Nurse Value-Added and Patient Outcomes in Acute Care

Olga Yakusheva, Richard Lindrooth, and Marianne Weiss

Objective. The aims of the study were to (1) estimate the relative nurse effectiveness, or individual nurse value-added (NVA), to patients' clinical condition change during hospitalization; (2) examine nurse characteristics contributing to NVA; and (3) estimate the contribution of value-added nursing care to patient outcomes.

Data Sources/Study Setting. Electronic data on 1,203 staff nurses matched with 7,318 adult medical–surgical patients discharged between July 1, 2011 and December 31, 2011 from an urban Magnet-designated, 854-bed teaching hospital.

Study Design. Retrospective observational longitudinal analysis using a covariate-adjustment value-added model with nurse fixed effects.

Data Collection/Extraction Methods. Data were extracted from the study hospital's electronic patient records and human resources databases.

Principal Findings. Nurse effects were jointly significant and explained 7.9 percent of variance in patient clinical condition change during hospitalization. NVA was positively associated with having a baccalaureate degree or higher (0.55, $p = .04$) and expertise level (0.66, $p = .03$). NVA contributed to patient outcomes of shorter length of stay and lower costs.

Conclusions. Nurses differ in their value-added to patient outcomes. The ability to measure individual nurse relative value-added opens the possibility for development of performance metrics, performance-based rankings, and merit-based salary schemes to improve patient outcomes and reduce costs.

Key Words. Nurse value-added, patient outcomes, nurse education

Robust research evidence of the contribution of nurses to patient outcomes has been generated from analyses of either nursing staffing intensity (hours-per-patient-day, staffing ratios) or characteristics of nurses (proportion of baccalaureate-educated nurses, experience), all measured at the hospital or unit level (Needleman et al. 2002; Aiken et al. 2003; Weiss, Yakusheva, and Bobay 2011; Blegen et al. 2013; McHugh and Ma 2013). This approach, while valuable from a workforce development and deployment perspective, assumes that all nurses in the aggregate pool (hospital, unit) are of equal value in achieving patient outcomes. It does not address two key questions: *Do all nurses have the*

same value in achieving positive patient outcomes? If not, what are the characteristics of high-performing nurses? Conceptualization and measurement of the value of each individual nurse's incremental contribution to improved patient-level outcomes has not been possible without a theoretical lens for studying marginal productivity and, until recently, in the absence of a practical way of linking patients with their direct care nurses in electronic hospital information systems.

The construct of "*Value-added*," or incremental increase in the value of output attributable to a productive input, is one of the fundamental concepts of economics. When applied to measuring a worker's contribution, it is usually referred to as "the marginal product of labor" or simply "*labor productivity*." Labor productivity, in turn, is determined by the worker's *human capital*, a term that reflects workers' skills, which are acquired through education and experience and enhanced by innate aptitude and ability. The classical human capital theory postulates that, conditional on the worker's innate characteristics, acquisition of additional human capital through education or experience will raise the worker's marginal productivity and increase the value-added (Becker 1962, 2009).

A collection of data analysis methods associated with this theoretical perspective is called value-added modeling, or *VAM* (McCaffrey et al. 2003). VAM was developed for the empirical measurement of productivity, or value-added, of an individual worker relative to other workers participating in the same production process. The relative ranking of workers based on their individual value-added determines *worker effectiveness*, with high value-added workers deemed more effective. This manuscript adapts this theoretical and methodological approach to the delivery of health care, by applying, for the first time, the concept of "value-added" and the methods of VAM to the measurement of nurse effectiveness (i.e., relative value-added) in improving in patient outcomes.

In the discipline of economics, VAM approaches to measuring the independent contributions of individual providers to recipient outcomes have been well developed and tested. These quantitative outcomes-based approaches have found practical applications in educational assessment of teacher effectiveness on the basis of student achievement (Murnane 1975;

Address correspondence to Olga Yakusheva, Ph.D., Division of Systems Leadership and Effectiveness Science, School of Nursing, Department of Health Management and Policy, School of Public Health, University of Michigan, 400 North Ingalls Building, Ann Arbor, MI 48109-5482; e-mail: yakush@umich.edu. Richard Lindrooth, Ph.D., is with the Department of Health Systems, Management and Policy, School of Public Health, University of Colorado, Aurora, CO. Marianne Weiss, D.N.Sc., R.N., is with the College of Nursing, Marquette University, Milwaukee, WI.

Hanushek 1979; Sanders and Rivers 1996; Sanders, Saxton, and Horn 1997; Wright, Horn, and Sanders 1997; Rivkin, Hanushek, and Kain 2000; Rowan, Correnti, and Miller 2002; McCaffrey et al. 2003). With direct parallels to nurse effectiveness and patient outcomes, VAM has potential to demonstrate the importance of individual nurses as a source of variance in patient outcomes and to produce actionable implications for health care reform through improved nurse evaluations or as part of performance-based accountability (McCaffrey et al. 2003). The methodological advantage of VAM lies in the use of statistical techniques to separate the effects of nurses from other factors; this is critical for accountability systems to work as intended. The advantage from a policy perspective is the possibility of linking differences in nurse effectiveness to characteristics of nurses, providing evidence for policy actions to invest in acquired nurse characteristics, such as education, with the goal of increasing nurse value-added (NVA) and effectiveness in achieving health care outcomes.

