Circulation: Cardiovascular Quality and Outcomes

ORIGINAL ARTICLE

Examining Hospital Variation on Multiple Indicators of Stroke Quality of Care

Peter C. Austin, PhD
Jiming Fang, PhD
Bing Yu, PhD
Moira K. Kapral, MD, MSc

BACKGROUND: Provider profiling involves comparing the performance of hospitals on indicators of quality of care. Typically, provider profiling examines the performance of hospitals on each quality indicator in isolation. Consequently, one cannot formally examine whether hospitals that have poor performance on one indicator also have poor performance on a second indicator.

METHODS: We used Bayesian multivariate response random effects logistic regression model to simultaneously examine variation and covariation in multiple binary indicators across hospitals. We considered 7 binary patient-level indicators of quality of care for patients presenting to hospital with a diagnosis of acute stroke. We examined between-hospital variation in these 7 indicators across 86 hospitals in Ontario, Canada.

RESULTS: The number of patients eligible for each indicator ranged from 1321 to 14 079. There were 7 pairs of indicators for which there was a strong correlation between a hospital's performance on each of the 2 indicators. Twenty-nine of the 86 hospitals had a probability higher than 0.90 of having worse performance than average on at least 4 of the 7 indicators. Seven of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on at least 5 indicators. Fourteen of the 86 of hospitals had a probability higher than 0.50 of having worse performance than average on at least 6 indicators. No hospitals had a probability higher than 0.50 of having worse performance than average on all 7 indicators.

CONCLUSIONS: These findings suggest that there are a small number of hospitals that perform poorly on at least half of the quality indicators, and that certain indicators tend to cluster together. The described methods allow for targeting quality improvement initiatives at these hospitals.

Key Words: heart failure ■ hospital ■ myocardial infarction ■ quality improvement ■ stroke

https://www.ahajournals.org/journal/circoutcomes

WHAT IS KNOWN

- Between-hospital variation on the performance of different quality of care indicators for patients with stroke has been shown in previous studies.
- Previous studies have examined between-hospital variation in indicators in isolation.

WHAT THE STUDY ADDS

- Bayesian methods for hospital profiling permit the simultaneous assessment of multiple indicators of stroke quality of care.
- There are several pairs of indicators such that hospitals that tended to perform well on one indicator also tended to perform well on the other indicator.
- Several hospitals had a high probability of having worse than average performance on 4 or more stroke quality indicators.

Provider profiling involves reporting on the performance of health care providers on indicators of quality of care. Indicators of health care quality include patient outcomes (eg, death) or measures of processes of care (eg, prescribing of appropriate medication). A health care provider can be an organization or institution such as a hospital or an individual health care worker, such as a surgeon. Examples of hospital report cards include reports on hospital-specific mortality rates for patients undergoing coronary artery bypass graft surgery in different United States and Canadian jurisdictions,[1–5] reports on hospital-specific mortality rates for patients hospitalized with acute myocardial infarction,[6–8] and the reporting on process of care measures in The EFFECT (Enhanced Feedback for Effective Cardiac Treatment) Study of patients hospitalized with acute myocardial infarction or heart failure.[9,10] The HospitalCompare website produced by www.Medicare.gov reports on hospital-specific risk-adjusted 30-day mortality rates for patients hospitalized with acute myocardial infarction, heart failure, and pneumonia and for those undergoing coronary artery bypass graft surgery.[11] Several reports are focused on quality of care for patients hospitalized with stroke. The American Heart Association's Get With The Guidelines program reports on hospital adherence to guidelines and on patient outcomes for patients hospitalized with stroke.[12] The United Kingdom stroke audit reports on hospital performance on an array of stroke quality of care indicators.[13] Riks-Stroke is a Swedish registry that reports on hospital performance on indicators of quality of care for patients with stroke.[14] The Australian Stroke Clinical Registry reports on hospital performance on the care of patients with stroke.[15] Finally, CorHealth Ontario reports on the quality of stroke care across health administrative regions in the province of Ontario.[16]

Provider profiling permits health care providers, administrators, and researchers to identify those providing quality of care that is significantly above or below average. Quality improvement interventions can be targeted at those providing poor quality care. Similarly, identifying health care providers that provide excellent quality of care permits investigation of the reasons for their excellent performance, so that information on best practice can be disseminated to others.

Provider profiling has historically focused on single indicators of quality in isolation. Many of the report cards described above focused on mortality as the primary indicator of quality. Those reports that considered multiple quality indicators tended to examine each indicator in isolation. Thus, variation in hospital performance for one indicator is examined separately from the examination of variation in hospital performance for the other indicators. Examination of each indicator in isolation precludes a formal examination of whether specific providers perform poorly on multiple indicators or whether providers that perform poorly on one indicator tend to perform poorly on other indicators.

A recently developed statistical method, based on a multivariate response Bayesian random effects logistic regression model, permits simultaneous provider profiling on an array of quality indicators.[17] This approach permits the formal evaluation of within-provider correlation on the performance of multiple quality indicators as well as determining the probability that a specific provider had worse (or better) than average performance on more than 1 quality indicator. The objective of the current study is to apply this recently described method to examine within-hospital variation in performance on multiple indicators of quality of care for patients with acute stroke.

METHODS

The datasets used in these analyses were linked using unique encoded identifiers and analyzed at ICES. While data sharing agreements prohibit ICES from making the data set publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at www.ices.on.ca/DAS. The use of data in this project was authorized under section 45 of Ontario's Personal Health Information Protection Act, which does not require review by a Research Ethics Board.

Data

We created a cohort of patients presenting to hospitals in Ontario, Canada with a diagnosis of acute stroke using methods identical to those used in a recent provincial report card on stroke care.[16] Hospital admissions were identified using the Canadian Institute for Health Information Discharge Abstract Database, while emergency department presentations were identified using the Canadian Institute for Health Information National Ambulatory Care Reporting System database.

Outpatient fillings of prescriptions by those over the age of 65 years were identified using the Ontario Drug Benefit database. Inpatient rehabilitation admission after acute care discharge were identified using National Rehabilitation Reporting System. Dr Austin had full access to all the data in the study and takes responsibility for its integrity and the data analysis.

Adults aged 18 and older presenting to hospital between April 1, 2017 and March 31, 2018 were identified using the Canadian Institute for Health Information Discharge Abstract Database and National Ambulatory Care Reporting System, using codes from the *International Classification of Diseases, 10th Revision, Canadian adaptation*. These codes included: H34.1, I60 (excluding I60.8), I61, I63 (excluding I63.6) and I64. The most responsible or main problem diagnosis was used to identify stroke records for adults aged 18 and older in the Canadian Institute for Health Information Discharge Abstract Database and National Ambulatory Care Reporting System databases, respectively.

