Lecture 8
Applications of Multilevel Models to Profiling of Health Care Providers

Outline

- What is profiling?
  - Definitions
  - Statistical challenges
  - Centrality of multi-level analysis
- Fitting Multilevel Models with Winbugs:
  - A toy example on institutional ranking
- Profiling medical care providers: a case-study
  - Hierarchical logistic regression model
  - Performance measures
  - Comparison with standard approaches

What is profiling?

- Profiling is the process of comparing quality of care, use of services, and cost with normative or community standards
- Profiling analysis is developing and implementing performance indices to evaluate physicians, hospitals, and care-providing networks

Objectives of profiling

- Estimate provider-specific performance measures:
  - measures of utilization
  - patients outcomes
  - satisfaction of care
- Compare these estimates to a community or a normative standard

U.S.News Best Hospitals 2002

Evaluating hospital performance

- Health Care Financing Administration (HCFA) evaluated hospital performance in 1987 by comparing observed and expected mortality rates for Medicare patients
- Expected Mortality rates within each hospital were obtained by :
  - Estimating a patient-level model of mortality
  - Averaging the model-based probabilities of mortality for all patients within each hospital
- Hospitals with higher-than-expected mortality rates were flagged as institutions with potential quality problems

Statistical Challenges

- Hospital profiling needs to take into account
  - Patients characteristics
  - Hospital characteristics
  - Correlation between outcomes of patients within the same hospital
  - Number of patients in the hospital
- 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

Within cluster correlation

- Hospital practices may induce a strong correlation among patient outcomes within hospitals even after accounting for patients characteristics
- Extend standard regression models to multi-level models that take into account the clustered nature of the data

Health care quality data are multi-level!

- Data are clustered at multiple-levels
  - Patients clustered by providers, physicians, hospitals, HMOs
  - Providers clustered by health care systems, market areas, geographic areas
- Provider sizes may vary substantially
- Covariates at different levels of aggregation: patient-level, provider level
- Statistical uncertainty of performance estimates need to take into account:
  - Systematic and random variation
  - Provider-specific measures of utilization, costs

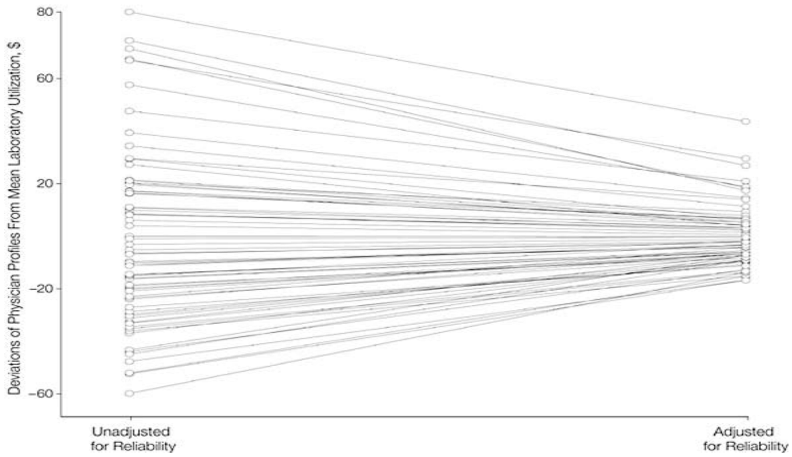
Sampling variability versus systematic variability

- "Sampling variability": statistical uncertainty of the hospital-specific performance measures
- "Systematic variability" : variability between hospitals performances that can be possibly explained by hospital-specific characteristics (aka "natural variability")
- Develop multi-level models that incorporate both patient-level and hospital-level characteristics

Borrowing strength

- Reliability of hospital-specific estimates:
  - *because of difference in hospital sample sizes, the precision of the hospital-specific estimates may vary greatly. 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

Each point represents the amount of laboratory costs of patients who have diabetes deviates from the mean of all physicians (in US dollars per patient per year). The lines illustrate what happens to each physician's profile when adjusted for reliability (Hofer et al JAMA 1999)