Identification of characteristics of high value-added nurses can help hospitals target investments in workforce development to increase the number of nurses with high value-added characteristics. A growing body of evidence relates higher average levels of nursing human capital (in a hospital or unit) to improved patient outcomes. These characteristics include a higher proportion of nurses with a baccalaureate degree (Aiken et al. 2003, 2011, 2014; Blegen et al. 2013), higher average nurse experience (Dunton et al. 2007), nurse certification (Kendall-Gallagher et al. 2011), and nurse dose (composite of education, experience, and skill mix) (Manojlovich et al. 2011). However, it is not known if, among nurses working on the same unit or in the same hospital, individual nurses with a baccalaureate degree or more experienced nurses, are, in fact, more effective in providing patient care than their less-educated and less-experienced peers. Measuring nurse effectiveness at the individual nurse level allows us to relax the implicit assumption that all nurses make equal contributions and examine not only the entire distribution of nurse effectiveness but also allow us to identify human capital characteristics that distinguish high performers from low performers.

Advances in health information technologies now incorporate identification of each care provider at each data point entered into the electronic health record (EHR), thus allowing direct linkages between patients and each individual nurse providing care over the course of hospitalization. Establishing the direct link between individual NVA and a patient's outcomes will demonstrate that the aggregate results are robust to alternative specifications, further strengthening the case for a causal interpretation of the associations consistently found in

a variety of settings. For individual nurses, the ability to measure a nurse's value-added contributions relative to other nurses is a critical, but as yet underdeveloped methodological foundation to performance metrics, performance-based rankings, and pay-for-performance schemes.

      With the objective of understanding the relationship of the relative value-added of individual nurses' contributions to their patients' outcomes, the aims of this study were to (1) estimate the relative nurse effectiveness, or value-added to their patients' clinical condition trajectories during hospitalization; (2) examine the nurse-level associations of nurse human capital measures (education, experience, expertise) and other nurse characteristics (age, gender) to NVA; and (3) examine, at the patient level, the association of nurse effectiveness to patient outcomes (change in clinical condition, length of stay [LOS], costs, 30-day readmission).

METHODS

The study was a retrospective observational analysis of individual nurse contributions to patient-level outcomes. We applied a VAM approach using longitudinal electronic inpatient clinical data to estimate nurse effects (McCaffrey et al. 2003).
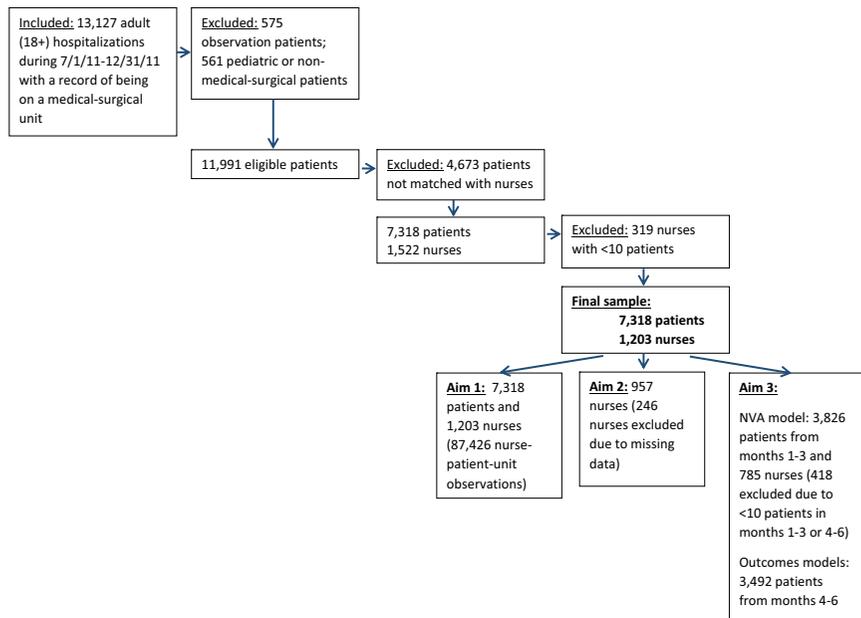
Data Sources

Following IRB review, we extracted deidentified data from electronic databases at the study hospital. We matched patient and nurse data using the nurse assessment inputs into the EHR database that retains the identifier of each nurse inputting data for each patient. Patient data were extracted from the financial database (patient characteristics), EHRs (clinical condition updates, nurse assessment inputs with nurse identifiers), and the Patient Activity Database (PAD) that records admission/discharge/transfers (unit placements). The nurse identifiers were used to extract nurse characteristics from the human resource database (nurse education, expertise level, age, hospital tenure, gender).

Sample and Setting

The sample was derived from adult nonpsychiatric inpatients with a discharge disposition recorded during the 6-month period between July 1,

2011 and December 31, 2011 at an urban academic medical center accredited by The Joint Commission and designated by the American Nurses Credentialing Center as a Magnet® facility. Our data extraction criteria included all adult (18 and older) patients with a record of being on a medical–surgical unit at any point during their hospitalization episode ($N = 13{,}127$). We then excluded observation patients ($n = 575$) and patients with pediatric and nonmedical/surgical discharge diagnoses ($n = 561$), resulting in 11,991 eligible inpatient discharges. After merging with nurse identifiers, the resulting sample was 7,318 patients and 1,522 nurses. Because we estimate nurse-specific effects, nurses matched with fewer than 10 patients each ($n = 319$) were eliminated. The final sample included 7,318 patient hospitalizations linked to 1,203 nurses, or 61 percent of all eligible inpatient discharges (Figure 1). Compared to all eligible patients, the patients in the final analysis sample were younger (55.7 vs. 58.9, $p < .01$), more likely to be privately insured (35 percent vs. 31 percent, $p < .01$), and less likely to be on Medicare (38 percent vs. 46 percent); the other patient characteristics were nonsignificantly different.

