures

Electronic invitation letters were sent to the deans of 72 schools after current contact information was obtained from the Web sites of the colleges. This was preceded by the principle investigator's telephone calls to most of the deans to explain the purpose of the study and request their permission and assistance in disseminating information to their faculty, students, and graduates. Deans were most helpful in referring the investigators to appropriate individuals at their schools who would distribute the questionnaire and serve as the colleges' liaisons. A set of e-mails with attachments that were to be sent to potential participants were then sent to the college liaisons. The set included a cover letter explaining the purpose of the study, institutional review board (IRB) approval letter, instructions about how to access the online QNDE questionnaire using the Web address (http://
gknf.or.kr/research/) along with two follow-up letters that needed to be sent 2 weeks and 4 weeks after the initial e-mail. Deans had a separate set of questionnaires within the QNDE questionnaire Web site that primarily asked for school characteristics such as their demographic and background information. Each school was given the option of either sending contact information of their faculty, students, and graduates to a research team member at our institution for centralized recruitment or using the instructions that the project staff sent to them to execute on-site recruitment and follow-up. Most of the schools opted to use the latter mechanism.

A research team member who was in charge of overseeing the administration of the on-line survey was available as a resource for college liaisons and potential participants throughout the data collection period. Data collection began in late March 2009 and closed in December 2010. The lengthy period was primarily because of the time needed to identify graduates and clarify questions. Some schools obtained their own IRB approval even though we offered them our institution’s approval.

Data Analysis

Descriptive statistics, two-group t test, and analysis of variance (ANOVA) were used to describe the characteristics of the survey participants and schools, compare the responses of providers and recipients, identify influencing factors to the QNDE, and compare it among schools in four quartiles categorized according to the average total funding amounts from the NIH during 2004–2007. Chi-square analysis was performed to examine relationships between number of correspondence and respondent rate and between QNDE domains and scholarly performance. Importance-performance analysis (IPA) was conducted to identify which domains of the QNDE needed attention for improvement. Seemingly unrelated multivariate regression (SUR) was applied to examine the relationship between multivariate QNDE and scholarly performance of nursing schools. The data envelopment analysis (DEA) and data mining methods were used for the additional analysis as presented in the Discussion section.

Results

Demographic Characteristics of Schools

Twenty-nine deans, 179 faculty, and 461 students/graduates (i.e., students, n = 297 and graduates, n = 164) responded to the questionnaire. Overall response rate of schools was 40% (29 out of 72). Twenty-two out of 29 responding deans of schools of nursing (76%) indicated that their programs were very closely aligned with the vision, goals, and mission of their parent institutions. The responses of the 29 schools were examined in four quartiles, per NIH funding amounts during the years 2004–2007. Six were in the top quartile, 8 in the second, 8 in the third, and 7 in the fourth. The average number of e-mail correspondence with all schools over the study period was 5.31, with the range of 1–16. There was no significant relationship between number of correspondence and response rate (P = .083). The mean number of enrolled doctoral students per school per year was 48.81, whereas the average number of graduates per year was 7.47. Time-to-degree responses were too few to be included in data analysis. Most of the graduates were employed after graduation (86.7%). Table 1 shows selected characteristics of schools in this study.

SA #1. Compare the Perspectives of Faculty and Students/Graduates on the QNDE.

Evaluations by faculty, students, and graduates showed positive ratings on QNDE questionnaire. Resources and evaluation domains were rated lower than the other two domains. The average rating on all four domains by faculty was 3.38 out of 4.0, which was the most positive (SD = 0.43, range = 1.93–4); by students was 3.37 (SD = 0.45, range = 1.85–4) and graduates was 3.37 (SD = 0.44, range = 1.06–4).

Because of the comparability among the three groups, an ANOVA was performed, revealing no significant differences among them (Table 2). Multivariate ANOVA was also conducted to confirm the comparability of the four domains among the three groups.

SA #2. Identify the Domains of the QNDE That Account for Overall Quality.

Quartile-group comparisons using a one-way ANOVA test followed by multiple comparisons showed that the quality measure differences among the four quartile groups (Gs) were significant for three domains of quality measures—all except the evaluation domain. The Bonferroni multiple-comparison test was applied, which revealed significant differences between G1 (top quartile) and G4 (bottom quartile). All other pair-wise comparisons were not significant (P > .05). This is shown in Table 3.

To identify domains that are important to the QNDE, we conducted IPA (Martilla & James, 1977). The performance was measured based on the survey response results, but the importance was estimated using indirect methods such as standardized regression coefficients (Dolinsky, 1991; Wittink & Bayer, 1994). While the importance measures signify the central role each domain plays in the QNDE, performance measures indicate achievements/outcomes of each school in four domains of the QNDE.