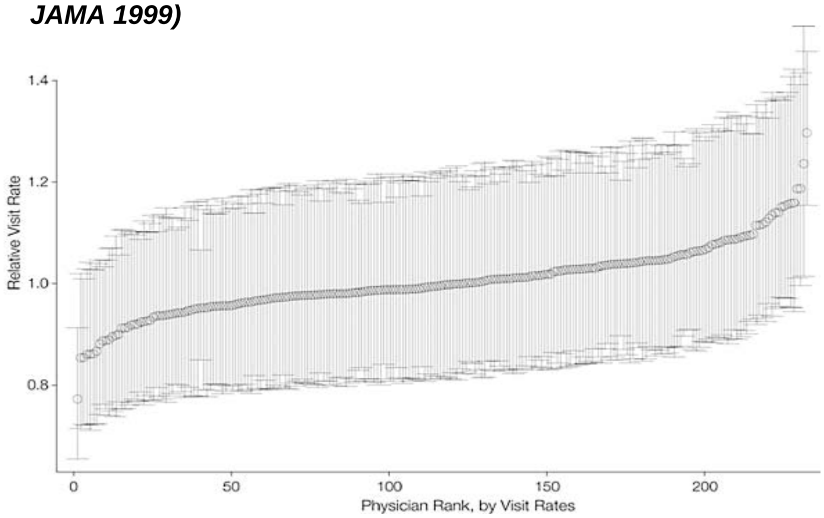


Adjusting Physician Laboratory Utilization Profiles for Reliability at the HMO Site

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 statistical significant (Hofer et al JAMA 1999)



Fitting Multilevel Models in Winbugs

Fitting Multi-Level Models

- SAS Proc Mixed
  - Maximum Likelihood Estimation (MLE)
  - Limitation: hard to estimate ranking probabilities and assess statistical uncertainty of hospital rankings
- BUGS and Bayesian Methods
  - Monte Carlo Markov Chains methods
  - Advantages: estimation of ranking probabilities and their confidence intervals is straightforward

Bayesian Inference and MCMC

- Goal: find $\hat{\eta}$ that maximizes the posterior distribution:

$$p(\boldsymbol{\eta} \mid \text{data}) = \frac{p(\boldsymbol{\eta})p(\text{data} \mid \boldsymbol{\eta})}{p(\text{data})}$$

  where

  $p(\eta)$ is the prior distribution

  $p(\text{data} \mid \eta)$ is the likelihood function

  $p(\text{data}) = \int p(\text{data} \mid \eta) \times p(\eta)$

- $p(\text{data})$ is hard to calculate when η is a vector of many parameters
- MCMC methods: are numerical approximation techniques that provide draws from the posterior distribution of the unknown parameters

$$\boldsymbol{\eta}^1, \ldots, \boldsymbol{\eta}^k, \ldots, \boldsymbol{\eta}^K \sim p(\boldsymbol{\eta} \mid \text{data})$$

- K is a large number ($K > 5000$) indicating the number of posterior samples

Estimating posterior mean and variance

- $\hat{\eta} = E[\eta \mid \text{data}]$ - easy to calculate by taking the average of the sampled values.
- $SE(\hat{\eta}) = \sqrt{V[\eta \mid \text{data}]}$ - easy to calculate by taking the variance of the sampled values
- In our case, WINBUGS will produce posterior samples of all parameters of interests.
- For example, let β_{0i}^k be the $k-th$ sample from $p(\beta_{0i} \mid \text{data})$, where β_{0i} is the hospital-specific log-odds ratio of death
- The posterior probability that β_{01} is larger than β_{02} (that is, hospital 1 is worse than hospital 2), can be easily estimated by counting how many times the posterior samples of β_{01} are larger than the posterior samples of β_{02}

$$\widehat{P}(\beta_{01} > \beta_{02} \mid \text{data}) = \frac{1}{K}\sum_{k=1}^{K} I(\beta_{01}^k > \beta_{02}^k)$$

Toy example on using WinBUGS for hospital performance ranking

This example considers mortality rates in 12 hospitals performing cardiac surgery in babies. The data are shown below.