Quality Indicators

We considered 7 binary indicators of quality of care for patients with acute stroke: (1) referral to secondary prevention services for patients with ischemic stroke discharged from the emergency department; (2) use of carotid imaging in admitted patients with ischemic stroke; (3) use of acute thrombolytic therapy (tissue-type plasminogen activator [tPA]) in patients with ischemic stroke; (4) treatment in a stroke unit at any time during the hospital admission; (5) discharge to an inpatient rehabilitation after acute care; (6) not admitted to a long-term care (LTC) facility or a complex continuing care (CCC) facility after discharge from acute care; (7) filling a prescription for anticoagulant therapy within 90 days of hospital discharge by patients with ischemic stroke and aged 65 years and older with atrial fibrillation. The first indicator is applicable to patients with ischemic stroke discharged from the emergency department. Indicators (2) to (4) apply to all hospitalized patients with stroke, while indicators (5) to (7) apply to patients with stroke discharged from hospital alive. Finally, the seventh indicator applies to patients over the age of 65 years with atrial fibrillation who were discharged alive from hospital. Indicators (1) to (3) and (7) are applicable to patients with acute ischemic stroke, while the remaining indicators are applicable to patients with acute ischemic or hemorrhagic stroke. The sixth indicator, not admitted to LTC/CCC after discharge from acute care, is usually reported as admission to LTC/CCC. However, this indicator, as typically reported, is a negative indicator, in that one wants to avoid admission to LTC/CCC after discharge. This indicator only applies to patients who did not originate from LTC/CCC. We have restructured the indicator so that it is a positive indicator, to be consistent with all the other indicators. The definitions of each indicator and their construction using administrative health databases have been described previously.[16]

Study Sample

We restricted the sample to those 86 hospitals at which at least one patient was eligible for each of the 7 indicators (the same patient did not have to be eligible for all 7 indicators). For each indicator, the minimum hospital-specific number of patients eligible for that indicator was <6 (numbers <6 must be suppressed because of ICES privacy policy). The median, 25th percentile, 75th percentile, and maximum number of patients eligible for each indicator across the 86 hospitals are reported in Table 1, along with the total number of patients eligible for each indicator.

Statistical Methods

We used a multivariate response Bayesian random effects logistic regression model.[17] This method fits a separate random effects logistic regression model for each of the 7 binary indicators. Each model incorporates hospital-specific random effects to account for within-hospital homogeneity in the performance on the given indicator. However, the random effects for the 7 logistic regression models are drawn from a multivariate normal distribution. Thus, the hospital-specific random effects for the different indicators can be correlated with one another. As each indicator is applicable for each eligible patient, none of the 7 regression models incorporated patient characteristics as there was no need for risk-adjustment. This is consistent with the approach taken by HospitalCompare program of www.Medicare.gov, in which outcomes such as mortality or readmission are subject to risk-adjustment, whereas process measures such as time to transfer to a specialized hospital and time to administration of fibrinolytic drugs are not subject to risk-adjustment.[11]

Let $Y_{ij}^{(k)}$ denote the kth binary indicator measured on the ith subject in the jth provider ($k = 1,\ldots,7$). $Y_{ij}^{(k)} = 1$ denotes success on the kth indicator for the ith patient in the jth provider. For each of the 7 binary indicators a random effects logistic regression model was fit:

$$\mathrm{logit}(\Pr(Y_{ij}^{(k)} = 1)) = \mathrm{logit}(p_{ij}^{(k)}) = \alpha_{0j}^{(k)} \quad (1)$$

Note that for the kth indicator, the intercept varies across providers. A multivariate normal distribution is then assumed for the distribution of the provider-specific intercepts for the 7 regression models:

Table 1. Number of Patients Eligible for Each Indicator at Each Hospital

Indicator	Median (25th–75th percentiles) across hospitals	Maximum across hospitals	Total number of patients
Carotid imaging	83 (26–188)	609	11 801
Anticoagulant use in those with atrial fibrillation	14 (4–35)	137	2200
tPA	90 (34–210)	662	13 393
Stroke unit care	87.5 (31–215)	794	14 079
Secondary prevention services	10.5 (4–21)	79	1321
Inpatient rehabilitation	88.5 (27–193)	693	12 688
No LTC/CCC admission	79 (25–175)	663	11 755

CCC indicates complex continuing care; LTC, long-term care; and tPA, tissue-type plasminogen activator.

$$\begin{pmatrix} \alpha_{0j}^{(1)} \\ \alpha_{0j}^{(2)} \\ \vdots \\ \alpha_{0j}^{(7)} \end{pmatrix} \sim \mathrm{MVN}\left(\mu = \begin{pmatrix} \alpha_{0}^{(1)} \\ \alpha_{0}^{(2)} \\ \vdots \\ \alpha_{0}^{(7)} \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{17} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{27} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{71} & \sigma_{72} & \cdots & \sigma_{77}^2 \end{pmatrix} \right) \quad (2)$$

The vector μ that parametrizes the multivariate normal distribution is the mean of the multivariate normal distribution, which has 7 components. The matrix Σ that parametrizes the multivariate normal distribution is the variance-covariance matrix and has 49 components (some of which are redundant since the matrix is symmetrical).

For a given indicator, the parameter $\alpha_0^{(k)}$ denotes the log-odds of the indicator being present (ie, of a successful patient outcome) at an average hospital. For a given indicator, hospitals whose random effects are greater than $\alpha_0^{(k)}$ (ie, $\alpha_{0j}^{(k)} > \alpha_0^{(k)}$) are hospitals at which the odds of the indicator being present are higher than at an average hospital, while hospitals whose random effects are less than $\alpha_0^{(k)}$ (ie, $\alpha_{0j}^{(k)} < \alpha_0^{(k)}$) are hospitals at which the odds of the indicator being present are lower than at an average hospital. A hospital for which $\alpha_{0j}^{(k)} < \alpha_0^{(k)}, k = 1,\ldots,7$ is a hospital with poorer performance than average on all 7 indicators.

Markov Chain Monte Carlo methods were used to estimate the posterior distribution of the model parameters.[18] The prior distribution for μ, the mean of the multivariate normal distribution of the random effects, was specified to be a multivariate normal distribution with mean zero and a variance-covariance matrix equal to $100\times I_{7\times7}$, where I denotes the 7×7 identity matrix. The prior distribution for Σ^{-1}, the precision matrix of the multivariate normal distribution of the random effects, was specified to be the Wishart distribution $W_7\left(\frac{1}{7}I_{7\times7}, 7\right)$. Three chains were run, each using different initial values for the model parameters. Each chain used an initial run of 2 500 000 burn-in iterations and was then monitored for an additional 2 500 000 iterations, with a thinning interval of 500 (ie, 5000 monitored iterations were retained from each of the 3 chains). Thus, a total of 15 000 monitored iterations were used to determine the posterior distributions of the parameters of interest. The Gibbs sampler was implemented using OpenBUGS version 3.2.3 using the R2OpenBUGS package for R.

Convergence of the Markov Chain Monte Carlo Process

A total of 637 model parameters were monitored: 7 parameters for the mean of the multivariate distribution of the random effects, 28 parameters for the precision matrix (inverse of the symmetrical variance-covariance matrix) for the multivariate distribution of the random effects, and 602 hospital-specific random effects for the 6 indicators (86 hospitals×7 indicators). Convergence of the Gibbs sampler was assessed by visual inspection of the trace plots for 70 parameters (the 7 components of the mean of the multivariate normal distribution, the 28 components of the variance-covariance matrix, and the 35 random effect parameters for the first 5 hospitals). The 3 separate chains starting at different starting values mixed well and displayed no lack of convergence. The convergence of each chain was also assessed using Geweke's statistic,[19] by which we tested the equality of the means of the sampled parameters in the first 25% of the chain with that in the last 25% of the chain. If the sampled values of a given parameter are drawn from the same stationary distribution, then the 2 means are equal, and the resultant test statistic will have a standard normal distribution. For each of the 3 chains, there was no evidence that the distribution of Geweke's test statistic was not normal across the 637 model parameters when using visual inspection of a normal quantile-quantile plot.

The statistical analyses were conducted using R (version 3.5.1) and OpenBUGS (version 3.2.3).

Sensitivity Analysis

Two of the 7 indicators (anticoagulant use in those with atrial fibrillation and referral to secondary prevention services) were applicable to fewer patients than the other 5 indicators. We repeated the above analyses using the same sample, but only considering the other 5 indicators.