Figure 1:   Sample Selection

Measures

Nurse Characteristics. Data on nurse characteristics were obtained from nurse profiles within the human resources electronic databases. Nurse education level represents the highest level of education attained and was coded as a categorical variable: Diploma, Associate's degree, and Bachelor's degree or higher. Nurse expertise was coded as a four-level categorical variable for each of the expertise level classifications in the hospital's nurse professional development framework based on the work of Benner (1984) (CNI = "Novice or Advanced Beginner"/CNII = "Competent"/CNIII = "Proficient"/CN4 = "Expert"). Nurse age and experience, measured as tenure at the hospital, were recorded in years; gender was coded as a dichotomous indicator with female as the reference category.

Patient Outcomes. In-hospital all-cause mortality and same-hospital all-cause 30-day readmission were coded as dichotomous variables with nonoccurrence as the reference category. LOS was computed in days from the time of admission to the time of discharge. Hospitalization cost represented the sum of direct and indirect costs derived from the hospital's billing, payroll, general ledger, and individual departments' resource utilization databases, at the patient level, using cost accounting software (Azoulay et al. 2007).

Clinical condition was measured with the Rothman Index (RI), a summary metric of the patient's clinical condition that integrates 26 clinical parameters available in the EHR, including nurse assessments (e.g., food/nutrition, skin, psycho-social, etc.), vital signs (e.g., temperature, blood pressure, etc.), heart rhythms (sinus bradycardia, sinus rhythm, etc.), and lab tests. Using proprietary software adopted by the study site, the RI tracks clinical condition over the course of hospitalization through real-time updates as new information is entered into the system. The RI score is a subtraction index from 100 with a theoretical range of -69 to 100 and higher values indicating better clinical condition and lower mortality risk. Construct validity has been evaluated in a prior study in three acute care hospitals by examining the relationship of the RI score to other indicators of clinical condition, resulting in area under the receiver operating characteristics curve (AUC) estimates of ≥ 0.92 for six discharge destination categories, AUC ≥ 0.93 when predicting 24 hours mortality, and AUC = 0.62 for 30-day readmission (Rothman, Rothman, and Solinger 2013). RI scores less than 40 were associated with a fivefold increase in the risk of 30-day mortality, and RI scores of less than 80

at discharge were associated with a nearly 2.5-fold risk of an unplanned readmission (Rothman et al. 2012, 2013; Bradley et al. 2013; Rothman, Rothman, and Solinger 2013). For this study, we used the first recorded RI as the admission RI, the last recorded score as the discharge RI, and the difference between admission and discharge RI as the change in clinical condition during hospitalization.

Patient Control Variables. Patient control variables were consistent with controls used in earlier studies (Kansagara et al. 2011): demographic characteristics (age in years, gender), insurance type (private/Medicare/Medicaid/none), an indicator for a surgical admission (based on the attending physician's department of record), and an indicator for a hospitalization within 30 days prior to the index admission. To ensure that the nurse contributions were not confounded by systematic variation in nurse characteristics and patient outcomes across different patient types and other factors (such as contributions of nonnurse medical team members and time trends), all models included sets of fixed effects for unit assignment, diagnosis, medical team assignments, and calendar week and day. Unit assignment fixed effects were based on the patient's location record from the PAD at the time of contact with each nurse (categorized using 33 unique "floor" PAD codes); discharge diagnosis groups were categorized based on the primary discharge diagnosis using the Agency for Healthcare Research and Quality Clinical Classification Software (49 categories); medical team assignment fixed effects were based on 43 codes from the EHR that designate the physician team responsible for the patient's care during the hospitalization; calendar week fixed effects (27–52) and day of the week fixed effects (1–7) were categorized based on the discharge date. For the diagnosis fixed effects, categories with 10 or more patients are assigned a unique fixed effect; the rest are combined into the respective reference categories. Fixed effects categories with a minimum of 10 observations per category reduced the finite sample bias related to inclusion of fixed effects in a nonlinear model (Greene 2004).

Analysis Methods

Aim 1. To obtain the NVA estimates, we used a covariate-adjustment value-added estimation technique with nurse fixed effects (Murnane 1975; Hanushek 1979; Sanders and Rivers 1996; Sanders, Saxton, and Horn 1997; McCaffrey et al. 2003). We estimated a linear regression of the change in

patient's RI score from admission to discharge on the set of nurse fixed effects, conditional on the patient's admission RI score and adjusting for the patient's characteristics, and the sets of fixed effects capturing non-nurse-related factors (unit, diagnosis, medical team assignment, calendar week and day). This allowed us to partition variation in the change in the clinical condition during hospitalization into the component explained by individual nurse fixed effects (NVA contributions), the component attributable to non-nurse-related factors (patient characteristics and other care-related factors like unit effects or time trends), and the unexplained error term.

Because patients were linked with multiple nurses and many were on multiple units during the course of hospitalization, we estimated the model at the patient–nurse-unit level. This means that a patient who received care on two units and was linked to two nurses on each unit would represent four observations, one for each patent–nurse-unit match. We estimated the model using the "survey" prefix in Stata11.0 using inverse frequency weights and clustering at the nurse level, to ensure that the estimates are not biased due to the structural properties of these data. The joint contribution of nurse fixed effects was examined using the change in the R-squared attributable to the fixed effects and the multiple F-tests of joint significance.

In fixed effects VAM, nurse fixed effect coefficients are estimated relative to the "reference nurse" arbitrarily omitted from the regression. To correct for bias associated with the selection of a particular nurse as the arbitrary reference category, the fixed effects were transformed by subtracting their mean. The transformation centered the estimates around zero, while preserving the relative ranking as well as the scale of differences among nurses (Murnane 1975; Hanushek 1979).