Hospital	No of ops	No of deaths
A	47	0
B	148	18
C	119	8
D	810	46
E	211	8
F	196	13
G	148	9
H	215	31
I	207	14
J	97	8
K	256	29
L	360	24

A Multi-level model for hospital ranking

- Let r_i the number of deaths for hospital i
- let n_i the number of surgeries performed in hospital i

We assume

$$\begin{aligned} r_i &\sim \text{Bin}(p_i, n_i) \\ \text{logit} p_i &= \mu + b_i \\ b_i &\sim N(0, \sigma^2) \end{aligned}$$

where

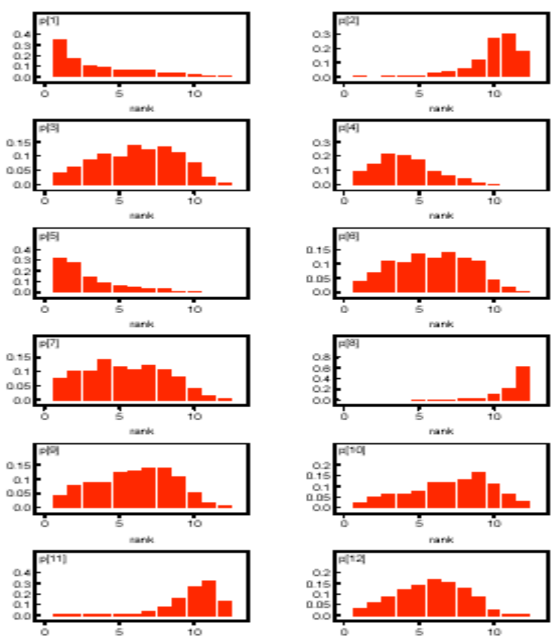
- p_i is the hospital-specific probability of death
- $\bar{p} = \exp(\mu)/(1 + \exp(\mu))$ is the probability of death for $b_i = 0$

Goal: identify the "aberrant" hospitals

```
model
{
        for( i in 1 : N ) {
            b[i] ~ dnorm(mu,tau)
            r[i] ~ dbin(p[i],n[i])
            logit(p[i]) <- b[i]
              }
          pop.mean <- exp(mu) / (1 + exp(mu))
          mu ~ dnorm(0.0,1.0E-6)
          sigma <- 1 / sqrt(tau)
          tau ~ dgamma(0.001,0.001)
}
```

node	mean	sd
p[1]	0.05357	0.01959
p[2]	0.1026	0.02203
p[3]	0.07102	0.01701
p[4]	0.05947	0.008078
p[5]	0.05252	0.01354
p[6]	0.06867	0.01401
p[7]	0.06796	0.01597
p[8]	0.1217	0.02196
p[9]	0.06943	0.01435
p[10]	0.07859	0.0193
p[11]	0.1019	0.01745
p[12]	0.06893	0.01185
pop.mean	0.07246	0.0105

Posterior distributions of the ranks – who is the worst?



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

Admission severity index (Normand et al 1997 JASA)

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
$(\text{Age—65})^2$	−.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

Hierarchical logistic regression model

I: *patient level, within-provider model*

– Patient-level logistic regression model with random intercept and random slope

- 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 ratio of death for hospital i having patients with severity equal to the average
- β_{1i} denotes the hospital-specific association between severity and logit of probability of death

Hospital-level model

- Without hospital-specific covariates

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

- With hospital-specific covariates

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

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

Normand et al JASA 1997

Table 1. Patient and Hospital Characteristics in the Study Cohort

	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

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

Hospital-Performance Measures

- 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
- 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:

$$\begin{aligned}
P_i^{A-S} &= P(\mu_i^A - \mu_i^S > \textbf{benchmark}), \\
&\quad \textbf{where} \\
\mu_i^A &= \frac{1}{n_i}\sum_{j=1}^{n_i} P(Y_{ij}=1 \mid \beta_{0i}, \beta_{1i}, \textbf{sev}) \\
&= \frac{1}{n_i}\sum_{j=1}^{n_i} \textbf{logit}^{-1}(\beta_{0i} + \beta_{1i}(\textbf{sev}_{ij} - \overline{\textbf{sev}})) \\
\mu_i^S &= \frac{1}{n_i}\sum_{j=1}^{n_i} P(Y_{ij}=1 \mid \gamma_0, \gamma_1, \textbf{sev}) \\
&= \frac{1}{n_i}\sum_{j=1}^{n_i} \textbf{logit}^{-1}(\gamma_0 + \gamma_1(\textbf{sev}_{ij} - \overline{\textbf{sev}}))
\end{aligned}$$

