Appendix A. Supplemental Tables and Figures

Supplemental Tables and Figures Legends

Supplemental Table 1. Model performance c-statistics.

Supplemental Table 2. Comparison of hospital quintile rankings based on direct standardized rates with and without shrinkage.

Supplemental Table 3. Hospital size in top 10% and bottom 10% ranking from each method.

Supplemental Figure 1. Observed rates for mortality and, major morbidity or mortality in isolated coronary artery bypass graft. The line in the middle of each box is the median. The box represents the middle 50% of the data. The box edges are the 25th and 75th percentiles. The circle inside the box represents the mean. The circle outside the box represents the outliers. The colors represent the hospital size.

Supplemental Figure 2. Hospital ranking from different methods. Each bubble represents a hospital, and the size of the bubble indicates the hospital volume. The rates are for hospital mortality. Y-axis and x-axis are position ranks from each method. The 45-degree line is the reference line for hospitals without changing ranks. (A-C) Ranking for operative mortality. A) Ranks from direct standardization rates (Dir_fixed vs. Bayesian rankings, r=0.869); B) Ranks from direct standardization rates (Dir_random vs. Bayesian, r=0.954); C) Ranks from Dir_random vs. Dir_fixed (r=0.939). (D-F) major morbidity or mortality. D) Ranks from direct standardization rates (Dir_fixed vs. Bayesian rankings, r= 0.9728); E) Ranks from direct standardization rates (Dir_random vs. Bayesian, r= 0.998); F) Ranks from Dir_random vs. Dir_fixed (r=0.976).

Supplemental Figure 3**.** Comparison of hospital performance outliers. Indir_logit: indirect standardization with logistic regression models; Indir_fixed: indirect standardization with fixed effect models; Indir_random: indirect standardization with random effect models. S: small hospitals; M: medium hospitals; L: large hospitals. Lower than average: better hospitals; Average: average hospitals; Higher than average: worse hospitals. A) Hospital operative mortality; B) Hospital major morbidity or mortality.

Supplemental Figure 4. The correlation between direct standardized rates with shrinkage Dir_random vs. indirect standardized rates with shrinkage Indir_random is shown using hospital mortality as the outcome. r= 0.998 ($p<0.0001$). Direct standardization rates with shrinkage are consistently with indirect standardized rates with shrinkage.

Supplemental Table 1. Model performance c-statistics.

Outcomes	Model types	Model development data	Ranking data
Mortality or morbidity	Logistic model	0.73	0.73
	Fixed effect model	0.74	0.75
	Random effect model	0.74	0.74
Mortality	Logistic model	0.79	0.81
	Fixed effect model	0.81	0.84
	Random effect model	0.81	0.82

Supplemental Table 2. Comparison of hospital quintile rankings based on direct standardized rates with and without shrinkage.

Supplemental Table 2A. Quintiles ranking changes for hospital mortality (Dir_fixed vs. Dir_random)

Quintiles based on Dir_random method with shrinkage	Quintiles based on Dir_fixed method without shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	13 (72.2%)	5 (27.8%)			
2	4 (22.2%)	11 (61.1%)	3 (16.7%)		
3	1 (5.6%)	2 (11.1%)	13 (72.2%)	2 (11.1%)	
4			2 (11.1%)	13 (72.2%)	3 (16.7%)
5 "High"				3 (16.7%)	15 (83.3%)
Hospitals with changing rankings	5 (27.8%)	7 (38.9%)	5 (27.8%)	5 (27.8%)	3 (16.7%)
small	5	2	0	2	1
medium	0	1	0	0	2
large	0	4	5	3	0

Supplemental Table 2B. Quintiles ranking changes for hospital mortality (Dir_fixed vs. Bayesian)

Quintiles based on Bayesian method with shrinkage	Quintiles based on Dir_fixed method without shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	13 (72.2%)	5 (27.8%)			
2	2 (11.1%)	10 (55.6%)	5 (27.8%)		1 (5.6%)
3	2 (11.1%)	3 (16.7%)	10 (55.6%)	3 (16.7%)	
4	1 (5.6%)		3 (16.7%)	13 (72.2%)	1 (5.6%)
5 "High"				2 (11.1%)	16 *88.9)
Hospitals with changing rankings	5 (27.8%)	8 (44.4%)	8 (44.4%)	5 (27.8%)	2 (11.1%)
small	5	2	5	1	2
medium	0	2	0	1	0
large	0	4	3	3	0