Quadrant B of the importance-performance matrix graph showed the highest results in both importance and performance, and both program and faculty domains of G1 (P1 and F1) found in this quadrant indicate that G1 needs to “keep up the good work.” In contrast, Quadrant C showed the lowest results in both importance and performance dimensions; as well, the resource domain of G1 (R1) and G4 (R4) and the evaluation domain of G4 were found in this quadrant. Quadrant A shows domains high in importance but low in performance; program and faculty domains of G4 found here indicate that they need to focus on these domains to improve the quality of doctoral education. Quadrant D shows domains low on
Because the program and the faculty domains were identified as most important by G1 (Quadrant B), we further examined questionnaire items under these two key domains that could give more specific guidance to schools for improvement. To identify such item(s), we applied data mining methods such as regression tree, Random Forest, and generalized boosting method to the data as a nonparametric estimation method. The most important item that was associated with the total program domain score was, Item #11 (environment is supportive of students' learning); whereas, for the faculty domain, it was Item #32 (faculty members mentor and assist students to understand the value of programs of research and scholarship). The domain of resources requires attention by both Groups 1 and 4 because the quality and availability of resources can affect doctoral program quality.

SA #3. Examine the Four Domains of the QNDE and Their Relationships With Scholarly Performance.

Recognizing the central role of faculty's scholarly performance in the QNDE led us to obtain national data on scholarly performance for the relevant study period from the CHE (2011), but as mentioned earlier, the CHE included data for only 18 schools out of the 29 schools we studied and only for one academic year out of the four in this study. Multivariate regression analysis was conducted with key variables of scholarly performance, such as annual publications per faculty member, number of citations, percentage of faculty with grants, time to degree, number of PhD graduates per year, and school type (CHE, 2011). All four regression models estimated with seemingly unrelated regression method were statistically significant according to the chi-square test. Furthermore, the Breusch–Pagan test of independence of residual terms was statistically significant to reject the null hypothesis of independent relations among the four qualities (i.e., the four domains of QNDE) equations residuals, $\chi^2(6) = 599.917$, $P < .0001$. In other words, there were significant relations among the four domains of QNDE even after controlling for the independent variables (i.e., scholarly performance and other school characteristics). To estimate more general and robust association of these indicators with the domains of QNDE, we performed significance tests that included each school's scholarly performance over four domains of QNDE. Neither number of annual publication per faculty member nor number of citations was significant, $F(4, 324) = 1.92$, $P = .1065$ and $F(4, 324) = 1.24$, $P = .2918$, respectively. However, the percentage of faculty members with research grant funding was significant for all domains of the QNDE, $F(4, 324) = 2.63$, $P = .0343$. Time to degree was also significant in explaining overall qualities, $F(4, 324) = 4.75$, $P = .0010$. It is unfortunate that our survey data on time to graduation were unusable because of limited and unreliable responses, but scholarly performance data from the CHE nevertheless provide a partial picture of the relationship of scholarly performance to the QNDE. These findings were confirmed with generalized estimation equation analysis that estimated school factors on aggregate quality while controlling for high correlations among the four measures of quality domains. Again, the percentage of faculty members with research grant funding and the time to degree were significant in explaining the overall qualities. In other words, schools that had higher percentage of faculty members with research grant funding and lower time to degree were significantly associated with all four domains of QNDE (Table 4).

### Discussion

This is part of an international study that examined the QNDE in seven countries (Australia, Japan, Korea, South Africa, Thailand, United States, and UK). This article reports the results of the QNDE survey conducted in the United States. As noted, this is the first study in the United States that attempted a comprehensive assessment of the QNDE from the perspectives of faculty and student/graduates and examined the relationship be-

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (SD)</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of enrolled doctoral students per year</td>
<td>48.81 (34.17)</td>
<td>13–134</td>
</tr>
<tr>
<td>No. of graduates per year</td>
<td>7.47 (5.73)</td>
<td>1–28</td>
</tr>
<tr>
<td>No. of employed students after graduation per year</td>
<td>6.48 (5.65)</td>
<td>1–26.67</td>
</tr>
</tbody>
</table>

*Obs = Observation.