Aim 2. The transformed coefficients of the nurse fixed effects were used as the estimates of individual nurses' NVAs to examine the association of NVA with nurse characteristics (education, expertise level, tenure at the hospital, age, and gender). A nurse-level multivariate linear regression of NVA on the nurse characteristics was used, controlling for unit fixed effects and adjusting for error term correlation at the unit level. Nurses with a missing record of education or expertise level ($n = 246$) were excluded from the analysis for this aim.

Aim 3. We estimated the models for patient outcomes as functions of the NVA of the patient's direct care nurses. This was achieved in a four-step

process designed to test for, and minimize, statistical biases from circular use of data and nonrandom nurse assignments.

First, because NVAs in Aim 1 were estimated from a change in patient RI scores, regressing patient outcomes back on NVAs could lead to incorrect estimates (Kupermintz 2002). To avoid this, we split the sample into two partitions based on the discharge date (months 1–3, $n = 3,826$/ months 4–6, $n = 3,492$) and obtained two sets of NVAs, one for each partition, using the VAM approach described in Aim 1. We examined the persistence of NVA between the two partitions using a correlation coefficient and a two-way scatter plot. Nurse fixed effect coefficients from the first partition were then used as measures of NVA for patients from the second partition. This method required that each nurse be linked to patients in both partitions; therefore, 418 nurses were excluded from the analysis for this aim, leaving 785 nurses.

Second, for the second partition ($n = 3,492$), we constructed a patient-level measure of NVA using a weighted average of nurse fixed effect estimates from the first partition, weighted by the relative frequencies of the patient's contact with each of the nurses during the patient's hospitalization. The number of direct care nurses used in computing patient-level NVA is correlated with hospital LOS because as the number of days in the hospital increases, the exposure to more nurses is also more likely. This causes patients with the shortest LOS to have the highest and the lowest NVA values and could bias the results (Yakusheva, Lindrooth, and Weiss 2014). To control for this bias, we included a cubic polynomial in the number of direct care nurses as additional controls in all models for this Aim.

Third, random assignment is a key identifying assumption of VAM (McCaffrey et al. 2003). We tested whether patients were randomly assigned to nurses, by estimating a linear regression of the patients' admission RI scores on NVA, conditional on patient characteristics, and diagnosis, unit, medical team assignment, and calendar fixed effects, and adjusted for correlation of the error term at the nurse level. For patients hospitalized on multiple units, unit fixed effects were categorized to include all unit assignments (Yakusheva, Lindrooth, and Weiss 2014). A lack of statistical significance of the NVA coefficient is consistent with random assignment.

Finally, we estimated multivariate models to examine the association of NVA with patient outcomes using the same approach. A linear regression model was used for the change in clinical condition and a logistic model was used for readmission. A GLM model with log-links was used for LOS and

costs, to reduce bias due to positively skewed distributions of the variables (Manning 1998; Manning and Mullahy 2001).

RESULTS

The sample was 50 percent male, with a mean age of 56 years, and comprised of 60 percent medical and 40 percent surgical patients (Table 1). The average RI score at admission was 75 (ranging from −23 to 99), which was consistent with other studies that used the RI score (Rothman et al. 2012; Bradley et al. 2013; Rothman, Rothman, and Solinger 2013). Fourteen percent of the sample had a prior hospitalization within 30 days of admission. The in-hospital mortality rate was 1.4 percent, and the readmission rate was 16 percent. Means for LOS and costs were 5.5 days (median, 3 days) and $20,203 (median, $10,949); the skewness test statistic was 14.80 for LOS and 9.60 for cost, suggesting a high degree of positive skewness. On average, patients were linked to 4.98 nurses over the course of hospitalization. The average nurse age was 35 years, and the average tenure at the hospital was 6 years. Close to 70 percent of the nursing staff with available education data had a baccalaureate degree or above.

Aim 1

The nurse fixed effects were jointly significant at $p < .01$ and explained 7.9 percent of variance in patient clinical condition trajectories. The transformed estimated coefficients of nurse fixed effects ranged from −23.46 to 22.0, with the standard deviation of 3.90. The frequency distribution of the fixed effects was symmetrical (sk-test = 0.086) and had a positive excess kurtosis (8.52); Shapiro–Wilk test rejected normality at $p < .01$ (Figure 2). Nurses ranked in the top 20 percent had an average NVA of 5.53 RI points, and nurses in the bottom 20 percent had an average NVA of −4.85 RI points.

Aim 2

Nurse characteristics explained 7.0 percent of variance in NVA (Table 2). NVA was positively associated with nurse education at the baccalaureate or higher levels. Relative to nurses with a diploma, nurses with a baccalaureate degree or higher had 0.55 ($p = .04$) RI points higher NVAs; NVA at the associate level was not significantly different from the NVA at the diploma level.

Table 1:   Descriptive Statistics of Sample

Variable	N (%)
Nurses, *N* = 1,203	
Age	
Mean (SD)	35.69 (10.93)
Education	
Diploma	100 (8.3)
Associate's degree	200 (16.6)
Bachelor's degree	611 (50.1)
Master's/doctoral degree	46 (3.8)
Unknown	246 (20.4)
Expertise level	
CN1	294 (29.3)
CN2	600 (49.9)
CN3	183 (18.2)
CN4	33 (2.7)
Unknown	93 (7.7)
Tenure at the hospital	
Mean (SD)	6.20 (8.1)
Patients, N = 7,318	
Age	
Mean (SD)	55.66 (19.32)
Sex	
Male	3,699 (50.5)
Female	3,619 (49.5)
Insurance type	
Medicare	2,592 (35.4)
Medicaid	1,692 (23.0)
Blue cross/commercial	2,786 (38.1)
Other/uninsured	248 (3.4)
Service type	
Medical	4,348 (59.4)
Surgical	2,970 (40.6)
Hospitalization within 30 days prior to admission	
No	6,319 (86.3)
Yes	999 (13.7)
Rothman index (RI) at admission	
Mean (SD)	75.78 (18.12)
RI at discharge	
Mean (SD)	78.59 (17.24)
RI change	
Mean (SD)	2.81 (14.88)
Patient died during hospitalization	
No	7,216 (98.6)
Yes	102 (1.39)

Table 1. *Continued*

Variable	N (%)
Patient readmitted within 30 days	
No	6,170 (84.3)
Yes	1,148 (15.7)
Length of stay, days	
Mean (SD)/median/retransformed mean*	5.54 (7.57)/3/3.48
Total cost of hospitalization, $	
Mean (SD)/median/retransformed mean*	20,203.48 (30,381.84)/10,949.25/11,887.67

*Retransformed smearing-adjusted mean of the log-transformed variable.

