Lecture 9

Evaluating "Quality"

Application to Ranking

(using Multi-Level Logistic Models)

Lecture 9 Outline

- Profiling/Ranking:
  - One key application for multilevel models
  - What is profiling?
    - Definitions
    - Statistical challenges
    - Examples of possible quality measures
- Historical example:
  - Profiling medical care providers: a case-study
    - Hierarchical logistic regression model
    - Side bar on fitting multi-level models using MCMC
    - Performance measures
    - Comparison with standard approaches
- What is happening today?

Markov
Chain
Monte
Carlo

Objectives of profiling/ranking in health care setting

- Estimate provider or hospital-specific performance measures:
  - measures of utilization
  - patient outcomes
  - satisfaction with care

- Compare provider-specific estimates to a community norm in order to identify preferred providers

First analysis!

- Health Care Financing Administration (HCFA, now CMS) in 1987
- Evaluated hospital performance by comparing observed and expected mortality rates for Medicare patients

- $SMR = \frac{\#\ observed\ deaths}{\#\ expected\ deaths}$

- Hospitals with higher-than-expected mortality rates were flagged as institutions with potential quality problems, SMR > 1

First analysis!

- Basic outline for analysis:
  - Take large sample of Medicare data representing M hospitals
  - Estimate a patient-level model of mortality using standard logistic regression
  - Calculate the "expected mortality rate"* for each hospital: average the model-based probabilities of mortality for patients within that hospital
  - Calculate the ratio of observed to "expected mortality" rates, called standardized mortality rate (SMR)

- Hospitals with SMR > 1 labeled as poor providers
  - Problems here?

* ignoring heterogeneity in mortality rates

Statistical Challenges

- Hospital profiling needs to take into account
  - Patients characteristics (compositional effects)
  - Correlation between outcomes of patients within the same hospital
  - Number of patients and/or events in the hospital
  - Hospital characteristics

- These data characteristics motivate the centrality of multi-level data analysis

“Case-mix” bias

- Estimating hospital specific mortality rates without taking into account patient characteristics
  - *Suppose that older and sicker patients with multiple diseases have different needs for health care services and different health outcomes independent of the quality of care they receive.*
  - *In this case, physicians who see such patients may appear to provide lower quality of care than those who see younger and healthier patients*

- Develop patient-level regression models to control for different case-mixes
  - Pre-admission, patient-level health characteristics

Where have we seen "case-mix" bias before?

compositional effects

inverse of ecologic effect

Within cluster correlation

- Hospital practices may induce a strong correlation among patient outcomes within hospitals even after accounting for patients' characteristics

  why? discuss!

- Extend standard regression models to multi-level models that take into account the naturally clustered structure of the data

Provider/Hospital Sample Sizes Vary

- Both number of patients and the number of events if outcome is binary ($y_{i\cdot}$ and $n_{i\cdot} - y_{i\cdot}$ are important)
- Reliability of hospital-specific estimates:
  - *because of difference in hospital sample sizes, the precision of the hospital-specific estimates may vary greatly. SEE ARCH INTERN MED papers*
  - *** *Large differences between observed and expected mortality rates at hospitals with small sample sizes may be due primarily to sampling variability*
- Implement shrinkage estimation methods: hospitals performances with small sample size will be shrunk toward the mean more heavily

Profiling vs. Decision Making

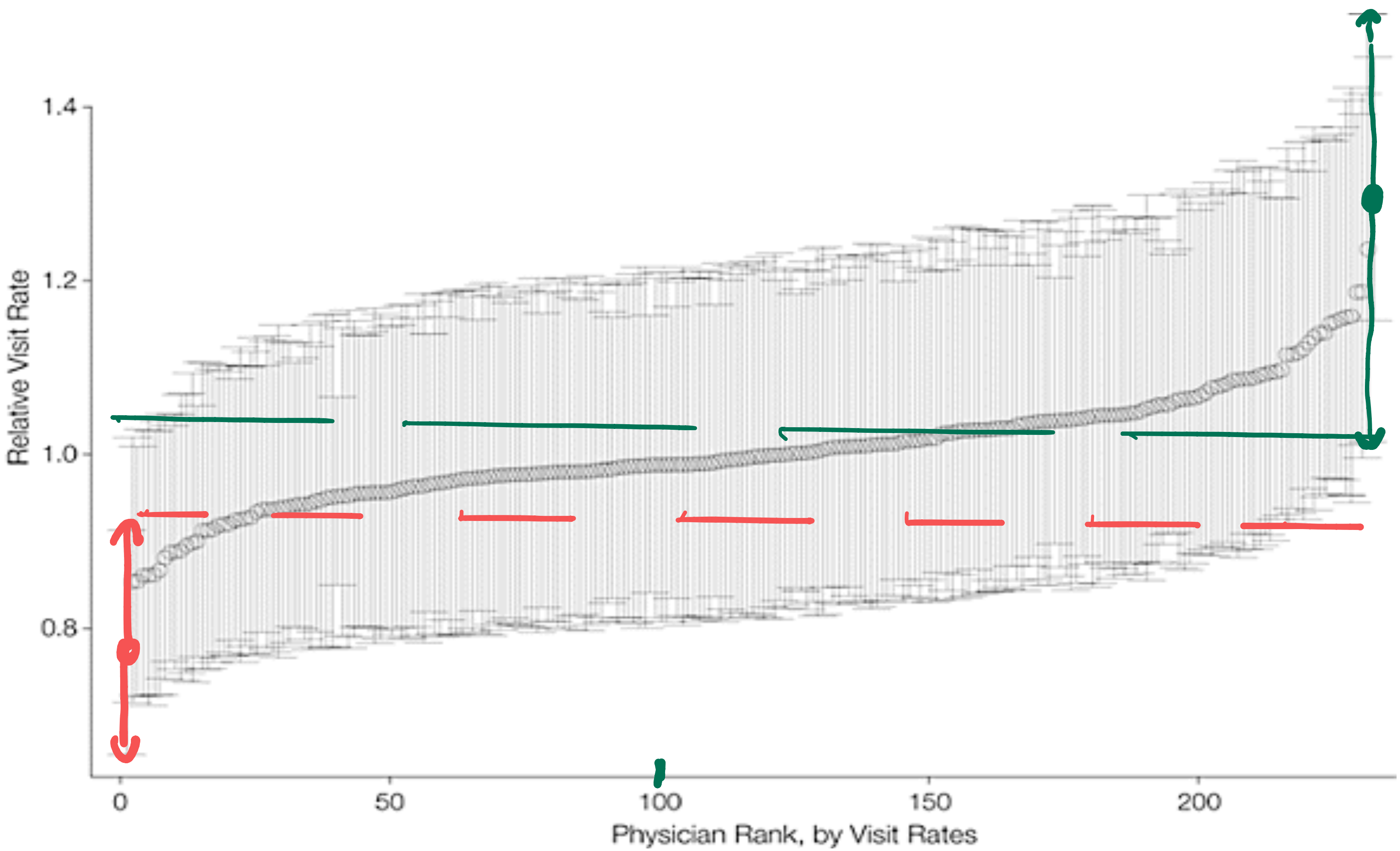
- Profiling:
  - How does this hospital's mortality for a particular procedure/condition compare to that predicted at the national level for the same kinds of patients seen at this hospital?
- Decision Making:
  - Given my medical status and needs, to which hospital should I go for a particular procedure or treatment of my condition?