Table 2. Within-Hospital Correlation in Performance on Pairwise Combinations of 7 Indicators

	Carotid imaging	Anticoagulant	tPA	Stroke unit	Secondary prevention services	Inpatient rehab	No LTC/CCC
Carotid imaging	1	−0.25*	0.84†	0.88†	0.39‡	0.59†	0.42‡
Anticoagulant	−0.25*	1	−0.23*	−0.33‡	−0.26*	−0.09	0
tPA	0.84†	−0.23*	1	0.9†	0.28*	0.57†	0.40‡
Stroke unit	0.88†	−0.33‡	0.9†	1	0.51†	0.55†	0.36‡
Secondary prevention services	0.39‡	−0.26*	0.28*	0.51†	1	0.13*	0
Inpatient rehab	0.59†	−0.09	0.57†	0.55†	0.13*	1	0.40‡
No LTC/CCC	0.42‡	0	0.40‡	0.36‡	0	0.40‡	1

CCC indicates complex continuing care; LTC, long-term care; and tPA, tissue-type plasminogen activator.
*A weak correlation.
†A strong correlation.
‡A moderate correlation.

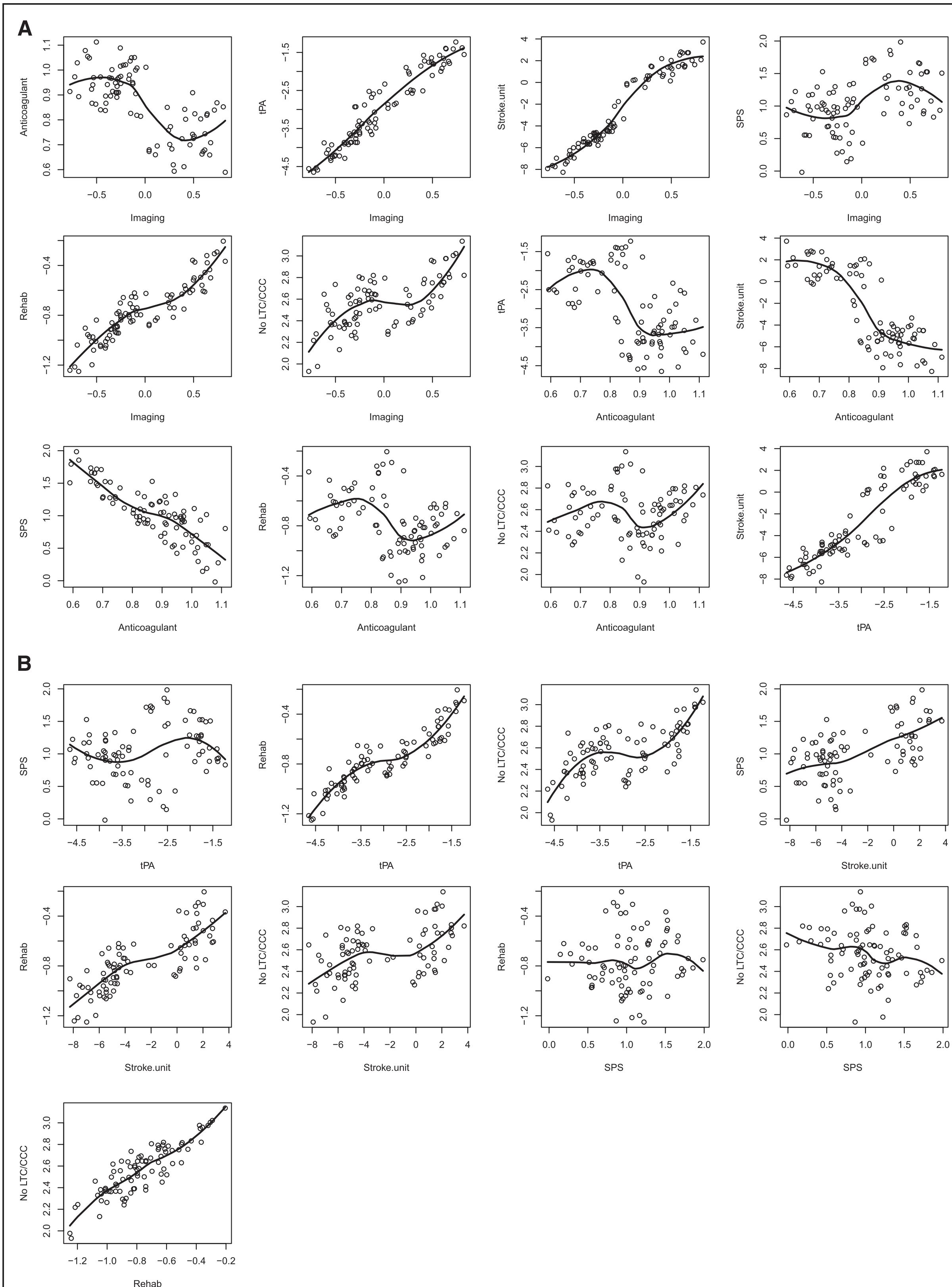

Figure 1. Correlation between hospital-specific random effects for the 7 indicators.
CCC indicates complex continuing care; LTC, long-term care; and tPA, tissue-type plasminogen activator**.**

RESULTS

Summary Statistics on the Quality Indicators

The overall probability of a successful outcome for each of the 7 indicators across all hospitals was 85% (carotid imaging), 72% (anticoagulant use for atrial fibrillation), 13% (tPA), 55% (stroke unit admission), 74% (referral to secondary prevention clinic after emergency department discharge), 33% (admission to in-patient rehabilitation), and 92% (no discharge to LTC/CCC). The hospital-specific prevalences of the indicators ranged from 0% (for all indicators other than no admission to LTC/CCC) to 100% (carotid imaging, anticoagulant use and referral to secondary prevention clinic), 27% (tPA), 98% (stroke unit admission), 67% (admission to in-patient rehabilitation). For no admission to LCT/CCC, the hospital-specific prevalences ranged from 50% to 100%.

The posterior means of the variances of the hospital-specific random effects were 0.043 (carotid imaging), 0.018 (anticoagulants), 0.177 (tPA), 2.191 (stroke unit), 0.041 (secondary prevention clinic), 0.022 (inpatient rehabilitation), and 0.025 (no LTC/CCC admission). These are equivalent to variance partition coefficients of 0.013, 0.005, 0.051, 0.400, 0.012, 0.007, and 0.008, respectively (using the latent variable formulation of the variance partition coefficient).[20–22] Thus, 1.3% of the variation in use of carotid imaging is due to systematic differences between hospitals, while 40% of the variation in stroke unit admission was due to systematic differences between hospitals. Most of the indicators displayed only minor between-hospital variation, one displayed moderate between-hospital variation (tPA), and one displayed strong between-hospital variation (stroke unit admission).