Supplemental Table 2C. Quintiles ranking changes for hospital mortality (Dir_random vs. Bayesian)

Quintiles based on Bayesian method with shrinkage	Quintiles based on Dir_random method with shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	18	0	0		
2		13 (72.2)	4 (22.2)	1 (5.6)	
3		5 (27.8)	9 (50.0)	4 (22.2)	
4			5 (27.8)	10 (55.6)	3 (16.7)
5 "High"				3 (16.7)	15 (83.3)
Hospitals with changing rankings	0	5 (27.8 %)	9 (50.0 %)	8 (44.4%)	3 (16.7 %)
small	0	3	9	2	1
medium	0	1	0	3	0
large	0	1	0	3	2

Supplemental Table 2D. Quintiles ranking changes for hospital major morbidity or mortality (Dir_fixed vs. Dir_random)

Quintiles based on Dir_random method with shrinkage	Quintiles based on Dir_fixed method without shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	15 (83.3%)	3 (16.7%)			
2	3 (16.7%)	13 (72.2%)	2 (11.1%)		
3		2 (11.1%)	15 (83.3%)	1 (5.6%)	
4			1 (5.6%)	16 (88.9%)	1 (5.6%)
5 "High"				1 (5.6%)	17 (94.4%)
Hospitals with changing rankings	3 (16.7%)	5 (27.8%)	3 (16.7%)	2 (11.1%)	1 (5.6%)
small	3	2	0	1	1
medium	0	1	1	0	0
large	0	2	2	1	0

Supplemental Table 2E. Quintiles ranking changes for hospital major morbidity or mortality (Dir_fixed vs. Bayesian)

Quintiles based on Bayesian method with shrinkage	Quintiles based on Dir_fixed method without shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	15 (83.3%)	3 (16.7%)			
2	3 (16.7%)	13 (72.2%)	2 (11.1%)		
3		2 (11.1%)	15 (83.3%)	1 (5.6%)	
4			1 (5.6%)	16 (88.9%)	1 (5.6%)
5 "High"				1 (5.6%)	17 (94.4%)

Hospitals with changing rankings	3 (16.7%)	5 (27.8%)	3 (16.7%)	2 (11.1%)	1 (5.6%)
small	3	2	0	1	0
medium	0	1	1	0	1
large	0	2	2	1	0

Supplemental Table 2F. Quintiles ranking changes for hospital major morbidity or mortality (Dir_random vs. Bayesian)

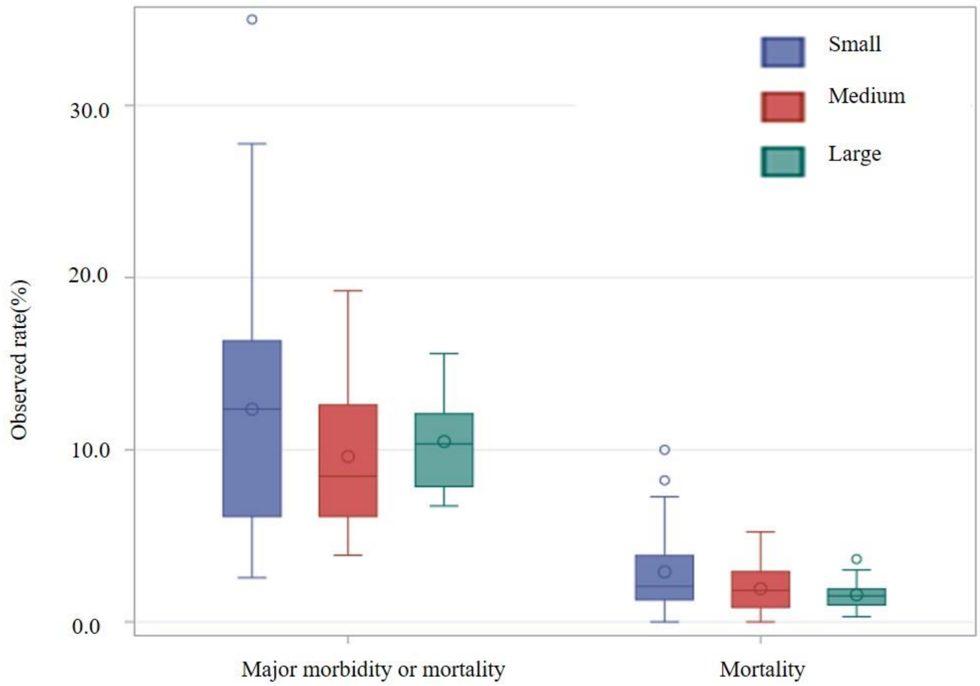
Quintiles based on Bayesian method with shrinkage	Quintiles based on Dir_random method with shrinkage				
	1 "Low"	2	3	4	5 "High"
1 "Low"	18	0	0		
2		18	0		
3		0	18		
4				17 (94.4%)	1 (5.6%)
5 "High"				1 (5.6%)	17 (94.4%)
Hospitals with changing rankings	0	0	0	1 (5.6%)	1 (5.6%)
small	0	0	0	1	0
medium	0	0	0	0	1
large	0	0	0	0	0