- Profiling aggregates over a mix of patients:
  - Low SMR may not mean that hospital performs well for a specific patient
  - High SMR may not mean hospital performs poorly for all patients
- Decision making goal would require case-mix adjustment and perhaps adjustment for hospital specific features (special labs or procedures) and prediction for the individual rather than the whole

Example of Measures of Performance

- Patient outcomes (e.g. patient mortality, morbidity, satisfaction with care)
  - For example: 30-day mortality among heart attack patients (*Normand et al JAMA 1996, JASA 1997)*
- Process (e.g were specific medications given or tests done, costs for patients)
  - For example: laboratory costs of patients who have diabetes (*Hofer et al JAMA, 1999)*
  - Number of physician visits (*Hofer et al JAMA, 1999)*

Relative visit rate by physician (with 1.0 being the average profile after adjustment for patient demographic and detailed case-mix measures). The error bars denote the CI, so that overlapping CIs suggest that the difference between the two physician visit rates is not ~~statisticly~~ statistically significant (Hofer et al JAMA 1999)



Hospital Profiling of Mortality Rates for Acute Myocardial Infarction Patients (*Normand et al JAMA 1996, JASA 1997*)

- Data characteristics
- Scientific goals
- Multi-level logistic regression model
- Definition of performance measures
- Estimation (MCMC)
- Results
- Discussion

Data Characteristics

- The Cooperative Cardiovascular Project (CCP) involved abstracting medical records for patients discharged from hospitals located in Alabama, Connecticut, Iowa, and Wisconsin (June 1992- May 1993)

- 3,269 patients hospitalized in 122 hospitals in four US States for Acute Myocardial Infarction

Data characteristics

- Outcome: mortality within 30-days of hospital admission

- Patients characteristics:
  - Admission severity index constructed on the basis of 34 patient characteristics   right ones?

- Hospital characteristics
  - Rural versus urban
  - Non-academic versus academic
  - Number of beds

Admission severity index (Normand et al 1997 JASA)

Jour Am Statistical Assoc methodologic not substantive

Table A.1 Admission Severity Variables and Weights Comprising the Admission Severity Index

X_p	$\hat{\beta}_p$	X_p	$\hat{\beta}_p$
Constant	5.5726	LV function proxies:	
Socio-demographic:		Cardiac arrest	.9069
(Age—65)	.0681	Gallop rhythm	−.0310
(Age—65)^2	−.0010	Cardiomegaly	−.0094
Admission history:		Hx CHF	−.1061
Hx cancer	−.1740	Rales and pulmonary edema	.1520
Admission severity:		Laboratory results:	
Mobility status		Albumin > 3 (g/dl)	−.4828
Walked independently	−.2740	Albumin missing	−.4793
Unable to walk	.4700	Log_{10}[BUN (mg/dl)]	1.0613
Mobility missing	.3669	BUN missing	1.4583
Body mass index (kg/m^2)	−.0259	Creatinine > 2 (mg/dl)	.3279
Body mass missing	−.1525	Creatinine missing	.1937
Respiration rate breaths/min		Diagnostic test results:	
Respiration (if ≥ 12)	.0429	Conduction disturbance	.4084
Respiration < 12	3.4840	No EKG (vs EKG reading)	.5050
Respiration missing	2.2666	No MI on EKG (vs MI on EKG)	−.1430
Ventricular rate > 100	.1564	Anterior MI (vs other MI)	.4384
Log_{10}(MAP)	−4.7101	Lateral MI (vs other MI)	.2908
MAP missing	−10.1796	Posterior MI (vs other MI)	.6416
Shock	1.6194	Lateral and posterior MI	−.8767

NOTE. Hx = history, MAP = mean arterial pressure; BUN = blood urea nitrogen level. Variables indicate the presence of the condition (coded 1 if present and 0 otherwise) with the exception of the following seven continuous covariates, which assume the observed values: age, body mass, respiration rate, MAP, albumin, BUN, and creatinine. The severity index is calculated as $\Sigma_p \hat{\beta}_p x_{ip}$ for the ith patient.

Scientific Goals

- Identify "aberrant" hospitals in terms of several performance measures
- Report the statistical uncertainty associated with the ranking of the "worst hospitals"
- Investigate if hospital characteristics explain heterogeneity of hospital-specific mortality rates

- what is the dimensionality of quality? of ranking?
- What role do politics + capitalism play in ranking methods?
- Why do we rank everything?

Multilevel logistic regression model
(hierarchical, mixed, random effects,….)

- I: *patient level, within-provider model*
  - Patient-level logistic regression model with random intercept and random slope

hospital-specific

- II: *between-providers model*
  - Hospital-specific random effects are regressed on hospital-specific characteristics

Patient-level model

- Y_{ij} is the binary indicator of death within 30 days of admission for patient j at hospital i
- severity_{ij} is the severity index for patient j at hospital i
- $\overline{\text{severity}}$ is average severity index

$$\text{logit}\, P^{*}(Y_{ij} = 1) = \beta_{0i} + \beta_{1i}(\text{severity}_{ij} - \overline{\text{severity}})$$

- β_{0i} and β_{1i} are random intercept and slope
- β_{0i} denotes the log-odds of death for hospital i ~~having~~ for patients with severity equal to the population average
- β_{1i} denotes the hospital-specific association between severity and logit of probability of death

$* \; Pr(Y_{ij} = 1 \mid \beta_{0i}, \beta_{1i}, x_j)$

Hospital-level models

- Without hospital-specific covariates (Exchangeable)

$$
\begin{aligned}
\beta_{0i} &= \gamma_{00}^{\star} + N(0, \sigma_0^{\star 2}) \\
\beta_{1i} &= \gamma_{10}^{\star} + N(0, \sigma_1^{\star 2})
\end{aligned}
$$

- With hospital-specific covariates (Non-Exchangeable)

$$
\begin{aligned}
\beta_{0i} &= \gamma_{00} + \gamma_{01}\mathbf{rural}_i + \gamma_{02}\mathbf{no\text{-}acad}_i + \gamma_{03}\mathbf{small} + \gamma_{04}\mathbf{medium}_i + N(0, \sigma_0^2) \\
\beta_{1i} &= \gamma_{10} + \gamma_{11}\mathbf{rural}_i + \gamma_{12}\mathbf{no\text{-}acad}_i + \gamma_{13}\mathbf{small} + \gamma_{14}\mathbf{medium}_i + N(0, \sigma_1^2)
\end{aligned}
$$

where rural, non-acad, small, medium are indicators of a rural, non academic, small, and medium size beds hospitals, respectively

Interpretation of the Random Effects

- Random intercept
  - inter hospital differences in "baseline" mortality rates (mortality among patients with population average severity)
    - Does the grand mean centering make sense here?