Figure 2:   Frequency Distribution of Nurse Value-Added (NVA), $N = 1,203$
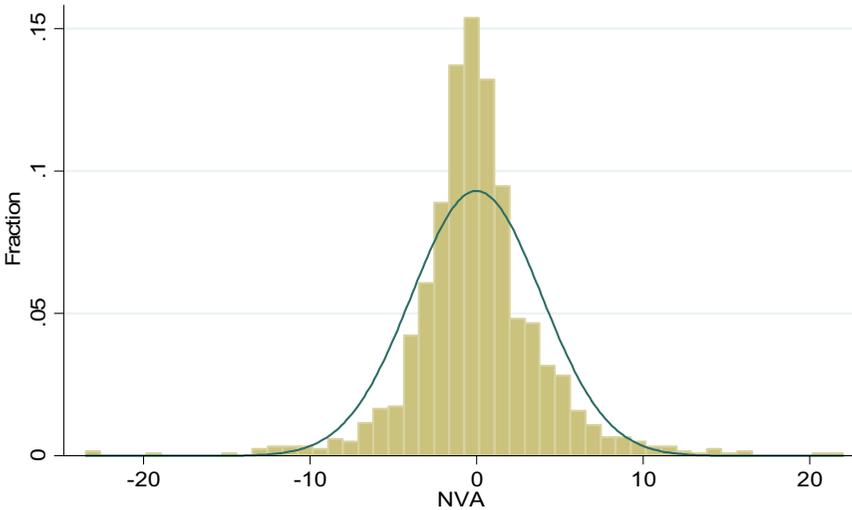
Note. The figure shows a frequency histogram of the standardized nurse fixed effect estimates, using change in patient condition as the dependent variable and a covariate adjusted value-added modeling (VAM) regression with nurse fixed effects. Other controls include patient characteristics and diagnosis, unit, care team, and calendar fixed effects. The superimposed distribution is normal with the same first and second moments.

There was also a significant effect of nurse expertise level; compared to "novice" or "advanced beginner" nurses, being at the "Proficient" or "Expert" levels was associated with a 0.66 ($p = .03$) RI points higher NVA. Nurse age, gender, and hospital tenure were nonsignificant.

Table 2:   Nurse Characteristics and Nurse Value-Added (NVA), $N = 957$

Variables	NVA
Education: Associate degree	0.364 (0.23)
Education: Bachelor degree or higher	0.554* (0.04)
Expertise level: CNII	−0.004 (0.99)
Expertise level: CNIII/CNIV	0.657* (0.03)
Tenure at hospital	0.036 (0.08)
Age	−0.003 (0.75)
Sex: male	0.298 (0.32)
R-squared	0.070

Note. Shown are regression coefficients from an ordinary least squares regression NVA and nurse characteristics using the sample of N = 957 nurses; the model also included unit level fixed effects; the estimates are adjusted for correlated errors at the unit level.
Robust *p*-values are in parentheses, *<.05.

Aim 3

NVA estimates obtained from the two partitions were significantly and positively correlated between the two partitions (nurse-level correlation coefficient, 0.23, $p < .01$). A two-way scatter plot showed an overall positive linear trend with heteroscedastic variance (Cook–Weisberg chi-square test, 232.85, $p < .01$). The coefficient of the NVA in the analysis of admission RI scores was nonsignificant ($−0.013$, $p = .36$), suggesting that nurses were essentially randomly assigned to patients. Controlling for patient characteristics and diagnosis, unit assignment, medical team, and calendar fixed effects, a one standard deviation increase in patient-level average of NVA (SD = 2.02, in the first partition) was associated with a 6.5 percent ($p < .01$) reduction in costs of hospitalization, and a 2.6 percent ($p < .01$) reduction in the LOS (Table 3). The effect of NVA on the RI change and 30-day readmission was nonsignificant.

DISCUSSION

The analysis of Aim 1 produced evidence of significant individual variation in nurse effectiveness, measured as relative value-added. While most nurses' value-added estimates clustered tightly around the mean, high-performing nurses (top 20 percent) had, on average, approximately 10 RI points greater value-added compared to low-performing nurses (bottom 20 percent), providing evidence that high performers produce greater improvement in their patients' clinical condition. With the average patient's end-of-hospitalization RI being close to 75 points, the 10-point difference between high- and low-per-