Correlation of Hospital-Specific Random Effects for the 7 Indicators

We determined the posterior mean of the precision matrix for the multivariate distribution of the hospital-specific random effects. This matrix was inverted to obtain the variance-covariance matrix of the distribution

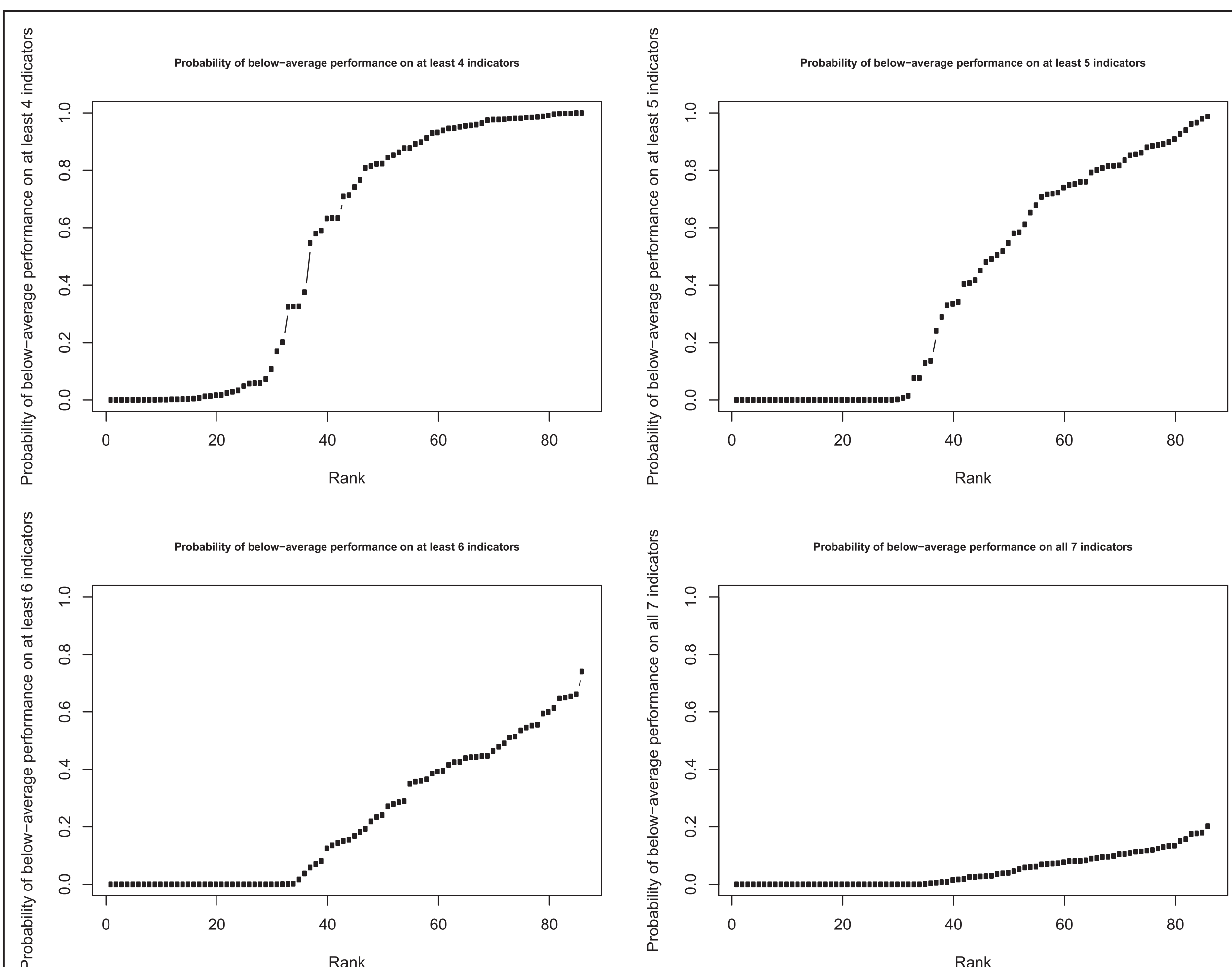

Figure 2. Probability of below-average performance on at least K indicators.

of the hospital-specific random effects. From this matrix we obtained the correlation matrix for the hospital-specific random effects. The correlation matrix of the hospital-specific random effects is reported in Table 2. In interpreting the magnitude of specific correlations, we used the following criteria, which are based on Cohen's discussion of effect sizes: $0.1 < \rho \leq 0.3$ denotes weak correlation; $0.3 < \rho \leq 0.5$ denotes moderate correlation; $\rho > 0.5$ denotes strong correlation.[23,24] The plots of the pair-wise relationships between the hospital-specific posterior means of the different random effects are presented in Figure 1A and 1B. There is one panel for each of the 21 pairwise comparisons. On each panel, we have superimposed a smooth curve, estimating using a loess regression model, describing the relationship between hospital-specific random effects for the 2 indicators.

There were strong positive correlations between a hospital's use of carotid imaging and: use of tPA, treatment in a stroke unit, and discharge to inpatient rehabilitation. Thus, a hospital that had better-than-average performance on use of carotid imaging tended to have better-than-average performance on use of tPA. Conversely, a hospital that worse-than-average performance on carotid imaging tended to have worse-than-average performance on use of tPA. There were strong positive correlations between a hospital's use of tPA and: treatment in a stroke unit and discharge to inpatient rehabilitation. Finally, there was a strong positive correlation between a hospital's use of stroke units and: referral to secondary prevention clinics and discharge to inpatient rehabilitation.

Probability of Having Below-Average Performance on Multiple Indicators Simultaneously

For each hospital, we determined the posterior probability that the hospital had worse performance than an average hospital on at least 4 indicators, on at least 5 indicators, on at least 6 indicators, and on all 7 indicators. Figure 2 depicts a snake plot in which the posterior probability of below-average performance on at least the given number of indicators is plotted against