- Random slope
  - effect of clinical burden (patient severity) on mortality differs across hospitals

- Models may or may not include hospital level covariates (as on prior slides) to help explain some of the heterogeneity in these two random effects

Normand et al JASA 1997

Table 1. *Patient and Hospital Characteristics in the Study Cohort*

$\bar{X}_{i\cdot}$

	25th percentile	*Median*	*Mean*	*75th percentile*
Observed Mortality				
Across hospitals	.14	.22	.22	.29
Admission severity				
Across patients	−2.47	−1.80	−1.65	− .99
Across hospitals	−1.47	−1.49	−1.47	−1.22

$\underset{\sim}{X}_i$

Hospital characteristics	*% of patients*		*% of Hospitals*
Rural (vs. urban)	54	?	76
Nonacademic (vs. academic)	79		88
Number of beds			
≤100 (small)	29	?	64
101–299 (medium)	27		21
≥300 (large)	44	?	15

Two ways to average

$\overline{Y}_{AP}$: "Across patients" : $\displaystyle\left(\frac{1}{\sum_{i=1}^{m} n_i}\right) \sum_{i=1}^{m} \sum_{j=1}^{n_i} Y_{ij}$

$\overline{Y}_{AH}$: Across hospitals : $\displaystyle\frac{1}{m} \sum_{i=1}^{m} \underbrace{\left(\frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}\right)}_{\overline{Y}_{i\cdot}}$

$$\overline{Y}_{AP} = \frac{1}{m \cdot \overline{n}_{\cdot}} \sum_i n_i \overline{Y}_{i\cdot} = \frac{1}{m} \sum_{i=1}^{m} \left(\frac{n_i}{\overline{n}_{\cdot}}\right) \overline{Y}_{i\cdot}$$

weight bigger
n_i more

$$\overline{Y}_{AH} = \frac{1}{m} \sum_i \overline{Y}_{i\cdot}$$

weight all the same

In this analysis, they defined two hospital performance measures, we will define these and then discuss estimation via MCMC

- Let μ_i^A be the "adjusted" mortality rate for hospital i
- Let μ_i^S be the "standardized" mortality rate for a "reference" hospital "like" hospital i
- We assume that a provider's performance is poor if the probability that $\mu_i^A - \mu_i^S$ being bigger than some benchmark value is large. We estimate:

$$P_i^{A-S} = P(\mu_i^A - \mu_i^S > \text{benchmark}),$$

At least recognize statistical uncertainty

- We assume that a provider's performance is poor if the probability that $\mu_i^A - \mu_i^S$ being bigger than some benchmark value is large. We estimate from our model:

$$\begin{aligned}
P_i^{A-S} &= P(\mu_i^A - \mu_i^S > \text{benchmark}), \\
&\quad \text{where} \\
\mu_i^A &= \frac{1}{n_i}\sum_{j=1}^{n_i} P(Y_{ij}=1 \mid \underline{\beta_{0i}}, \underline{\beta_{1i}}, \text{sev}) \\
&= \frac{1}{n_i}\sum_{j=1}^{n_i} \text{logit}^{-1}(\underline{\beta_{0i}} + \underline{\beta_{1i}}(\text{sev}_{ij} - \overline{\text{sev}}))
\end{aligned}$$

$$\boxed{\begin{aligned}
\mu_i^S &= \frac{1}{n_i}\sum_{j=1}^{n_i} P(Y_{ij}=1 \mid \gamma_0, \gamma_1, \text{sev}) \\
&= \frac{1}{n_i}\sum_{j=1}^{n_i} \text{logit}^{-1}(\underline{\gamma_0} + \underline{\gamma_1}(\text{sev}_{ij} - \overline{\text{sev}}))
\end{aligned}}$$

treats hospitals the same; adjusts for case-mix

Hospital-Performance Measures

- Let $Y_i^\star = E[Y_i \mid, \overline{\text{sev}}, \beta]$ be the probability of death for an "average" patient with severity index equal to $\overline{\text{sev}}$ treated in hospital i

"les homme moyen"

- Let M be the median probability of death for the same "average" patient across all hospitals

$M = \text{med}\, Y_i^\star$

- We define a hospital performance measure as the probability of excess mortality for the average patient
- The performance of hospital i is poor if the probability of death for an "average" patient treated in hospital i is large compared to M. More specifically, we are interested to estimate:

$$
\begin{aligned}
P_i^\star &= P(Y_i^\star > M) \\
&= P(\text{logit}^{-1}(\beta_{0i}) > M)
\end{aligned}
$$

Thoughts?


Quetelet


THIS
IS THE
WORLD'S
most
TYPICAL
PERSON

Fitting Multi-Level Models

- SAS, Stata, R
  - Maximum Likelihood Estimation (MLE)
  - Limitation: hard to estimate ranking probabilities and assess statistical uncertainty of hospital rankings

- Bayesian Computational Methods
  - Monte Carlo Markov Chains methods
  - Advantages: estimation of ranking probabilities and their confidence intervals is straightforward

What is MCMC?

- Markov Chain Monte Carlo
  - A class of algorithms for sampling from probability distributions based on constructing a Markov chain that has the desired distribution as its limit.

- In MLMs, the goal is to estimate the distribution of the unknown parameters and random effects given the observed data.
  - This is known as the posterior distribution of the unknowns

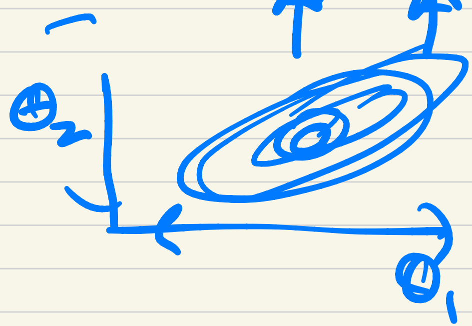
Likelihood $Pr(\underset{\sim}{Y} \mid \underset{\sim}{\theta}) = \dfrac{1}{\sqrt{2\pi}\,\theta_2} \exp\left(-\dfrac{1}{\theta_2^2}(Y-\theta_1)^2\right)$

$[\underset{\sim}{Y} \mid \theta]$

$Y \sim N(\theta_1, \theta_2^2)$

Prior : $\pi(\underset{\sim}{\theta})$

$[\underset{\sim}{\theta}]$



Belief

$\theta_1 >> 0$

$+10, 100$

Posterior

$[\underset{\sim}{\theta} \mid \underset{\sim}{Y}]$: $\underbrace{Pr(\theta \mid Y)}$ $= \dfrac{Pr(Y \mid \theta)\,\pi(\theta)}{\underbrace{\int Pr(Y \mid \theta)\,\pi(\theta)\,d\underset{-}{\theta}}}$ $\leftarrow Pr(Y \text{ and } \theta)$ $\leftarrow Pr(Y)$

How does MCMC work?