Table 3:   Exposure to Nurse Value-Added (NVA) and Patient Outcomes, $N = 3,492$

Variable	(1) RI Change	(2) Cost	(3) LOS	(4) Readmission
NVA, z-score	0.025 (0.10)	−0.065** (<0.01)	−0.026** (<0.01)	0.878 (0.09)
Patient characteristics				
RI score at admission	−0.529** (<0.01)	−0.052** (<0.01)	−0.037** (<0.01)	0.834* (<0.01)
Sex: Female	−0.099** (<0.01)	0.050* (0.0369)	0.047** (<0.01)	0.867 (0.263)
Age: 45–54	0.011 (0.745)	0.139** (<0.01)	−0.047** (<0.01)	0.956 (0.809)
Age: 55–64	−0.093* (0.0108)	−0.034 (0.325)	−0.050** (<0.01)	0.693 (0.0684)
Age: 65–74	−0.128** (<0.01)	0.092* (0.0149)	−0.077** (<0.01)	0.636 (0.0665)
Age: 75+	−0.439** (<0.01)	−0.034 (0.483)	−0.103** (<0.01)	0.633 (0.0724)
Type of admission: surgical	0.389 (0.0779)	0.400* (0.0372)	0.145 (0.0621)	1.373 (0.0635)
Prior hospitalization: no	0.158** (<0.01)	−0.035 (0.312)	−0.079** (<0.01)	0.457** (<0.01)
Goodness-of-fit	0.534	1.241	1.025	0.196

Note. Shown are regression coefficients from an ordinary least squares regression (1), GLM regression (2)-(3), and odds ratios from a logistic regression (4), using the sample of $n = 3,492$ patients hospitalized during the months 4–6 of the study. The goodness-of-fit statistic is adjusted R-squared for model (1), scaled Pearson chi-squared for models (2) and (3), and pseudo-*R*-squared for model (4). LOS and cost coefficients represent approximate percentage changes. Other controls included are diagnosis, unit assignment profile, and calendar week and day fixed effects. The estimates are adjusted for correlated errors at the nurse level.
Robust *p*-values are in parentheses, *<.05, **<.01.

forming nurses could have significant implications for the patient's risk of mortality (RI <65) and readmission (RI <80) (Rothman, Rothman, and Solinger 2013). Split-sample analysis revealed a significant correlation in NVA between the samples, indicating that the individual nurse effects were persistent and nurse-specific. The observed differences among nurses, and their robustness across samples, serve as "proof of concept" for measuring relative NVA, lending credibility to the idea of using similar approaches to create unique and meaningful individual nurse effectiveness rankings.

Significant differences in individual NVAs were observed in Aim 2 by the nurse characteristics. NVA was strongly and positively associated with baccalaureate education, suggesting that investments in upgrading nurse education and preferential hiring practices have a potential of improving the value-added of nursing care. This finding provides evidence at the individual nurse–patient level, expanding and validating hospital-aggregate level results

and improving causal inference about the relationship of BSN education to improved patient outcomes. Nurse expertise level was also positively associated with NVA. The concept of expertise has been difficult to measure as hospital classifications are typically organization-specific as is the classification in this study. A few studies have explored expertise, often using experience and education as surrogate metrics (McHugh and Lake 2010); however, no prior studies using expertise classifications have found quantitative evidence of an impact on patient outcomes. The dual findings of positive associations of education and expertise levels with NVA in this study speak to the importance of upgrading the nurse workforce to BSN level and to organizational practices focused on developing and retaining expert nurses. However, the nurse characteristics, combined, explained only 7 percent of the variance in NVA, pointing out the potential importance of other nurse attributes not captured in this study (e.g., aptitude, ability).

The findings for Aim 3 highlight the role of high value-added nursing in improving patient outcomes. Patients who received care from high value-added nurses had shorter hospitalizations and lower costs. While the effects were small, the split-sample design resulted in measuring only the effect of the time-invariant component of nurse effectiveness, eliminating any nurse-specific time-variant (e.g., experience and education increasing over time) or transient (e.g., variation in overtime hours, or illness) factors. Our estimates therefore are likely to be conservative.

Collectively, the three study aims contribute to a more granular understanding of the nurses' contributions to patient outcomes. Aim 1 shows that nurses differ in their value-added and effectiveness as direct care providers to patients; Aim 2 reveals that these differences are only partly explained by education and expertise, and that many factors that distinguish high-performing nurses are yet to be measured and explained; Aim 3 demonstrates that the differences in individual nurse effectiveness may have nontrivial implications for patient outcomes and costs. Taken together, the results highlight the need to recognize individual nurses as each being a unique and integral component of the care-delivery process. Achieving high-quality patient care and reducing costs will require that high-performing nurses are identified, recognized, and rewarded.

The NVA approach could potentially be applied to performance measurement and compensation enhancement models. Traditionally, performance appraisals use institutionally developed criteria (Weathers and Raleigh 2013) and are conducted through communication with managers, peer review, and self-assessment of performance (Stonehouse 2013). While certification has been used as a metric for competency and has been associated

with improved patient outcomes (Kaplow 2011; Kendall-Gallagher et al. 2011), the NVA approach is more granular, producing individualized nurse performance assessments. The potential for unintended consequences of this approach needs careful exploration in considering its utility for application in performance evaluation and merit payment schemes.

The study was subject to several limitations. First, the RI score was based on clinical indicators and nurse assessments only; these point-in-time measures of the patient's clinical condition serve as proxies, but not direct measures of severity of illness. In addition, primary diagnosis fixed effects do not control for secondary diagnoses and complications. However, variation in secondary diagnoses and procedures are captured, in part, by variation in the RI. Insufficient controls for severity of illness would bias our estimates downward (i.e., toward the null hypothesis) if high performers were systematically assigned to more severely ill patients and will likely lead to underestimates of the importance of NVA and human capital. Second, the nurse experience variable reflects only hospital-specific human capital because it was derived from nurse tenure at the hospital; the total amount of experience was not available in the study data. Specialty certification was not available in the dataset and could not be evaluated as a nurse characteristic related to expertise. Not fully capturing experience and certification could bias our results in Aim 2; however, the inclusion of the nurse's age in the Aim 2 specification will in part control for experience at other hospitals. Third, the estimates of value-added of high-performing nurses could be biased down by low-performing nurses caring for the same patient, and, conversely, the contributions of low-performing nurses could be biased upward by high-performing nurses. However, because nurses in our sample were matched, on average, with 108 patients, these biases are not likely to affect relative rankings of nurses in the distribution of nurse effectiveness. Moreover, the potential underestimation of high performers (and overestimation of low performers) means that the true underlying differences in value-added between high- and low-performing nurses are likely higher than our study estimates, and their implications for patient outcomes are likely greater. Fourth, the split-sample approach was used to ensure that the estimates are not subject to reusing the same data and are not biased due to the effects of unobserved nonnurse factors (e.g., seasonal trends). While this enabled us to reduce potential bias, it came at the cost of reduced efficiency due to a smaller estimation sample. Thus, our estimates are conservative and further contribute to understated statistical significance of NVA for patient outcomes. Fifth, the readmission measure did not include return visits to other