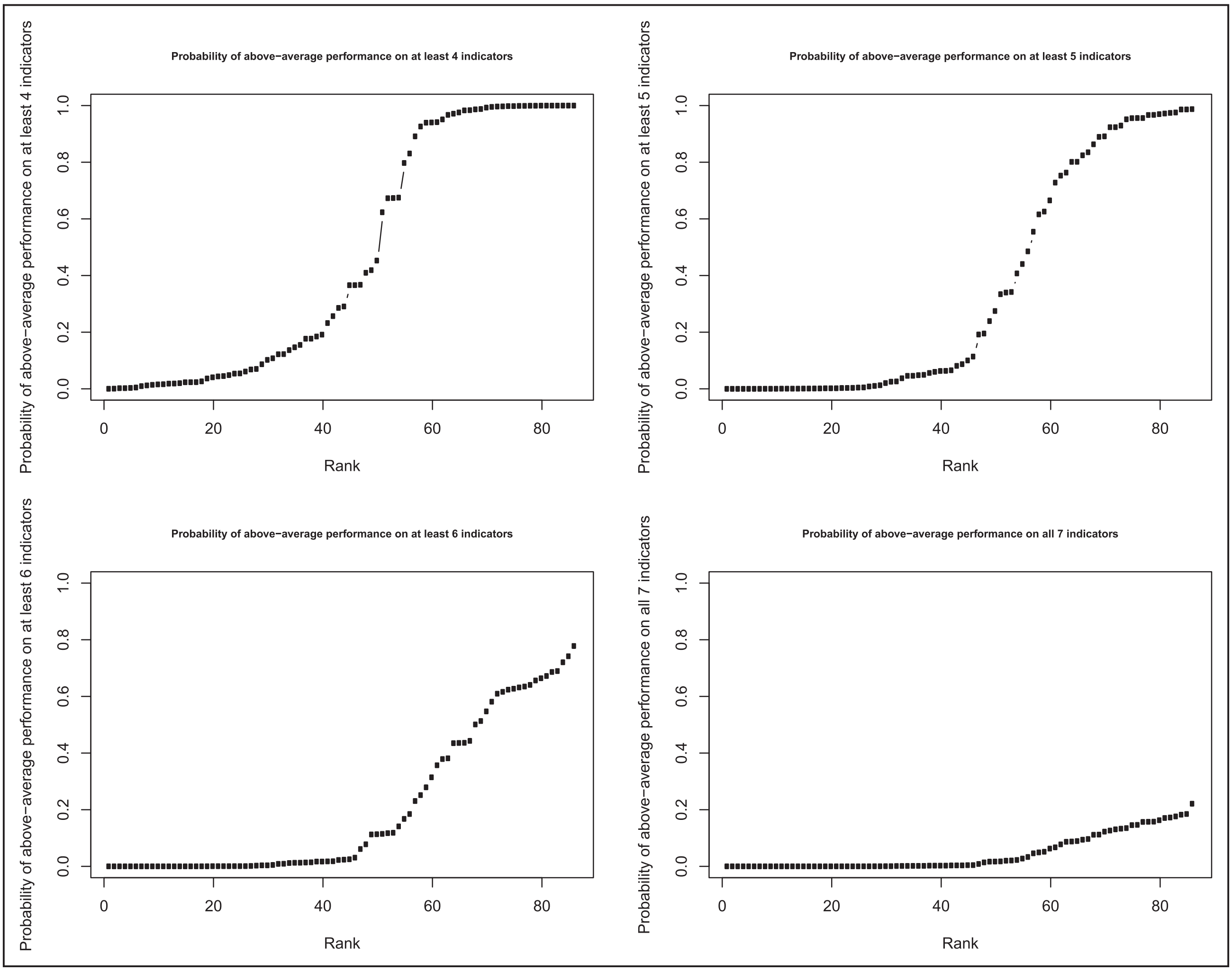

Figure 3. Probability of above-average performance on at least K indicators.

the hospital's rank on these probabilities. Twenty-nine of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on at least 4 indicators. Seven of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on at least 5 indicators. Fourteen of the 86 of hospitals had a probability higher than 0.50 of having worse performance than average on at least 6 indicators. No hospitals had a probability higher than 0.50 of having worse performance than average on all 7 indicators.

Probability of Having Above-Average Performance on Multiple Indicators Simultaneously

We repeated the analyses described in the previous section but examined the probability of having above-average performance on multiple indicators simultaneously. Results are reported in Figure 3, which is similar in structure to Figure 2. Twenty-nine of the 86 of hospitals had a probability higher than 0.90 of having better performance than average on at least 4 indicators. Sixteen of the 86 of hospitals had a probability higher than 0.90 of having better performance than average on at least 5 indicators. Nineteen of the 86 of hospitals had a probability higher than 0.50 of having better performance than average on at least 6 indicators. No hospitals had a probability higher than 0.50 of having better performance than average on all 7 indicators.

Sensitivity Analysis

We repeated the above analyses after excluding 2 indicators that were applicable to fewer patients (anticoagulant use in those with atrial fibrillation and referral to secondary prevention services). The correlation matrix of the hospital-specific random effects is reported in Table 3. Eight of the pairwise correlations were strong, while the remaining 2 were moderate.

Results are reported graphically in Figures 4 through 6. The plots of the pairwise relationships between the hospital-specific posterior means of the different random effects are presented in Figure 4. There is 1 panel for each of the 10 pairwise comparisons. On each panel, we have superimposed a smooth curve, estimating using a loess regression model, describing the relationship between hospital-specific random effects for the 2 indicators.

Thirty-five of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on at least 3 of the 5 indicators. Eleven of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on at least 4 of the 5 indicators. One of the 86 of hospitals had a probability higher than 0.90 of having worse performance than average on all 5 indicators.

Thirty-two of the 86 of hospitals had a probability higher than 0.90 of having better performance than average on at least 3 of the 5 indicators. Eighteen of the 86 of hospitals had a probability higher than 0.90 of having better performance than average on at least 4 of the 5 indicators. Two of the 86 of hospitals had a probability higher than 0.90 of having better performance than average on all 5 indicators.

Table 3. Within-Hospital Correlation in Performance on Pairwise Combinations of 5 Indicators

	Carotid imaging	tPA	Stroke unit	Inpatient rehab	No LTC/CCC
Carotid Imaging	1	0.88*	0.92*	0.68*	0.51*
tPA	0.88*	1	0.95*	0.58*	0.38†
Stroke unit	0.92*	0.95*	1	0.62*	0.43†
Inpatient rehab	0.68*	0.58*	0.62*	1	0.51*
No LTC/CCC	0.51*	0.38†	0.43†	0.51*	1

CCC indicates complex continuing care; LTC, long-term care; and tPA, tissue-type plasminogen activator.

*Strong correlation.

†Moderate correlation.

DISCUSSION

The objective of this study was to examine within-hospital correlation in performance on a set of 7 binary indicators of quality of care for patients with acute stroke. We found that there were several pairs of quality indicators for which hospitals that performed well on one indicator tended to perform well on the other indicator. Furthermore, we found that several hospitals had a high probability of having worse than average performance on 4 or more quality indicators. Conversely, several hospitals had a high probability of having better than average performance on 4 or more quality indicators.

We found that hospitals that performed well on 1 indicator tended to perform well on other indicators. There are at least 2 possible reasons for this phenomenon. First, that the indicators themselves are correlated or represent the same underlying construct. Second, that there is an underlying construct of hospital quality, and that high-quality hospitals perform well on multiple indicators because they are high-quality hospitals. We would argue that the first explanation is the less likely of the 2 in our context, as the 7 indicators represent diverse aspects of quality of care, with some denoting use of resources internal to the hospital and some denoting use of resources external to the hospital.

The simultaneous assessment of multiple indicators provides information on how quality measures cluster