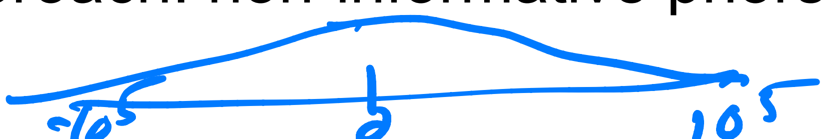
- Y_{ij} is the indicator of death within 30-days of admission for subject *j* from hospital *i*
- $Y_{ij} \sim Bernoulli(p_i)$
- $logit[Y_{ij} = 1 | b_{oi}] = \beta_0 + b_{0i}$
- $b_{0i} \sim N(0, \sigma^2)$

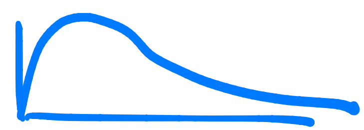
$[\beta_0, \sigma^2, b_{0i}, | i=1,\dots,n$

$Y_{ij}]$

$n = 1,\dots,n$

- In fully Bayesian approach, we need to specify priors for β_0 and σ^2
  - Common approach: non-informative priors
  - $\beta_0 \sim N(0, big)$
  - $\tau = {}^1/_\sigma \sim Gamma(small, small)$





How does MCMC work?

GOAL: update the priors on $\beta_0,\ \sigma^2$ by incorporating the data (Y_{ij})

- This updating will generate the posterior distribution for each of these parameters

$$[\beta_0, \sigma^2, b_i \mid \underset{\sim}{Y}] = \frac{[Y \mid \beta_0, \sigma^2][\beta_0, \sigma^2]}{\int [\quad] \, d\beta_0 \, d\sigma^2}$$

General Case:

$$Y_{ij} = X_{ij}\underline{\beta} + Z_{ij}\underline{b_i} + \varepsilon_{ij}, \quad \varepsilon_{ij} \overset{iid}{\sim} G(0, \underline{\sigma^2}), \quad b_i \perp \varepsilon_{ij}$$

$$b_i \overset{iid}{\sim} G(0, \underline{D})$$

e.g. "parameters": β, D, σ^2, b_i, $i = 1, \dots, m$

data (Y_{ij}, X_{ij}, Z_{ij}) $j = 1, \dots, n_i$; $i = 1, \dots, m$

We seek the posterior dist of

$$\left[\underset{\sim}{\beta}, \sigma^2, D, \underset{\sim}{b_i} \;\middle|\; \underset{\sim}{Y_i}, \underset{\sim}{X_i}, \underset{\sim}{Z_i}\right]$$

$i = 1, \dots, m$ (under b_i) $\qquad i = 1, \dots, m$ (under the data)

Complicated!

How does MCMC work? - $[\beta_0, \sigma^2 | \underset{\sim}{Y}_i]$

simpler case

0. Start with initial values for $\beta_0^{(0)}$ and $\sigma^{2^{(0)}}$

1a. Sample $\beta_0^{(1)} \sim f(\beta_0^{(1)}|Y_{ij}, \sigma^{2^{(0)}})$

1b. Sample $\sigma^{2^{(1)}} \sim f(\sigma^{2^{(1)}}|Y_{ij}, \beta_0^{(1)})$

2a. Sample $\beta_0^{(2)} \sim f(\beta_0^{(2)}|Y_{ij}, \sigma^{2^{(1)}})$

2b. Sample $\sigma^{2^{(2)}} \sim f(\sigma^{2^{(2)}}|Y_{ij}, \beta_0^{(2)})$ ….

$\begin{pmatrix} \beta_0^{(1)} \\ \sigma^{2(1)} \end{pmatrix} \begin{pmatrix} \beta_0^{(2)} & \cdots \\ \sigma^{2(2)} & \cdots \end{pmatrix}$

We continue to sample and then the sampled values are draws from the target distribution.

At each iteration (k), we can compute

- b_{0i}, for i =1, …,m $\quad b_{0i}^{(k)} = \bar{Y}_{i\cdot} - \beta_0^{(k)}$
- any other statistic of interest derived from the model

How does MCMC work?

- We take as our "point estimates" of β_0, σ^2, and other statistics (e.g. b_{0i} *i = 1, …, m*)
  - The mean of the posterior samples
  - Our standard error estimates represent the standard deviation of the posterior distribution
- What else can we calculate?
  - Median of posterior distribution
  - Distribution of the ranks of b_{0i}
    - For each posterior sample, rank b_{0i} and store the rank information
    - After many posterior samples, display the distribution of the ranks.

https://chi-feng.github.io/mcmc-demo/

Estimating Hospital Profiling Measures with MCMC

- **Probability of a large difference between adjusted and standardized mortality rates:**

$$
\begin{aligned}
R_i^k &= \mu_i^{A^k} - \mu_i^{S^k} = \\
&= \frac{1}{n_i}\sum_{j=1}^{n_i}\left[\text{logit}^{-1}(\beta_{0i}^k + \beta_{1i}^k x_{ij}) - \text{logit}^{-1}(\gamma_0^k + \gamma_1^k x_{ij})\right] \\
\hat{P}_i^{A-S} &= \frac{1}{K}\sum_{k=1}^{K} I(R_i^k > H)
\end{aligned}
$$

where:

- $x_{ij} = \text{sev}_{ij} - \overline{\text{sev}}$ } person-specific deviation from grand-mean
- H is a benchmark value that can be calculated based upon the distribution of R_i^k across hospitals

Estimating Hospital Profiling Measures with MCMC

- **Probability of excess mortality for the average patient:**

$$\hat{P}_i^{\star} \;=\; \tfrac{1}{K}\textstyle\sum_{k=1}^{K} I(\text{logit}^{-1}(\beta_{0i}^{k}) > M)$$

where:

- M is the median of $\left\{\text{logit}^{-1}(\beta_{0i}^{k}),\ i = 1,\ldots,96\right\}$

Hospital-Performance Measures
Standard Logistic Regression Approach

- The HCFA's algorithm to identify "aberrant hospitals" does not rely on multilevel models
- A logistic regression is fitted to the data, and z-scores were derived from the standardized difference between observed and expected mortality in each hospital. More specifically

$$z_i = n_i(\bar{Y}_i - \bar{p}_i) / \sqrt{\textstyle\sum_{j=1}^{n_i} \hat{p}_{ij}(1 - \hat{p}_{ij})}$$

where

$$\hat{p}_{ij} = \text{logit}^{-1}(\hat{\beta}_{0i} + \hat{\beta}_{1i}(\text{sev}_{ij} - \overline{\text{sev}}))$$