Supplemental Table 3. Hospital size in top 10% and bottom 10% ranking from each method.

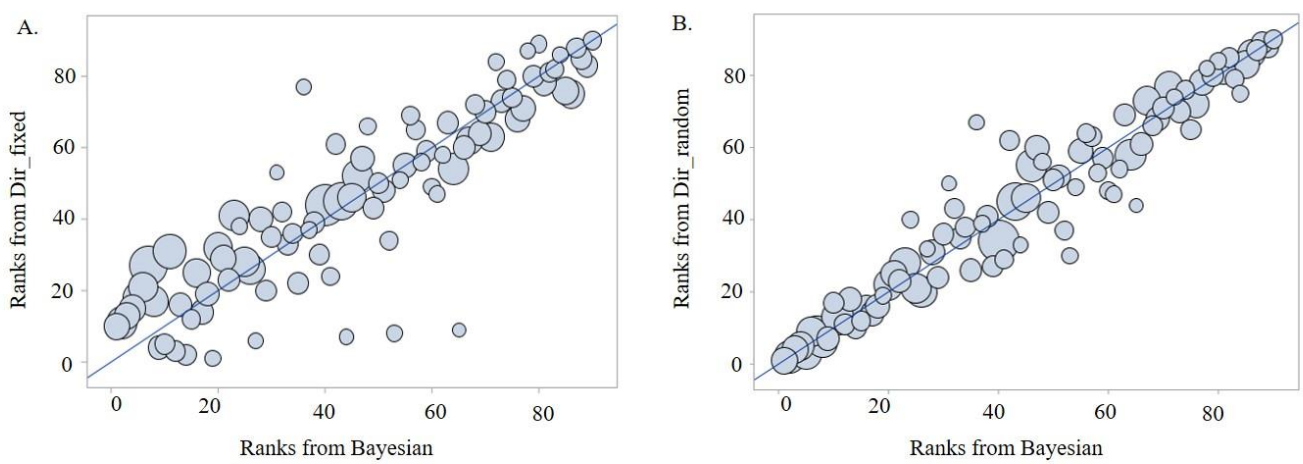
Outcome: mortality	Top 10%			Bottom 10%		
Methods	Small	Medium	Large	Small	Medium	Large
Observed	5	4	0	6	3	0
Indir_logit	5	4	0	6	3	0
Indir_fixed	5	4	0	6	3	0
Indir_random	0	1	8	4	3	2
Dir_fixed	5	4	0	6	3	0
Dir_random	0	1	8	4	3	2
Bayesian	0	1	8	4	3	2