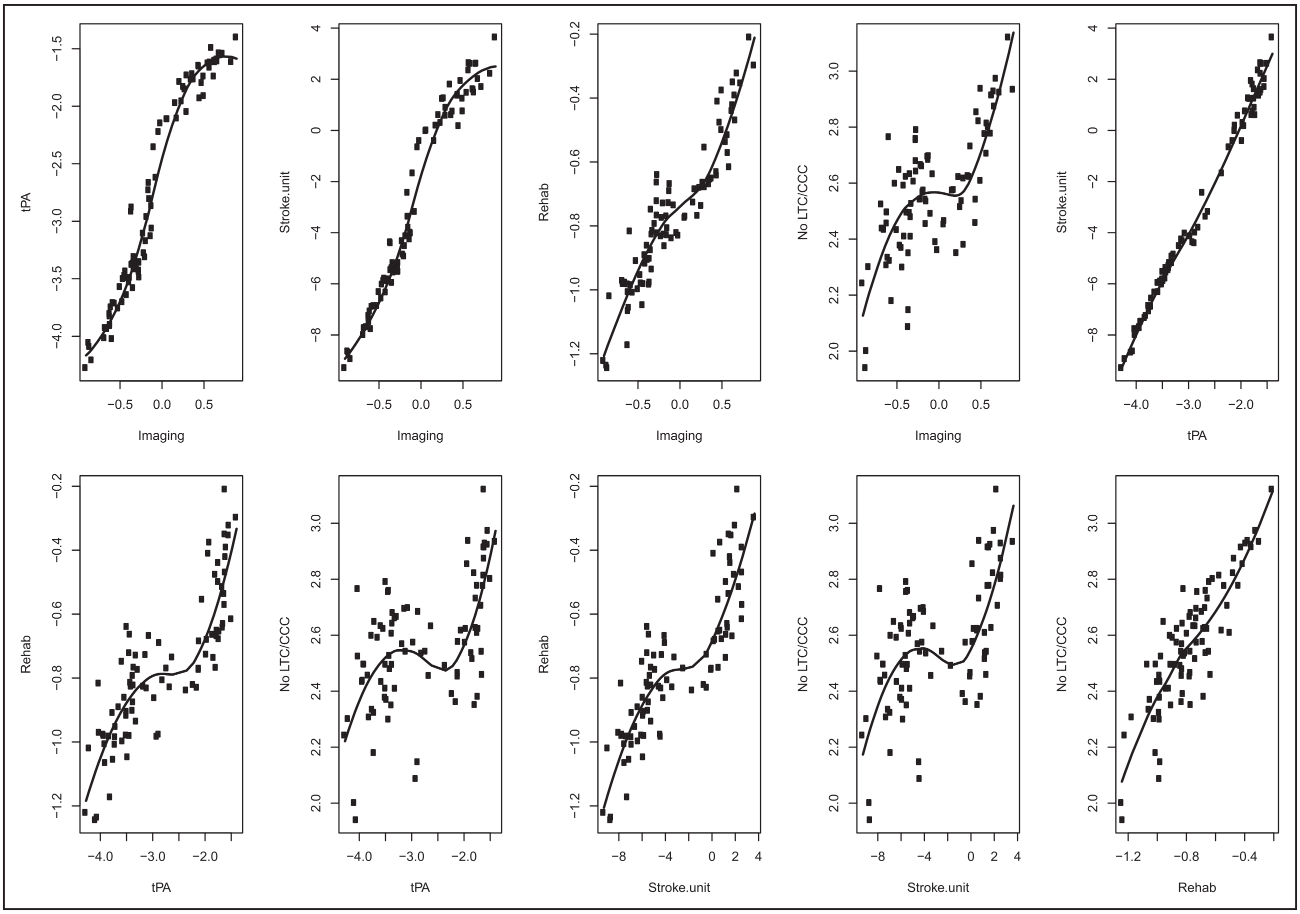

Figure 4. Correlation between hospital-specific random effects for the 5 indicators.
CCC indicates complex continuing care; LTC, long-term care; and tPA, tissue-type plasminogen activator.

together and may be useful for understanding patterns at high- and low-performing hospitals. There were 4 indicators (carotid imaging, use of tPA, treatment in a stroke unit, and discharge to in-patient rehabilitation) for which hospitals had a strong correlation on their performance on each of the 6 possible pairwise comparisons. Thus, hospitals that performed well on any 1 of these 4 indicators tended to perform well on the other 3 indicators. Similarly, hospitals that performed poorly on any 1 of these 4 indicators tended to perform poorly on the other 3 indicators. Performing well on these indicators requires access to specific resources such as neuroimaging, stroke units, and rehabilitation facilities. Thus, to a certain extent, performing well on multiple indicators simultaneously can serve as a marker of access to adequate resources for delivering comprehensive stroke care. These findings suggest that quality may be best envisaged as an overall construct of a hospital rather than being process-specific. Hospitals that strive to improve the quality of stroke care across multiple domains may wish to focus on a comprehensive indicator such as stroke unit care, anticipating that the performance on other indicators will improve in parallel with this.

Christiansen and Morris[25] suggest that an advantage to the use of Bayesian methods for profiling compared with conventional Frequentist approaches is that they allow for profiling to be based on medically based criteria for assessing provider profiling. A similar argument was made by Normand et al.[26] In particular, Bayesian methods allow for computation of the probability of acceptable provider performance. This contrasts with conventional Frequentist methods that typically focus on testing the null hypothesis of whether the performance of a given provider differed from that of an average provider. Christiansen and Morris[25] suggest that "this hypothesis is not very useful: Taken literally, it means that if the true hospital mortality rates differ even by tiny amounts (which one would expect), many of the hospitals would have true rates that exceed the population mean". Our use of Bayesian methods allowed us to compute the probability that a hospital had worse than average performance on a given number of quality indicators. This is a quantity that cannot be computed using conventional frequentist statistics. The computation of this probability is important as it allows one to assess the strength of the evidence that a hospital has poor performance on one or more indica-

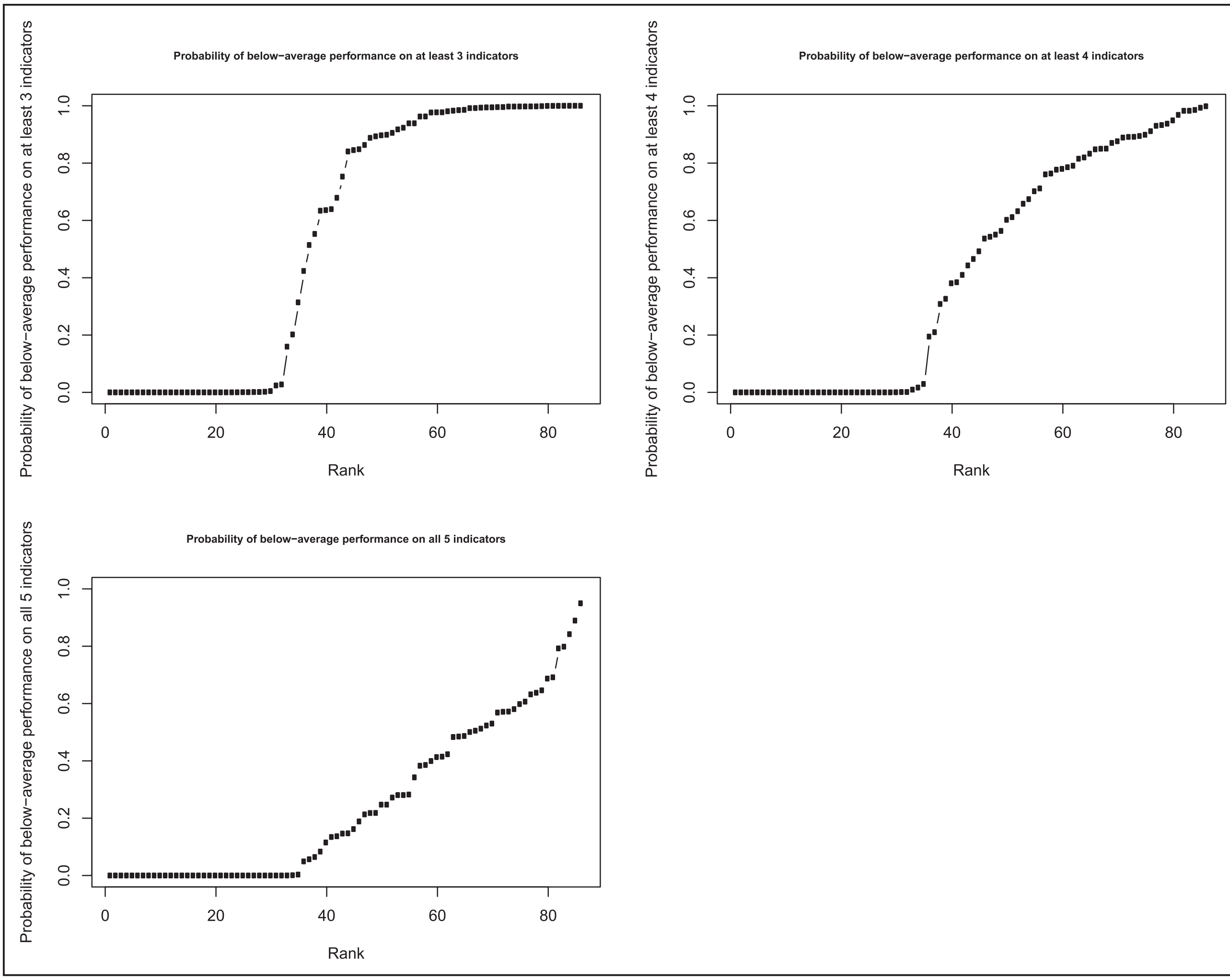

Figure 5. Probability of below-average performance on at least K indicators (sensitivity analysis).