- hospitals with $z_i \geq 1.645$ (top 5%) were classified as "aberrant"

Comparing measures of hospital performance

- Three measures of hospital performance
  - Probability of a large difference between adjusted and standardized mortality rates
  - Probability of excess mortality for the average patient
  - Z-score

Results

- Estimates of regression coefficients under three models:
  - Random intercept only
  - Random intercept and random slope
  - Random intercept, random slope, and hospital covariates

- Hospital performance measures

Table 2. *Regression Estimates*

Level I parameter	Level II parameter	Estimated posterior summaries: Mean	SD	Mean/SD	Percentiles (2.5, 97.5)
Exchangeable model: Random-intercept model					
β_{0i}: Intercept	γ: Intercept	−1.70	.07	−24.29	(−1.85, −1.57)
	$\sigma^2_{\beta_0}$: Variance	$(.31)^2$	.05		(.01, .22)
β_1: Severity − $\overline{\text{severity}}$		1.03	.05	20.60	(.93, 1.13)
Exchangeable model: Random-intercept and slope model					
β_{0i}: Intercept	γ_{00}: Intercept	−1.72	.08	−21.53	(−1.87, −1.56)
β_{1i}: Severity − $\overline{\text{severity}}$	γ_{10}: Intercept	1.03	.05	19.67	(.94, 1.15)
				Estimated posterior mean	
	D: Variance			$\begin{pmatrix} (.42)^2 & -.03 \\ -.03 & (.21)^2 \end{pmatrix}$	
Nonexchangeable model: Random-intercept and slope model					
β_{0i}: Intercept	γ_{00}: Intercept	−1.79	.17	−10.29	(−2.15, −1.45)
	γ_{01}: Rural	.55	.20	2.76	(.15, .93)
	γ_{02}: Non-Academic	−.27	.27	−1.24	(−.71, .14)
	γ_{03}: Small	−.27	.25	−1.06	(−.74, .27)
	γ_{04}: Medium	.29	.20	1.46	(−.10, .67)
β_{1i}: Severity − $\overline{\text{severity}}$	γ_{10}: Intercept	1.22	.13	9.18	(.96, 1.52)
	γ_{11}: Rural	.05	.16	.33	(−.27, .36)
	γ_{12}: Nonacademic	−.11	.17	−.64	(−.44, .23)
	γ_{13}: Small	−.08	.20	−.39	(−.50, .28)
	γ_{14}: Medium	−.29	.15	−1.88	(−.58, .01)
				Estimated posterior mean	
	D: Variance			$\begin{pmatrix} (.35)^2 & -.03 \\ -.03 & (.22)^2 \end{pmatrix}$	

Estimates of log-odds of 30-day mortality for a ``average patient''

- Exchangeable model (without hospital covariates), random intercept and random slope:
  - **95% posterior interval** of the log-odds of 30-day mortality ***for a patient with average admission severity*** is equal to **(-1.87,-1.56)** ~~(-1.80 -1.64)~~
  - corresponding to **(0.13,0.17)** in the probability scale

.15 $e^{-1.87}/(1+e^{-1.87}) = .15$

- Non-Exchangeable model (with hospital covariates), random intercept and random slope:
  - **95% posterior interval** of the log-odds of 30-day mortality ***for a patient with average admission severity treated in a large, urban, and academic hospital*** **is equal** to **(-2.15,-1.45)**
  - corresponding to **(0.10,0.19)** in probability scale

Estimates of II-stage regression coefficients (intercepts)

$$\beta_{0i} = \gamma_{00} + \gamma_{01}\text{rural}_i + \gamma_{02}\text{no-acad}_i + \gamma_{03}\text{small} + \gamma_{04}\text{medium}_i + N(0, \sigma_0^2)$$

Nonexchangeable model: Random-intercept and slope model

β_{0i}: Intercept	γ_{00}: Intercept	−1.79	.17	−10.29	(−2.15, −1.45)
	γ_{01}: Rural	.55	.20	2.76	(.15, .93)
	γ_{02}: Non-Academic	−.27	.27	−1.24	(−.71, .14)
	γ_{03}: Small	−.27	.25	−1.06	(−.74, .27)
	γ_{04}: Medium	.29	.20	1.46	(−.10, .67)

Effect of hospital characteristics on baseline log-odds of mortality

- Baseline log-odds of mortality: log-odds of mortality for patients with average severity
- Rural hospitals have higher odds of mortality than urban hospitals for an average patient
- This is an indication of inter-hospital differences in the baseline mortality rates

Estimates of II-stage regression coefficients (slopes)

$$\beta_{1i} = \gamma_{10} + \gamma_{11}\text{rural}_i + \gamma_{12}\text{no-acad}_i + \gamma_{13}\text{small} + \gamma_{14}\text{medium}_i + N(0, \sigma_1^2)$$

β_{1i}: Severity − $\overline{\text{severity}}$	γ_{10}: Intercept	1.22	.13	9.18	(.96, 1.52)
	γ_{11}: Rural	.05	.16	.33	(−.27, .36)
	γ_{12}: Nonacademic	−.11	.17	−.64	(−.44, .23)
	γ_{13}: Small	−.08	.20	−.39	(−.50, .28)
	γ_{14}: Medium	−.29	.15	−1.88	(−.58, .01)

Effects of hospital characteristics on associations between severity and mortality (slopes)

- The association between severity and mortality is ``modified'' by the size of the hospitals

weak!

- Medium-sized hospitals having smaller severity-mortality associations than large hospitals

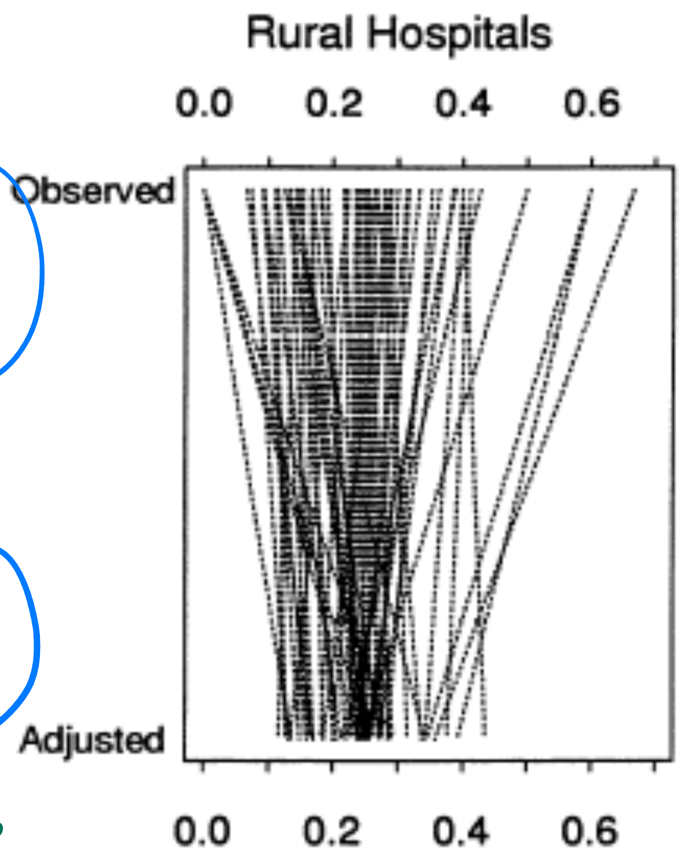
?