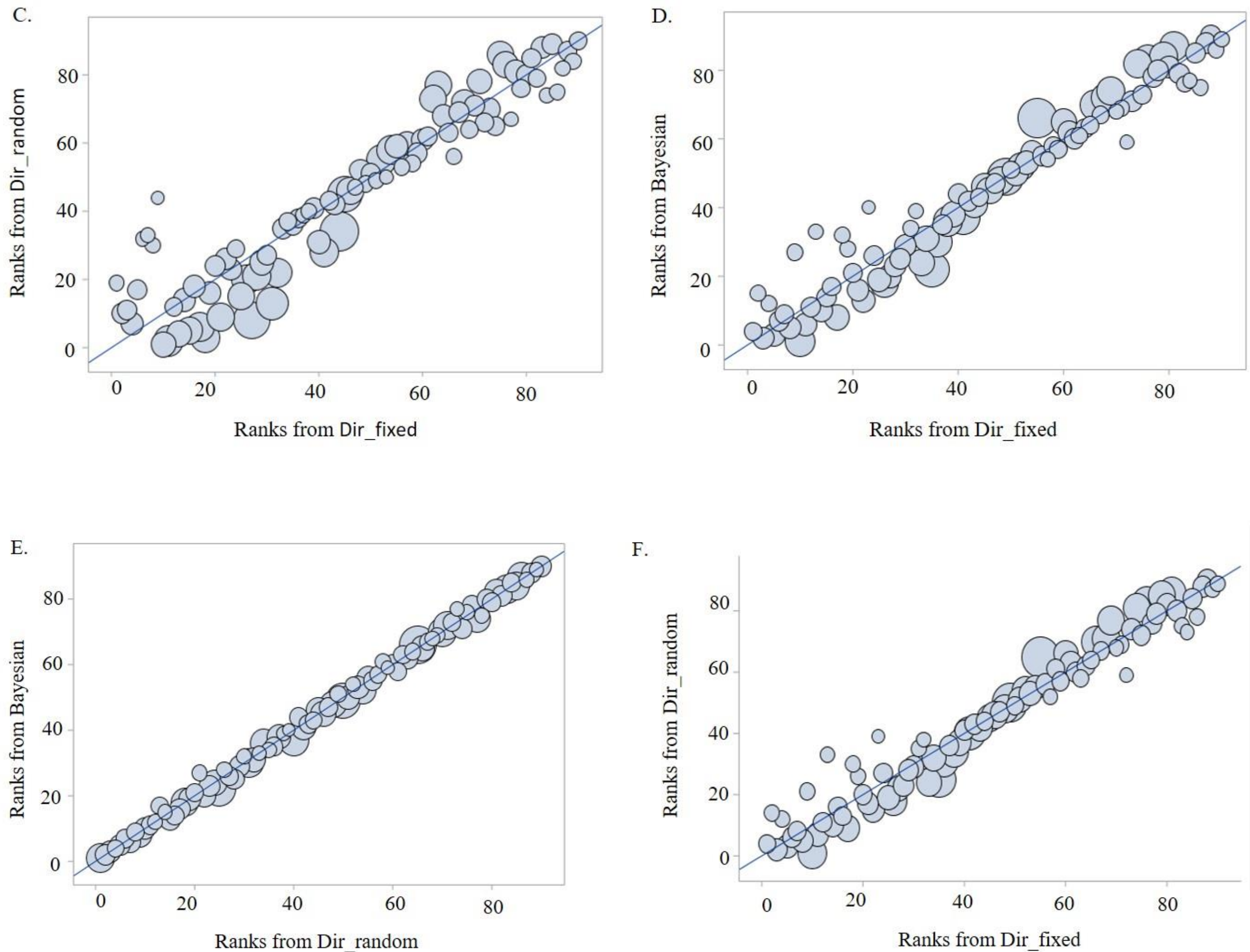
* The number represents the number of hospitals.

Supplemental Figure 1.