facilities, which accounted for 19 percent of heart failure readmissions in 2006 CMS data (Nasir et al. 2010) and could have further contributed to the lack of significant effects on readmission. Sixth, hospitalization cost computation uses a complex formula of hospital-wide and unit fixed cost allocations as well as direct care costs, potentially causing a measurement error in the cost model. Any bias resulting from unit variation in allocation method is minimized by inclusion of unit-level fixed effects. If the error is nonsystematic, our approach is subject to attenuation bias but practical in large-sample analyses (Azoulay et al. 2007). Lastly, the study data came from a single facility with a high proportion of BSN-educated nursing staff, limiting generalizability of our findings to other settings. The effect of NVA may be substantially different at hospitals with different quality improvement processes, management practices, cost measurement methods, and other unmeasured characteristics that may be correlated with the outcomes and, at the same, time influence the effectiveness of nursing care.

Despite these limitations, this paper presents a novel approach to assessing differences in nurse productivity and understanding its relationship with measured attributes of human capital, and to patient outcomes; unmeasured variables make our estimates conservative. NVA and human capital would likely continue to be an important determinant of outcomes in future studies that have access to more variables. We hope that this approach serves as the basis for future studies using different samples, outcomes, and controls, so that the validity of the VAM approach and generalizability of our results to other hospitals and work environments can be assessed.

Conclusions

The findings of this study strengthen the body of evidence in support of strategic federal, state, and health system initiatives to build toward a baccalaureate-educated workforce. Identification of baccalaureate education as a characteristic of high value-added nurses, and targeting investments in workforce development to ensure that a complement of high value-added nurses are deployed to direct patient care, could improve patient outcomes and reduce costs of care.

Further refinement and use of NVA methods is an innovative approach to measuring nurse effectiveness and differentiating nurses based on their relative value-added. Performance ranking and performance-based compensation are possible future applications of this novel approach, but they need careful scrutiny.

ACKNOWLEDGMENTS

REFERENCES

Aiken, L. H., S. P. Clarke, R. B. Cheung, D. M. Sloane, and J. H. Silber. 2003. "Educational Levels of Hospital Nurses and Surgical Patient Mortality." *Journal of the American Medical Association* 290 (12): 1617–23.

Aiken, L. H., J. Cimiotti, D. M. Sloane, H. L. Smith, L. Flynn, and D. Neff. 2011. "The Effects of Nurse Staffing and Nurse Education on Patient Deaths in Hospitals with Different Nurse Work Environments." *Medical Care* 49 (12): 1047–53.

Aiken, L. H., D. M. Sloane, L. Bruyneel, K. Van den Heede, P. Griffiths, R. Busse, M. Diomidous, J. Kinnunen, M. Kózka, E. Lesaffre, M. D. McHugh, M. T. Moreno-Casbas, A. M. Rafferty, R. Schwendimann, P. A. Scott, C. Tishelman, T. van Achterberg, and W. Sermeus. 2014. "Nurse Staffing and Education and Hospital Mortality in Nine European Countries: A Retrospective Observational Study." *The Lancet* 383: 1824–30. Early Online Publication, February 26, 2014. doi:10.1016/S0140-6736(13)62631-8.

Azoulay, A., N. M. Doris, K. B. Filion, J. Caron, L. Pilote, and M. J. Eisenberg. 2007. "The Use of the Transition Cost Accounting System in Health Services Research." *Cost Effectiveness and Resource Allocation* 5 (11): 1–9.

Becker, G. S. 1962. "Investment in Human Capital: A Theoretical Analysis." *Journal of Political Economy* 70 (5): 9–49.

——— 2009. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. London: University of Chicago Press.

Benner, P. 1984. *From Novice to Expert: Excellence and Power in Clinical Nursing Practice*. Menlo Park, CA: Addison-Wesley.

Blegen, M. A., C. J. Goode, S. H. Park, T. Vaughn, and J. Spetz. 2013. "Baccalaureate Education in Nursing and Patient Outcomes." *Journal of Nursing Administration* 43 (2): 89–94.

Bradley, E., O. Yakusheva, L. I. Horwitz, H. Sipsma, and J. Fletcher. 2013. "Identifying Patients at Increased Risk for Unplanned Readmission." *Medical Care* 51 (9): 761–6.

Dunton, N., B. Gajewski, S. Klaus, and B. Pierson. 2007. "The Relationship of Nursing Workforce Characteristics to Patient Outcomes." *Online Journal of Issues in Nursing* 12 (3): 3.

Greene, W. 2004. "The Behaviour of the Maximum Likelihood Estimator of Limited Dependent Variable Models in the Presence of Fixed Effects." *Econometrics Journal* 7: 98–119.

Hanushek, E. 1979. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *Journal Human Research* 14 (1): 351–88.

Kansagara, D., H. Englander, A. Salanitro, D. Kagen, C. Theobald, M. Freeman, and S. Kripalani. 2011. "Risk Prediction Models for Hospital Readmission: A Systematic Review." *Journal of the American Medical Association* 306 (15): 1688–98.