- This indicates that the effect of clinical burden (patient severity) on mortality differs across hospitals

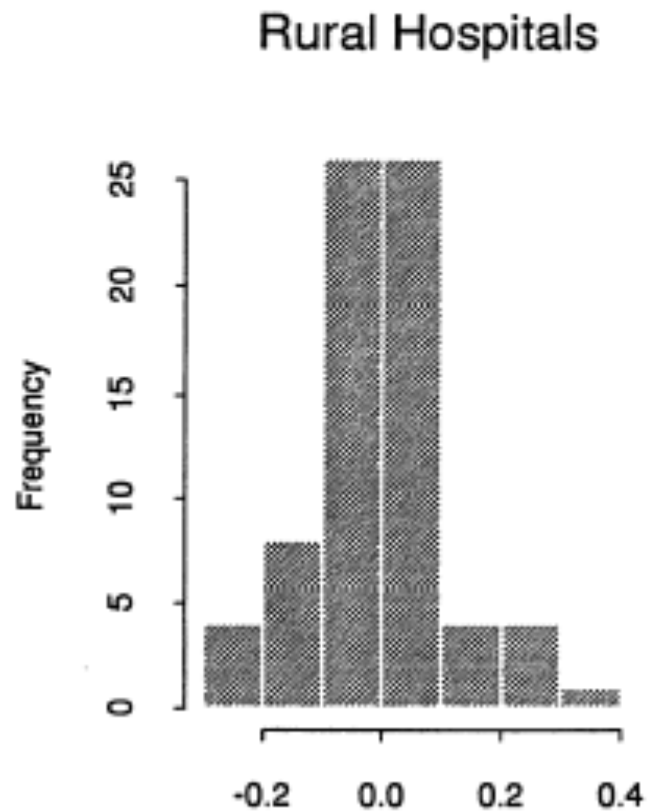
weakly

Observed and risk-adjusted hospital mortality rates: Crossover plots
Display the observed mortality rate (upper horizontal axis) and
Corresponding risk-adjusted mortality rates (lower horizontal line).
Histogram represents the difference = observed - adjusted

M.L. $\bar{Y}_{i\cdot}$

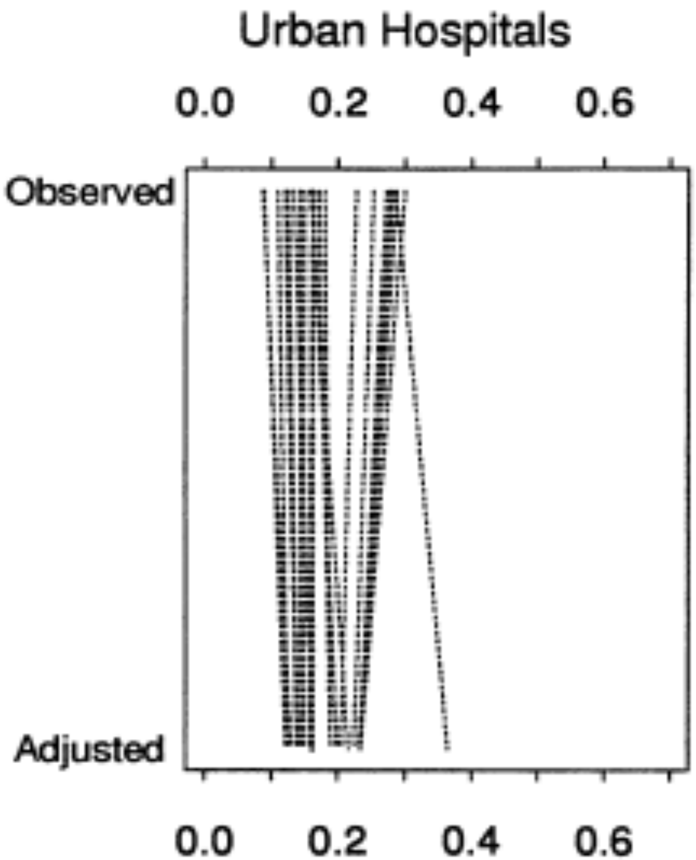


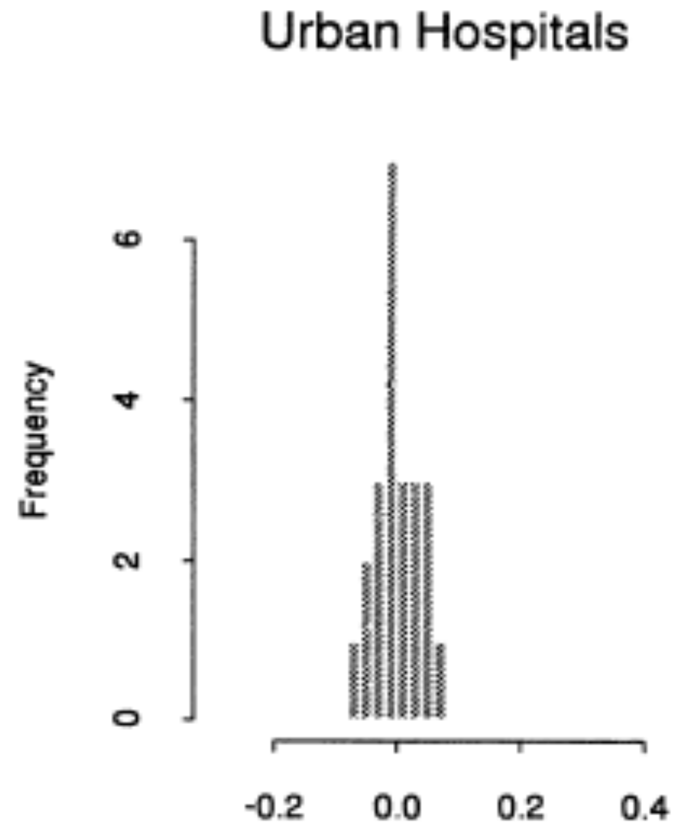
$\hat{\mu}_i$ Bayes



Substantial adjustment for severity!