tors. While Bayesian analyses can be more complex to implement than conventional Frequentist analyses, they allow analysts to answer clinically meaningful and policy-relevant questions about hospital performance that cannot be addressed using conventional Frequentist methods. Furthermore, models that incorporate multiple indicators, such as those described in this article and elsewhere are much easier to implement from a Bayesian perspective than from a Frequentist approach.[17,27]

Several authors have described the use Bayesian methods for hospital profiling. Normand et al[26] described the use of Bayesian hierarchical models for provider profiling. They developed methods to estimate the probability that adjusted mortality rate at a specific hospital exceeded specified thresholds. Racz and Sedransk[28] explored the use of Bayesian methods for provider profiling using data from the New York State coronary artery bypass graft report cards. Berlowitz et al[29] examined the use of Bayesian hierarchical models for profiling nursing homes on their rates of pressure ulcers. Both Staggs and Gajewski[30] and Gajewski et al[31] used Bayesian models to profile hospital nursing units. Finally, a series of methodological articles by the first author explored different issues in the use of Bayesian methods for provider profiling.[32–35]

There are certain limitations to the current study. The principal limitation is the absence of eligible denominators for some indicators. Stroke severity and other baseline characteristics will influence eligibility for rehabilitation, thrombolysis, and other quality indicators. However, information on stroke severity is not available from the data sources that were used. Furthermore, performance on some indicators, such as discharge to in-patient rehabilitation, may be influenced by factors outside of the hospitals' control.

In conclusion, we found that using Bayesian methods for hospital profiling permitted the simultaneous assessment of multiple indicators of stroke quality of care and identified groups of indicators for which hospitals tended to have a consistent performance. We anticipate that this methodology will serve as a useful tool for jurisdictions that seek to identify hospitals with high and low performance on groups of indicators to identify targets for quality improvement activities, and

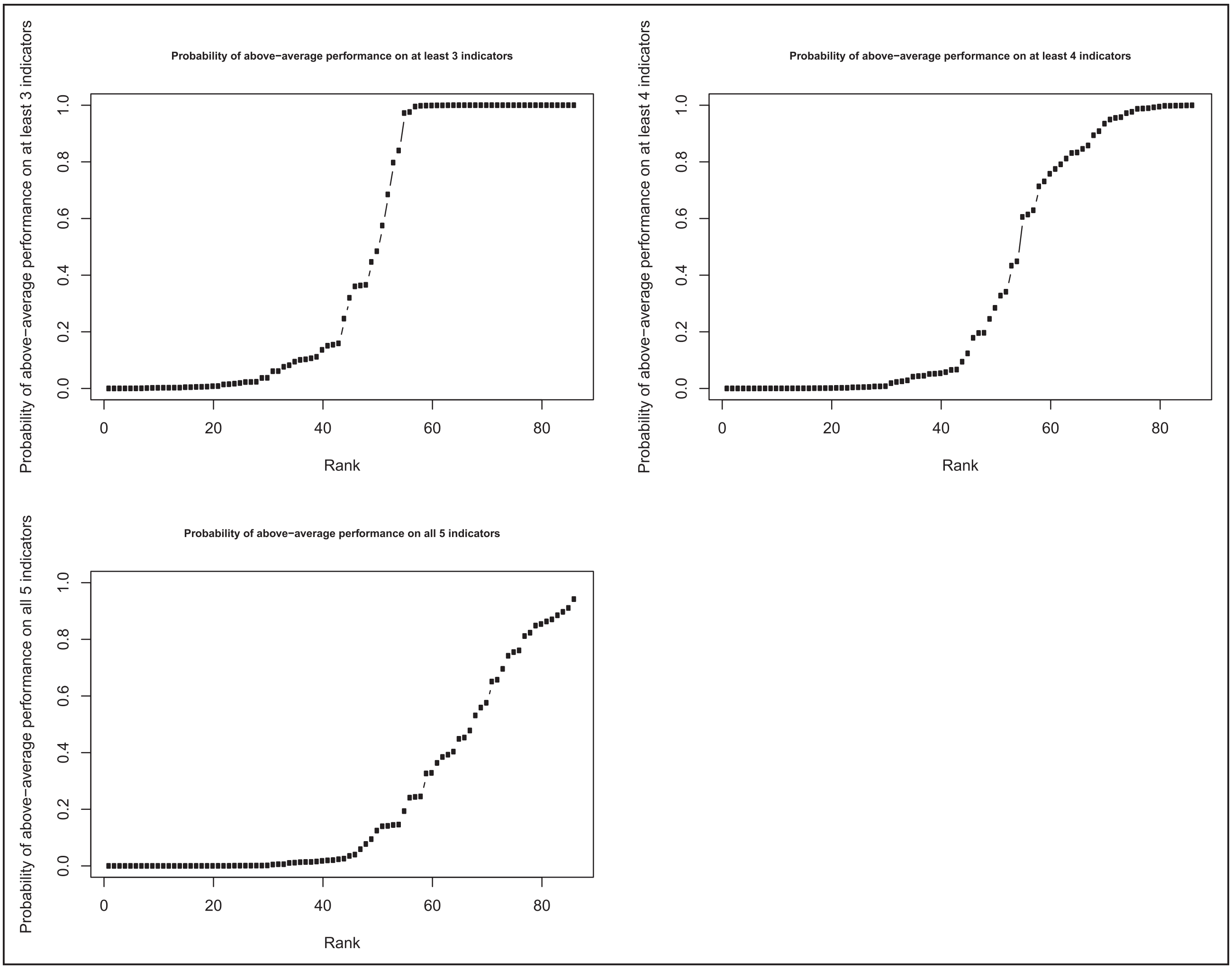

Figure 6. Probability of above-average performance on at least K indicators (sensitivity analysis).

that the identification of clustered quality measures will allow hospitals to focus on indicators, such as stroke unit care, which may be most likely to provide improvements in quality of care across multiple domains.

ARTICLE INFORMATION

Received May 11, 2020; accepted October 1, 2020.

Correspondence

Peter C. Austin, PhD, ICES, G106, 2075 Bayview Ave, Toronto, ON M4N 3M5, Canada. Email peter.austin@ices.on.ca

Affiliations

ICES, Toronto, ON, Canada (P.C.A., J.F., B.Y., M.K.K.). Institute of Health Policy, Management and Evaluation (P.C.A., M.K.K.) and Department of Medicine (M.K.K.), University of Toronto, ON, Canada. Schulich Heart Research Program, Sunnybrook Research Institute, Toronto, ON, Canada (P.C.A.).

Sources of Funding

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). The opinions, results, and conclusions reported in this article are those of the authors and are independent from the funding sources. No endorsement by ICES or the Ontario MOHLTC is intended or should be inferred. This research was supported by an operating grant from the Canadian Institutes of Health Research (CIHR) (MOP 86508). Drs Austin and Kapral are supported by Mid-Career Investigator awards from the Heart and Stroke Foundation. Parts of this material are based on data and/or information compiled and provided by CIHI. However, the analyses, conclusions, opinions, and statements expressed in the material are those of the author(s) and not necessarily those of Canadian Institute for Health Information. We thank IMS Brogan, Inc for use of their Drug Information Database. These datasets were linked using unique encoded identifiers and analyzed at ICES. While data sharing agreements prohibit ICES from making the data set publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at www.ices.on.ca/DAS.