### Table 1. Characteristics of Schools During 2004–2007 *(n = 29)*

<table>
<thead>
<tr>
<th>Obs.</th>
<th>Faculty (n = 179)</th>
<th>Students (n = 297)</th>
<th>Graduates (n = 164)</th>
<th>F (P value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program</td>
<td>651</td>
<td>3.41 (0.49)</td>
<td>3.36 (0.48)</td>
<td>3.33 (0.62)</td>
</tr>
<tr>
<td>Faculty</td>
<td>640</td>
<td>3.45 (0.52)</td>
<td>3.44 (0.53)</td>
<td>3.43 (0.59)</td>
</tr>
<tr>
<td>Resource</td>
<td>639</td>
<td>3.17 (0.54)</td>
<td>3.23 (0.56)</td>
<td>3.20 (0.60)</td>
</tr>
<tr>
<td>Evaluation</td>
<td>509</td>
<td>3.39 (0.56)</td>
<td>3.33 (0.54)</td>
<td>3.31 (0.56)</td>
</tr>
</tbody>
</table>

*Obs = Observation.

importance but high on performance; only the evaluation domain of G1 is found here, although it is at the borderline to Quadrant B. This is shown in Figure 1.
tween scholarly performance indicators and quality domains. It also examined the quality of doctoral education from individual and institutional levels, whereas previous studies assessed quality based on either individual level or school level alone (Minnick & Halstead, 2002; Minnick et al., 2010). This study affirms the importance of including students and graduates for comprehensive evaluation of PhD programs (Morrison, Zumeta, Rudd, & Nerad, 2011). In addition, this study expands the understanding of quality assessment of research-intensive nursing doctoral programs in the United States by demonstrating both traditional scholarly performance and stakeholders’ quality assessment.

In this study, 22 deans of 29 school of nursing (76%) indicated that their program was very closely related to the vision, goals, mission, and objectives of their universities. This finding contrasts to the findings of Minnick et al., although their focus was not exclusively on research-focused doctoral program (Minnick & Halstead, 2002).

SA #1. Favorable ratings given by both faculty (provider) and students/graduates (receivers) on all four domains of the QNDE indicate that the overall QNDE in the United States need not be a source of concern. Lack of significant differences between the two groups noted in this study is in contrast to the findings of the Korean and Japanese studies (Kim et al., 2012; Nagata et al., 2012). Kim et al. 2012 reported in the Korean study that faculty’s perceptions of the quality of the domains of program, faculty, and resources were higher than the perceptions of the graduates and students of the same domains. In the case of the Japanese study (Nagata et al., 2012), the evaluations provided by faculty members were more negative than that provided by graduates and students. This might be related to attitudes of students and graduates toward faculty in general and hierarchical relationships between the faculty and the students/graduates. It is plausible that the students and graduates in Japan wanted to show their respect to the faculty by providing positive responses rather than negative ones.

SA #2. High correlations among the four domains of the QNDE found in this study were not a surprise because these domains together constituted the overall QNDE. The evaluation domain did not show a significant difference among the four quartile groups; this may be explained by the small number of responses to the questions related to this domain. Significant quality differences found between G1 and G4 indicate that quality measures were significantly different between the top quartile group (G1) and lowest quartile (G4) group.

The results of IPA may help schools in G4 develop more specific strategies for improvement in the elements of the program and faculty domains of the QNDE. Figure 1 shows the program and faculty domains of the G1 as having the highest importance and performance (by virtue of being in Quadrant B). Interestingly, the resource domains of G1 and G4 were found in Quadrant C, indicating that both had low importance and low performance. No domains were found in Quadrant D, except for the evaluation domain of G1, which was at the borderline to Quadrant B in terms of importance. High importance and performance shown in the program and the faculty domains in G1 according to the IPA suggest that schools in G4 particularly need to give focused attention to these domains to improve the QNDE.

SA #3. In examining the relationship between the QNDE and scholarly performance, the percentage of faculty members with research grant funding emerged as one of the most significant factors among the indicators of scholarly performance and epitomized its critical role in quality. It played a major role in distinguishing top-quartile schools in research-intensive universities and substantiated the importance of faculty’s competitiveness in generating innovative ideas, their successful grantsmanship, and the schools’ support system, including infrastructure that accounts for success in grantsmanship. Although the percentage of faculty members with research grant funding (data from the CHE 2010) may not be the same as the total funded amounts of research grants in schools of nursing (data we used in this study to represent scholarly performance), we considered them as representing similar profiles of scholarship performance (CHE, 2011).