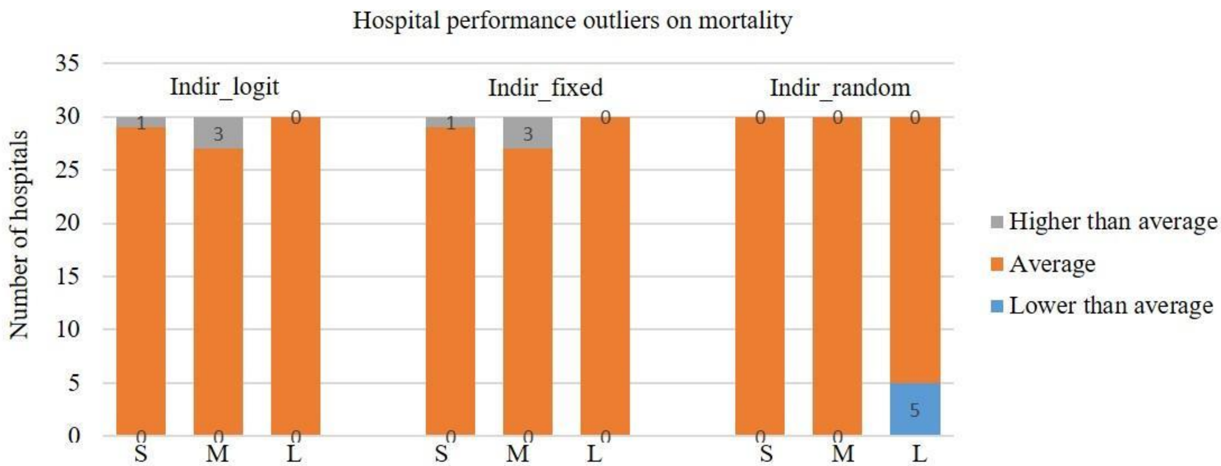
Supplemental Figure 2.



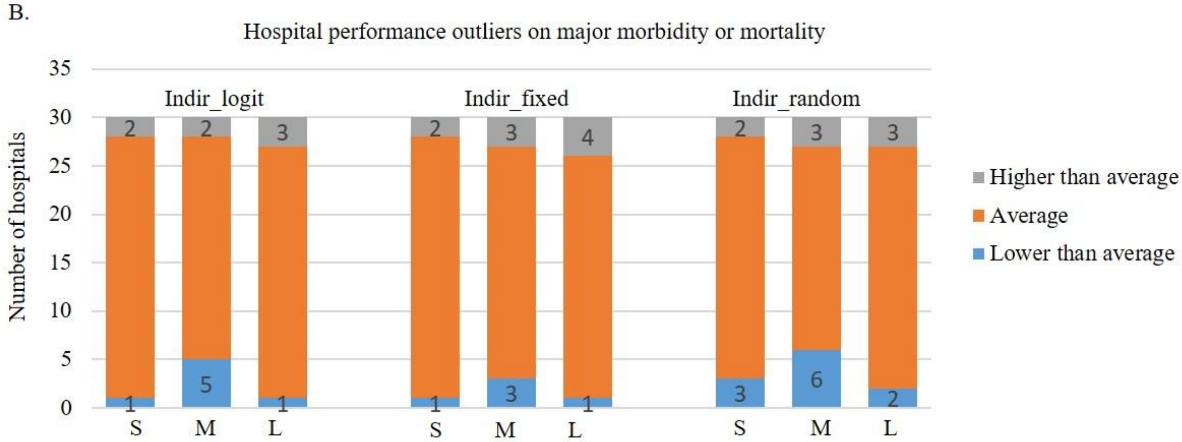

C.
Ranks from Dir_random
Ranks from Dir_fixed
D.
Ranks from Bayesian
Ranks from Dir_fixed
E.
Ranks from Bayesian
Ranks from Dir_random
F.
Ranks from Dir_random
Ranks from Dir_fixed
0
20
40
60
80