Kaplow, R. 2011. "The Value of Certification." *AACN Advanced Critical Care* 22 (1): 25–32.

Kendall-Gallagher, D., L. H. Aiken, D. M. Sloane, and J. P. Cimiotti. 2011. "Nurse Specialty Certification, Inpatient Mortality, and Failure to Rescue." *Journal of Nursing Scholarship* 43 (2): 188–94.

Kupermintz, H. 2002. *Teacher Effects as a Measure of Teacher Effectiveness: Construct Validity Considerations in TVAAS (Tennessee Value-Added Assessment System).* CSE Technical Report 563. Los-Angeles: Center for the Study of Evaluation, University of California.

Manning, W. G. 1998. "The Logged Dependent Variable, Heteroscedasticity, and the Retransformation Problem." *Journal of Health Economics* 17 (3): 283–95.

Manning, W. G., and J. Mullahy. 2001. "Estimating Log Models: To Transform or Not to Transform." *Journal of Health Economics* 20 (4): 461–94.

Manojlovich, M., S. Sidani, C. L. Covell, and C. L. Antonakos. 2011. "Nurse Dose: Linking Staffing Variables to Adverse Patient Outcomes." *Nursing Research* 60 (4): 214–20.

McCaffrey, D. F., J. R. Lockwood, D. M. Koretz, and L. S. Hamilton. 2003. *Evaluating Value-Added Models for Teacher Accountability.* Santa Monica, CA: RAND.

McHugh, M. D., and E. T. Lake. 2010. "Understanding Clinical Expertise: Nursing Education, Experience, and the Hospital Context." *Research in Nursing and Health* 33 (4): 276–87.

McHugh, M. D., and C. Ma. 2013. "Hospital Nursing and 30-Day Readmissions among Medicare Patients with Heart Failure, Acute Myocardial Infarction, and Pneumonia." *Medical Care* 51: 52–9.

Murnane, R. J. 1975. *The Impact of School Resources on the Learning of Inner City Children.* Cambridge, MA: Ballinger Publishing Co.

Nasir, K., Z. Lin, H. Bueno, S. Normand, E. E. Drye, P. Keenan, et al. 2010. "Is Same-Hospital Readmission Rate a Good Surrogate for All-Hospital Readmission Rate?" *Medical Care* 48 (5): 477–81.

Needleman, J., P. Buerhaus, S. Mattke, M. Stewart, and K. Zelevinsky. 2002. "Nurse-Staffing Levels and the Quality of Care in Hospitals." *New England Journal of Medicine* 346 (22): 1715–22.

Rivkin, S. G., E. A. Hanushek, and J. F. Kain. 2000. NBER Working Paper #W6691. *Teachers, Schools, and Academic Achievement.* Cambridge, MA: National Bureau of Economic Research.

Rothman, S. I., M. J. Rothman, and A. B. Solinger. 2013. "Placing Clinical Variables on a Common Linear Scale of Empirically-Determined Risk: A Step Toward

Construction of a General Patient Condition Score from the Electronic Health Record." *BMJ Open* 3 (5): e002367. doi:10.1136/bmjopen-2012-002367.

Rothman, M. J., A. B. Solinger, S. I. Rothman, and G. D. Finlay. 2012. "Clinical Implications and Validity of Nursing Assessments: A Longitudinal Measure of Patient Condition from Analysis of the Electronic Medical Record." *BMJ Open* 2 (4): e000849. doi:10.1136/bmjopen-2012-000849.

Rothman, M. J., S.I. Rothman, and J. Beals IV. 2013. "Development and Validation of a Continuous Measure of Patient Condition Using the Electronic Medical Record." *Journal of Biomedical Informatics* 46 (5): 837–48.

Rowan, B., R. Correnti, and R. J. Miller. 2002. "What Large-Scale Survey Research Tells Us about Teacher Effects on Student Achievement: Insights from the *Prospects* Study of Elementary Schools." *Teacher College Record* 104: 1525–67.

Sanders, W. L., and J. C. Rivers. 1996. *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement.* Knoxville, TN: University of Tennessee Value-Added Research Center.

Sanders, W. L., A. Saxton, and B. Horn. 1997. "The Tennessee Value-Added Assessment System: A Quantitative Outcomes-Based Approach to Educational Assessment." In *Grading Teachers, Grading Schools: Is Student Achievement a Valid Evaluation Measure?*, edited by J. Millman, pp. 137–62. Thousand Oak, CA: Corwin Press Inc.

Stonehouse, D. 2013. "Appraisal: A Key Tool to Delivering Quality Care." *British Journal of Healthcare Management* 19 (7): 338–41.

Weathers, S. M., and E. D. H. Raleigh. 2013. "1-Year Retention Rates and Performance Ratings: Comparing Associate Degree, Baccalaureate, and Accelerated Baccalaureate Degree Nurses." *Journal of Nursing Administration* 43 (9): 468–74.

Weiss, M. E., O. Yakusheva, and K. L. Bobay. 2011. "Nursing Staffing, Readiness for Hospital Discharge, and Post-Discharge Utilization." *Health Services Research* 46 (5): 1473–94.

Wright, S. P., S. P. Horn, and W. L. Sanders. 1997. "Teacher and Classroom Context Effects on Student Achievement: Implications for Teacher Evaluation." *Journal of Personnel Evaluation in Education* 11: 57–67.

Yakusheva, O., R. Lindrooth, and M. Weiss. 2014. "Economic Evaluation of the 80 percent Baccalaureate Nurse Workforce Recommendation: A Patient-Level Analysis." *Medical Care* 52: 864–69.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.