It is interesting to note that, although Groups 1, 2, and 3 had different NIH funding levels, their QNDE scores were not significantly different on the four domains; hence, efficiency of scholarly performance for each of the 29 participating schools was examined to gain insight.
into its relationship with the QNDE. We reasoned that the most efficient schools would indicate that they have leadership, infrastructure, and mechanisms in place that facilitate faculty's scholarly performance. School-level efficiency estimated with DEA yielded a relative efficiency score for each school and identified the benchmarking schools (Charnes & Cooper, 1978). The top four schools were the most efficient. It was interesting to note that all top four schools with the highest efficiency scores belonged to G1 of this study. In other words, the top-ranked schools in terms of research funding also showed the highest efficiency. It is noteworthy that time to degree (i.e., shorter time to degree) was also a predictor of overall quality. This finding may be characterized as part of efficiency. The correlation between funded research amount in dollars and efficiency scores was highly significant ($r = 0.7243$). This finding reaffirms that faculty's NIH funding amount or the percentage of faculty with research grants is the main factor in determining the QNDE. It also justifies our use of one as a stand-in for the other. This supports our initial premise that led us to select 72 schools on the basis of NIH funding amounts in this study.

As to why annual faculty publications and number of citations were not significant predictors of overall quality, we can only conjecture. It may be because of the fact that these two elements (publications and citations) may indicate quantitative rather than qualitative measures. Regardless of where the explanation may reside, it reflects our difficulty in obtaining precise data on these matters sufficient for drawing definitive conclusions.

We note that NIH funding distribution patterns have changed since 2007, the last year for which data were used for this study. The number of schools that received NIH funding beyond our study period (2004–07) varied greatly in the years since, showing a declining trend (63 schools in 2008, 36 in 2009, and 38 in 2010). Data for the years 2008 through 2010 were obtained by ranking the sum of total funding of each school using the NIH Web site (http://report.nih.gov/index.aspx) because the rankings were not provided by the Web site during this period. Such a pattern indicates that maintaining NIH funding level is a real challenge to many schools and that it requires serious commitment by the schools and the universities to foster faculty's research and scholarly performance and productivity. Supportive academic environments for students' learning (program domain) and faculty's mentorship for students' scholarly development (faculty domain) of the QNDE may hold the key for successful research funding and attaining a high QNDE. Hence, faculties are advised to not only obtain funding from external sources such as NIH but also that colleges of nursing may develop policy on student mentoring, both of which would help improve the QNDE.

### Conclusions

This study assessed the QNDE as a measure of nursing doctoral program quality comprehensively by examining the perspectives of student/graduates and faculty and taking into account factors pertaining to the schools themselves. The results of this study suggest that the quality of doctoral nursing education in the United States is generally good, although there is room for improvement, for all groups in the four quartiles, especially those in the lower quartiles. Specifically, the faculty's mentorship of students' research and scholarly development and supportive academic environment for students' learning are the two areas that deserve special attention for quality improvement. This study showed that annual publications and citations were not significant indicators of faculty quality but faculty members' research grants were. Schools that had the highest research funding also showed the highest efficiency in scholarly performance. Most of the efficient schools indicated that they have leadership, infrastructure, and mechanisms in place that help faculty's scholarly performance and overall doctoral program quality.

Several limitations of the study can be noted. One relates to the questionnaire to deans/department heads. Some items lacked clarity and specificity, even though these were examined by content experts and pilot tested before use. Another limitation was procedural, in that we relied on school liaisons for the implementation of data collection, such as for sending out initial invitation letters to the three groups (faculty, students, and graduates) with two follow-up letters at 2-week intervals; these features made the data collection uneven, and investigators had no
control over whether the designated liaisons were in fact sending the reminders in a timely manner or responding to any queries accurately. This may have contributed to low response rates. Another consideration was that some institutions may not keep detailed data on matters such as the number of citations of faculty publications per year; as well, many institutions may not have up-to-date contact information on their graduates and may rely on central institutional data, which may be several years old—a situation that may account for the excessive length of time it took to obtain responses. Lack of sufficient data provided by the institutions led the investigators to employ nationally available data such as CHE; in some instances, these did not always match the period of interest for the study. Thus, caution needs to be used in interpreting the results. We recommend that individual schools collect on regular basis some of the questions that have bearing on the QNDE such as time to graduation, the level of funding students received, whether they studied full time or part time, and type of position they obtained upon graduation. Evaluation of the QNDE could have been more complete if up-to-date data about scholarly performance and other areas were available at the institutional level or in the literature. We recommend further research in this area in order to gain a fuller understanding of the QNDE.

References


Nagata, S., Gregg, M. F., Miki, Y., Arimoto, A., Murashima, S., & Kim, M. J. (2012). Evaluation of doctoral nursing education in Japan by students, graduates, and faculty: A comparative study based on a cross-sectional questionnaire survey. Nurse...