Observed and risk-adjusted hospital mortality rates: Crossover plots
Display the observed mortality rate (upper horizontal axis) and
Corresponding risk-adjusted mortality rates (lower horizontal line).
Histogram represents the difference = observed – adjusted
(Normand et al JASA 1997)





What are these pictures telling us?

- Adjustment for severity on admission is substantial (mortality rate for an urban hospital moves from 29% to 37% when adjusted for severity)

- There appears to be less variability in changes between the observed and the adjusted mortality rates for urban hospitals than for rural hospitals

Hospital Ranking: *Normand et al 1997 JASA*

Table 4. HCFA Highest and Lowest Ranked Hospitals

						Random intercept				Random intercept and slope			
	No. of AMI	No.	Hospital	Academic	Hospital	HCFA		$\hat{P}_i^{A-S}$		$\hat{P}_i^{A-S}$		$\hat{P}_i^{*}$	
Hospital	patients	dead	location	(Y/N)	size	z_i	Rank	(%)	Rank	(%)	Rank	(%)	Rank
1	54	19	R	N	M	3.83	1	36	1	25	1	89	1
28	6	4	R	N	M	2.55	2	10	7	15	3	70	3
2	18	7	R	N	S	2.55	3	12	5	19	2	32	9
10	62	18	R	N	M	2.51	4	16	2	7	19	71	2
90	8	4	R	N	S	2.00	5	15	3	13	5.5	13	28
43	27	6	R	N	S	1.95	6	3	43	11	8	22	17
15	81	22	U	Y	L	1.82	7	9	11	5	26	10	31
44	7	3	R	N	S	1.75	8	6	20	5	33	11	30
95	22	8	R	N	S	1.68	9	12	6	14	4	16	21
29	31	5	U	N	S	−1.75	93	0	84.5	1	74	0	94
39	6	0	R	N	S	−1.77	94	2	54	3	48.5	2	77
19	46	4	U	N	L	−1.80	95	0	90.5	0	90.5	0	94
42	70	11	U	Y	L	−2.01	96	0	94.0	0	94.5	0	94

NOTE: HCFA highest-ranked ($z_i > 1.65$) and lowest-ranked ($z_i < -1.65$) hospitals. The rank of each measure is from worst (1) to best (96). L denotes hospitals with $\geq$300 beds, M denotes hospitals with 101–299 beds, S denotes hospitals with fewer than 101 beds, R denotes rural hospitals, and U denotes urban hospitals.

What type of statistical information would you suggest adding ?

Ranking of hospitals

- There was moderate disagreement among the criteria for classifying hospitals as ``aberrant”

- Despite this, hospital 1 is ranked as the worst. This hospital is rural, medium sized non-academic with an observed mortality rate of 35%, and adjusted rate of 28%

Summary of Lecture 9

- Profiling medical providers is a multi-faced and data intensive process with significant implications for health care practice, management, and policy

- Major issues include data quality and availability, choice of performance measures, formulation of statistical analyses, and development of approaches to reporting results of profiling analyses

Summary of Lecture 9

- Performance measures were estimated using a unifying statistical approach based on multi-level models

- Multi-level models:
  - take into account the hierarchical structure usually present in data for profiling analyses
  - permitting the impact of patient severity on outcome to vary by provider
  - adjusting for within-provider correlations
  - accounting for differential sample size across providers

Summary of Lecture 9

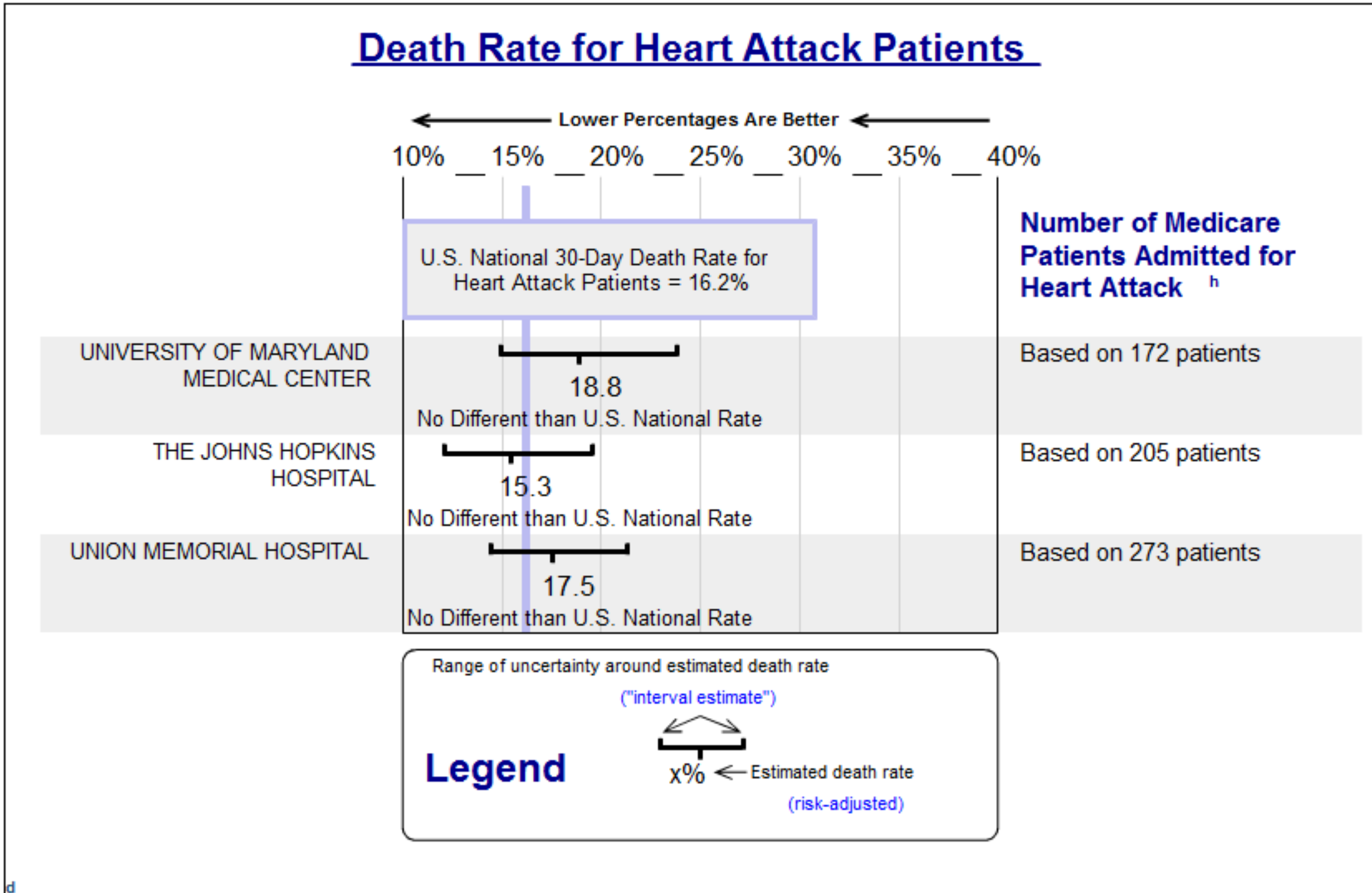
- The consideration of provider characteristics as possible covariates in the second level of the hierarchical model
  - May explain a large fraction of the natural variability in the observed data
  - Not clear whether or not to include these in national profiling systems (see White paper)
  - For sure you don't want to condition on the hospital profile, i.e. you want to compare urban and rural hospitals not just performance within urban or rural, so the "reference" should ignore hospital characteristics and be for the "typical" or "average" hospital.
  - Also, it is not clear the role played by some standard hospital characteristics (specifically: volume of procedure), this is likely in the causal pathway between patient characteristics and the outcome. Should it be included?