Disclosures

None.

REFERENCES

1. *Coronary artery Bypass Graft Surgery in New York State 1989-1991*. New York State Department of Health; 1992.
2. Jacobs FM. *Cardiac Surgery in New Jersey in 2002: A Consumer Report*. Department of Health and Senior Services; 2005.
3. *Adult Coronary Artery Bypass Graft Surgery in the Commonwealth of Massachusetts: Fiscal Year 2010 Report*. Massachusetts Data Analysis Centre, Department of Health Care Policy, Harvard Medical School; 2012.
4. Naylor CD, Rothwell DM, Tu JV, Austin PC; the Cardiac Care Network Steering C. Outcomes of coronary artery bypass surgery in Ontario. In: Naylor

CD, Slaughter PM, eds. *Cardiovascular Health and Services in Ontario: An ICES Atlas*. Institute for Clinical Evaluative Sciences; 1999:189–198.
5. *Consumer Guide to Coronary Artery Bypass Graft Surgery*. Pennsylvania Health Care Cost Containment Council; 1995. Volume 4.
6. Luft HS, Romano PS, Remy LL, Rainwater J. *Annual Report of the California Hospital Outcomes Project*. California Office of Statewide Health Planning and Development; 1993.
7. *Focus on Heart Attack in Pennsylvania: Research Methods and Results*. Pennsylvania Health Care Cost Containment Council; 1996.
8. Tu JV, Austin PC, Naylor CD, Iron K, Zhang H. Acute myocardial infarction outcomes in Ontario. In: Naylor CD, Slaughter PM, eds. *Cardiovascular Health and Services in Ontario: An ICES Atlas*. Institute for Clinical Evaluative Sciences; 1999:83–110.
9. Tu JV, Donovan LR, Lee DS, Wang JT, Austin PC, Alter DA, Ko DT. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. *JAMA*. 2009;302:2330–2337. doi: 10.1001/jama.2009.1731
10. Tu JV, Donovan LR, Lee DS, Austin PC, Ko DT, Wang JT, Newman AM. *Quality of Cardiac Care in Ontario*. Institute for Clinical Evaluative Sciences; 2004.
11. Medicare.gov. *Hospital Compare*. https://www.medicare.gov/HospitalCompare/Data/Death-rates.html. Accessed October 1, 2020.
12. Ormseth CH, Sheth KN, Saver JL, Fonarow GC, Schwamm LH. The American Heart Association's get with the guidelines (GWTG)-stroke development and impact on stroke care. *Stroke Vasc Neurol*. 2017;2:94–105. doi: 10.1136/svn-2017-000092
13. King's College London. *Sentinel Stroke National Audit Programme*. https://www.strokeaudit.org/About-SSNAP/SSNAP-Acute-Organisational-Audit.aspx. Accessed October 1, 2020.
14. Asplund K, Hulter Asberg K, Norrving B, Stegmayr B, Terént A, Wester PO; Riks-Stroke Collaboration. Riks-stroke - a Swedish national quality register for stroke care. *Cerebrovasc Dis*. 2003;15(suppl 1):5–7. doi: 10.1159/000068203
15. Breen S, Cadilhac DA, Lannin NA, Kim J, Dalli L, Anderson CS, Kilkenny M, Shehata S, Faux S, Dewey H, et al. *The Australian Stroke Clinical Registry Annual Report*. The Florey Institute of Neuroscience and Mental Health; 2018.
16. Pardhan A, Sharp S, Yearwood K, Britton K, Linkewich B, Bayley MT, Brown G, Fang J, Fearn J, French E, et al. *Ontario and Local Health Integration Network (LHIN) 2017/18 Stroke Report Cards and Progress Reports*. Corhealth; 2019.
17. Austin PC, Lee DS, Leckie G. Comparing a multivariate response Bayesian random effects logistic regression model with a latent variable item response theory model for provider profiling on multiple binary indicators simultaneously. *Stat Med*. 2020;39:1390–1406. doi: 10.1002/sim.8484
18. Gilks WR, Richardson S, Spiegelhalter D. *Markov Chain Monte Carlo in Practice*. Chapman & Hall; 1996.
19. Geweke J. Evaluating the accuracy of sampling based approaches to the calculation of posterior moments. In: Bernardo JM, Berger JO, Dawid AP, Smith AFM, eds. *Bayesian Statistics 4*. Oxford University Press; 1992:169–193.
20. Snijders T, Bosker R. *Multilevel Analysis: An introduction to Basic and Advanced Multilevel Modeling*. Sage Publications; 2012.
21. Goldstein H, Browne W, Rasbash J. Partitioning variation in generalised linear multilevel models. *Understanding Statistics*. 2002;1:223–232.
22. Austin PC, Merlo J. Intermediate and advanced topics in multilevel logistic regression analysis. *Stat Med*. 2017;36:3257–3277. doi: 10.1002/sim.7336
23. Cohen J. *Statistical Power Analysis for the Behavioural Sciences*. Lawrence Erlbaum Associates; 1988.
24. Sheskin DJ. *Handbook of Parametric and Nonparametric Statistical Procedures*. Chapman & Hall/CRC; 2004.
25. Christiansen CL, Morris CN. Improving the statistical approach to health care provider profiling. *Ann Intern Med*. 1997;127(8 pt 2):764–768. doi: 10.7326/0003-4819-127-8_part_2-199710151-00065
26. Normand SLT, Glickman ME, Gatsonis CA. Statistical mertods for profiling providers of medical care: issues and applications. *J Am Stat Assoc*. 1997;92:803–814.
27. Teixeira-Pinto A, Normand SL. Statistical methodology for classifying units on the basis of multiple-related measures. *Stat Med*. 2008;27:1329–1350. doi: 10.1002/sim.3187
28. Racz MJ, Sedransk J. Bayesian and frequentist methods for provider profiling using risk-adjusted assessments of medical outcomes. *J Am Stat Assoc*. 2010;105:48–58.
29. Berlowitz DR, Christiansen CL, Brandeis GH, Ash AS, Kader B, Morris JN, Moskowitz MA. Profiling nursing homes using Bayesian hierarchical modeling. *J Am Geriatr Soc*. 2002;50:1126–1130. doi: 10.1046/j.1532-5415.2002.50272.x
30. Staggs VS, Gajewski BJ. Bayesian and frequentist approaches to assessing reliability and precision of health-care provider quality measures. *Stat Methods Med Res*. 2017;26:1341–1349. doi: 10.1177/0962280215577410
31. Gajewski BJ, Mahnken JD, Dunton N. Improving quality indicator report cards through Bayesian modeling. *BMC Med Res Methodol*. 2008;8:77. doi: 10.1186/1471-2288-8-77
32. Austin PC. Bayes rules for optimally using Bayesian hierarchical regression models in provider profiling to identify high-mortality hospitals. *BMC Med Res Methodol*. 2008;8:30. doi: 10.1186/1471-2288-8-30
33. Austin PC. A comparison of Bayesian methods for profiling hospital performance. *Med Decis Making*. 2002;22:163–172. doi: 10.1177/0272989X0202200213
34. Austin PC. Are (the log-odds of) hospital mortality rates normally distributed? Implications for studying variations in outcomes of medical care. *J Eval Clin Pract*. 2009;15:514–523. doi: 10.1111/j.1365-2753.2008.01053.x
35. Austin PC. The reliability and validity of bayesian methods for hospital profiling: a Monte Carlo assessment. *J Stat Plan Inference*. 2005;128:109–122.