Supplemental Figure 3A

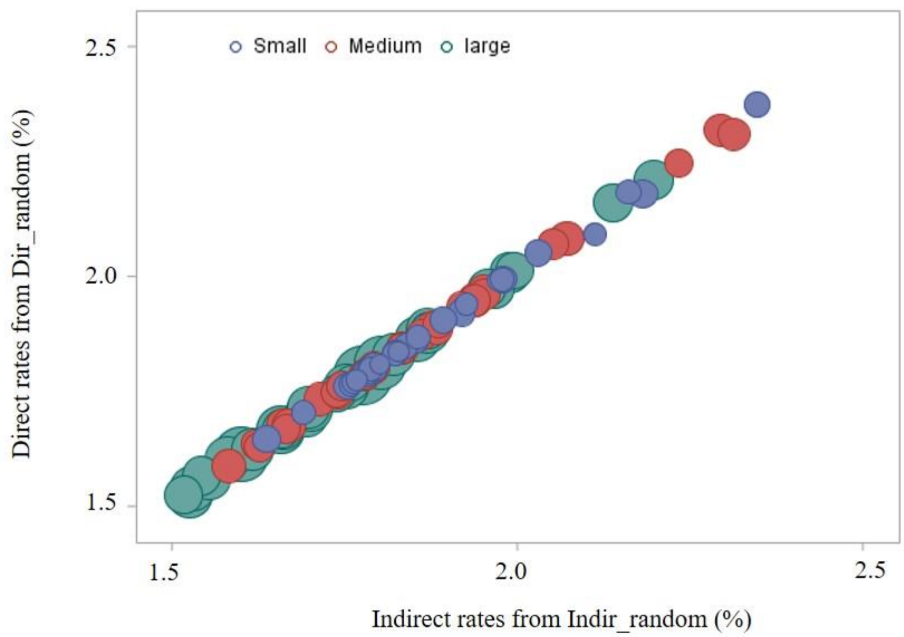
A.



B.



Supplemental Figure 4.



Appendix B. Supplemental Methods and Codes

We applied six statistical approaches (Table 2) to estimate hospital mortality, major morbidity or mortality, adjusting for patient risk. These approaches include: Indir_logit, indirect standardization with logistic regression models without hospital effect; Indir_fixed, indirect standardization with hospital fixed effect models; Indir_random, indirect standardization with hospital random effect models; Dir_fixed, direct standardization with fixed effect models; Dir_random, direct standardization with random effect models; Bayesian, the Bayesian method.

Let N be total number of patients from 90 hospitals, n_j be the number of patients in hospital j, p_j be the number of observed outcomes in hospital j , R be overall observed event rate in 90 hospitals: $R = \sum_{j=1}^{90} p_j / N$.

Method 1: Indir_logit

Step 1: Obtain the patients covariates estimates β from standard logistic model without hospitals effect using the developing dataset.

$$logit\ (P(Y_{ij}|X_{ij}) = \alpha + \beta X_{ij}$$

α : intercept; β : regression coefficients; X_{ij} : patients covariates describing characteristics of patients i in hospital j; Y_{ij}: response of patients i in hospital j.

Step 2: Offset the β from step 1, and fit the logistic model to ranking data (2016 data), and obtain the intercept α_{2016}

Step 3: Calculate hospital indirect standardized rates using Observed to expected (O/E) ratio:

$r_j = \frac{O_j}{E_j} \times R$. O_j: sum of the observed number of outcome in hospital j. E_j: sum of the expected number of outcomes in hospital j. E_j is estimated summing the individual predictive probability from the logistic model, $E_j = \sum_{i=1}^{n_j} \frac{\exp(\alpha_{2016}+\beta X_{ij})}{1+\exp(\alpha_{2016}+\beta X_{ij})}$. Clopper-Pearson exact 95% binomial confidence interval (CI) was used to construct the 95% CI for O/E ratios.

Method 2: Indir_fixed

Step 1. Obtain the patients covariates coefficients from fixed effect models, adjusted for patient covariates and accounted for hospital fixed effects, using 2014-2016 data

$$logit\ (P(Y_{ij}|X_{ij}) = \alpha + \beta X_{ij} + \theta_j$$

α : intercept; β : regression coefficients; X_{ij}: patients covariates describing characteristics of patients i in hospital j; θ_j: fixed effect of hospital j, which are the coefficients of hospital dummy variables

Step 2. offset the β from step 1, and fit the fixed effect model to data 2016, and obtain the intercept α_{2016} and hospital effect θ_{2016j}

Step 3. calculate hospital indirect standardized rates using Observed to expected (O/E) ratio:

$r_j = \frac{O_j}{E_j} \times R$. O_j: sum of the observed number of outcome in hospital j. E_j: sum of the expected number of outcomes in hospital j. E_j is estimated summing the individual predictive probability from the fixed model with the median hospital effect, $E_j = \sum_{i=1}^{n_j} \frac{\exp(\alpha_{2016}+\beta X_{ij}+median\ (\theta_{2016j})}{1+\exp(\alpha_{2016}+\beta X_{ij}+median\ (\theta_{2016j})}$. Clopper-Pearson exact 95% binomial CI was used to construct the 95% CI for O/E ratios

Method 3: Indir_random

Step 1: Random effects models, adjusted for patient covariates and accounted for hospital random effects with empirical Bayes estimates.

$$logit\ (P(Y_{ij}|X_{ij}) = \alpha + \beta X_{ij} + \theta_j$$

α : intercept; β : regression coefficients; X_{ij} : patients covariates describing characteristics of patients i in hospital j; θ_j: random effect of hospital j, which are the random intercepts for hospitals, assuming the hospital effects are drawn from a normal distribution with mean of zero and variance.

Step 2. offset the β from step 1, and fit the random effect model to data 2016, and obtain the intercept α_{2016} and hospital random intercepts θ_{2016j} for hospital random effect in 2016.

Step 3. calculate hospital indirect standardized rates using predicted to expected (P/E) ratios:

$r_j = \frac{P_j}{E_j} \times R$. P_j: sum of the predicted number of outcomes in hospital j, $P_j = \sum_{i=1}^{n_j} \frac{\exp(\alpha_{2016}+\beta X_{ij}+\theta_{2016j})}{1+\exp(\alpha_{2016}+\beta X_{ij}+\theta_{2016j})}$. E_j: sum of the expected number of outcomes in hospital j. In another word, E_j is estimated summing the individual predictive probability from the random model with the average hospital effect, $E_j = \sum_{i=1}^{n_j} \frac{\exp(\alpha_{2016}+\beta X_{ij})}{1+\exp(\alpha_{2016}+\beta X_{ij})}$. Bootstrapping 95% CIs were constructed to identify significant performance outliers[12.]

Method 4: Dir_fixed

Step 1 and 2. Same as step 1 and 2 in Method 2 - Indir_fixed:

Step 3. Calculate the direct standardized rates using all patients (N from all hospitals) as the reference.

Dir_fixed rates for hospital j= $\frac{\sum_{i=1}^{N} \frac{\exp(\alpha_{2016}+\beta X_{ij}+\theta 2016_j)}{1+\exp(\alpha_{2016}+\beta X_{ij}+\theta_{2016j})}}{N}$

Method 5: Dir_random

Step 1 and 2. Same as step 1 and 2 in Method 3 - Indir_fixed:

Step 3. Calculate the direct standardized rates using all patients (N from all hospitals) as the reference.

Dir_random rates for hospital j= $\frac{\sum_{i=1}^{N} \frac{\exp(\alpha_{2016}+\beta X_{ij}+\theta 2016_j)}{1+\exp(\alpha_{2016}+\beta X_{ij}+\theta_{2016j})}}{N}$

Method 6: Bayesian methods

The Bayesian method was implemented based on the STS approach [11]. Diffuse prior was specified. Hospital performance was assigned to average if risk standardized rates was statistically indistinguishable from the average rate based on 95% Bayesian certainty criterion.

[11] O'Brien SM, Shahian DM, DeLong ER, Normand SL, Edwards FH, Ferraris VA, Haan CK, Rich JB, Shewan CM, Dokholyan RS, Anderson RP and Peterson ED. Quality measurement in adult cardiac surgery: part 2--Statistical considerations in composite measure scoring and provider rating. *Ann Thorac Surg*. 2007;83:S13-26.

[12] Arlene S. Ash PSEF, PhD; Thomas A. Louis, PhD; Sharon-Lise T. Normand, PhD; Th´er`ese A. Stukel, PhD; Jessica Utts, PhD. STATISTICAL ISSUES IN ASSESSING HOSPITAL PERFORMANCE. *The COPSS-CMS White Paper Committee*. 2012.

Supplemental Codes

Please find the analytic R codes in the following link:

https://github.com/wu-git/hospital_benchmarking-.git