What is happening now?

- Hospital Compare
  - http://www.hospitalcompare.hhs.gov
  - http://www.hospitalcompare.hhs.gov/staticpages/for-professionals/ooc/calculation-of-30-day-risk.aspx
  - http://www.hospitalcompare.hhs.gov/staticpages/for-professionals/ooc/statistcal-methods.aspx


U.S. Department of Health & Human Services
HHS.gov
Sign In to MyMedicare.gov
Search Medicare.gov
Search
FAQ
Email | Print | Bookmark & Share | RSS | Español (Spanish) | A A A
Medicare.gov
Manage Your Health
Medicare Basics
Resource Locator
Help & Support
? Help
For Consumers
For Professionals
Medicare.gov » Hospital Compare Home
Hospital Compare
Where do you want to find a hospital?
Search Information
Location - ZIP Code or City, State
21211
e.g. 10009 or New York, NY
Search type [?]
General
Medical Conditions
Surgical Procedures
Find Hospitals
Hospital Spotlight
Medicare releases new data on Hospital Acquired Conditions. Click here for more information.
You can now visit Medicare's Hospital Value Based Purchasing Program page and learn more about potential future measures.
Additional Information
View a list of Hospital Compare Contacts
Download the Hospital Compare Database (Data Last Updated: April 11, 2011)

These percentages were calculated from Medicare data on patients discharged between July 01,2006 and June 30, 2009. They don't include people in Medicare Advantage Plans (like an HMO or PPO) or people who don't have Medicare.



HHS.gov

Close Window

Medicare.gov | Hospital Compare | Print

Information for Professionals

Value Based Purchasing

Process of Care

Outcome of Care Measure

Calculation of 30-Day Risk-Standardized Mortality Rates and Rates of Readmission

Data Collection Methods

Risk-Adjustment and Covariates Included in the Model

Statistical Methods Used to Calculate Rates

Mortality Measures

Hierarchical Regression Model

The statistical model for computing 30-day risk-adjusted mortality rate measures is a "hierarchical regression model." This type of model is based on the assumption that any heart attack or heart failure or pneumonia patients treated at a particular hospital will experience a level of quality of care that applies to all patients treated for the same condition in that hospital. In other words, the expected risk of death for two similar heart attack or heart failure or pneumonia patients treated in the same hospital would be more alike than the risk of death for the same two patients treated in two different hospitals. The likelihood that an individual patient will die is therefore a combination of:

- his or her individual risk characteristics (for example, gender, comorbidities, and past medical history) and
- the hospital's unique quality of care for all patients treated for that condition in that hospital.

The model estimates the effects of both of these components on mortality.

Calculating Mortality Rates

Each hospital's "30-day risk-adjusted mortality rate" (also called the "Risk Standardized Mortality Rate" or RSMR) is computed in several steps. First, the predicted 30-day mortality for a particular hospital obtained from the hierarchical regression model is divided by the expected mortality for that hospital, which is also obtained from the regression model. Predicted mortality is the rate of deaths from heart attack or heart failure or pneumonia that would be anticipated in the particular hospital during the 12-month period, given the patient case mix and the hospital's unique quality of care effect on mortality. Expected mortality is the rate of deaths from heart attack or heart failure or pneumonia that would be expected if the same patients with the same characteristics had instead been treated at an "average" hospital, given the "average" hospital's quality of care effect on mortality for patients with that condition. This ratio is then multiplied by the national unadjusted mortality rate for the condition for all hospitals to compute a "risk-adjusted mortality rate" for the hospital. So, the higher a hospital's predicted 30-day mortality rate, relative to expected mortality for the hospital's particular case mix of patients, the higher its adjusted mortality rate will be. Hospitals with better quality will have lower rates.

(Predicted 30-day mortality/Expected mortality) * U.S. National mortality rate = RSMR

For example, suppose the model predicts that 10 percent of Hospital A's heart attack patients would die within 30 days of admission in a given year, based on their ages, gender mix, and pre-existing health conditions, and based on the estimate of the hospital's specific quality of care. Then, suppose that the expected rate of 30-day deaths for those same patients were higher – say, 15 percent – if they had instead been treated at an "average" U.S. hospital. If the actual mortality rate for the 12-month period for all heart attack patients in all hospitals in the U.S. is 12 percent, then the hospital's risk-adjusted 30-day mortality rate would be 8 percent.

(10%/15%)* 12% = RSMR for Hospital A 8%

If, instead, 9 percent of these patients would be expected to have died if treated at the average hospital, then the hospital's mortality rate would be 13.3 percent.

(10%/9%)* 12% = RSMR for Hospital A 13.3%

In the first case, the hospital performed better than the average hospital and had a relatively low risk-adjusted mortality rate (8 percent); in the second case it performed worse and had a relatively high rate (13.3 percent).

What is happening now?

- CMS (formerly HCFA):
  - Updated recommendations for assessing quality:
    - **STATISTICAL ISSUES IN ASSESSING HOSPITAL PERFORMANCE (White Paper)**
    - Commissioned by the Committee of Presidents of Statistical Societies
  - Medicare Quality Improvement Organization (QIO) Program
    - See Annals of Internal Medicine 2006 paper.
  - Physician Quality Reporting System
    - http://www.cms.gov/pqrs/01_overview.asp

- Agency for Healthcare Research and Quality
  - Measuring Healthcare Quality
    - http://www.ahrq.gov/qual/measurix.htm
    - National Healthcare Quality report
      - http://www.ahrq.gov/qual/nhqr10/nhqr10.pdf
  - Trends in Hospital Risk-Adjusted Mortality for Select Diagnoses by Patient Subgroups, 2000–2007, Statistical Brief #98, Healthcare Cost and Utilization Project (HCUP) Statistical Briefs
    - http://www.ncbi.nlm.nih.gov/books/NBK52657/