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

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

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:

- $x_{ij} = \text{sev}_{ij} - \overline{\text{sev}}$
- 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:

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

where:

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

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

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

where

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

- 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

Normand et al JASA 1997

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:
  - We found that the 2.5 and 97.5 percentiles of the log-odds of 30-day mortality ***for a patient with average admission severity*** is equal to **(-1.87,-1.56),** corresponding to **(0.13,0.17)** in the probability scale
- Non-Exchangeable model (with hospital covariates), random intercept and random slope:
  - We found that the 2.5 and 97.5 percentiles for 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

Effect of hospital characteristics on baseline log-odds of mortality

- Rural hospitals have higher odds ratio of mortality than urban hospitals for an average patient γ
- This is an indication of inter-hospital differences in the baseline mortality rates

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)

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

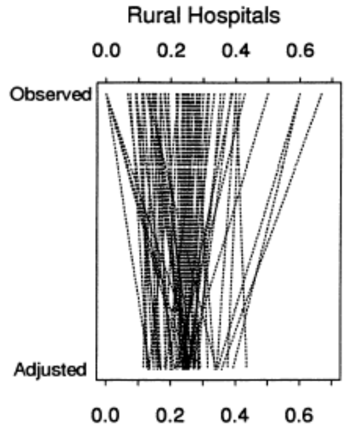
- The association between severity and mortality is ``modified'' by the size of the hospitals
- Medium-sized hospitals having smaller severity-mortality associations than large hospitals
- This indicates that the effect of clinical burden (patient severity) on mortality differs across hospitals

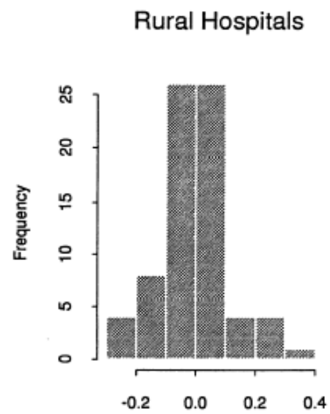
Estimates of II-stage regression coefficients (slopes)

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

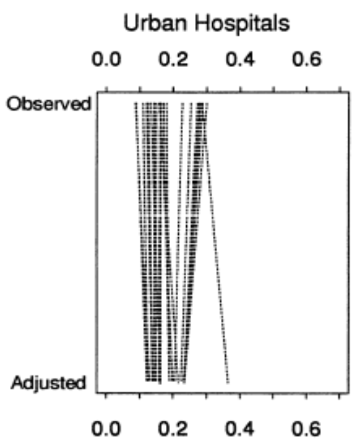
β_{1i}: Severity − $\overline{\text{severity}}$	γ_{10}: Intercept	1.22	.13	9.18	(.96, 1.52)
	γ_{11}: Rural	.05	.16	.33	(−.27, .36)
	γ_{12}: Nonacademic	−.11	.17	−.64	(−.44, .23)
	γ_{13}: Small	−.08	.20	−.39	(−.50, .28)
	γ_{14}: Medium	−.29	.15	−1.88	(−.58, .01)

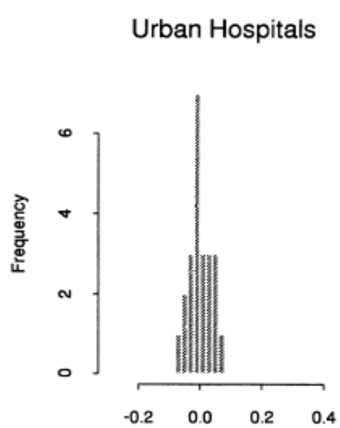
Observed and risk-adjusted hospital mortality rates: Crossover plo
Display the observed mortality rate (upper horizontal axis) and
Corresponding risk-adjusted mortality rates (lower horizontal line)
Histogram represents the difference = observed - adjusted





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			
						HCFA		$\hat{P}_i^{A-S}$		$\hat{P}_i^{A-S}$		$\hat{P}_i^{*}$	
Hospital	No. of AMI patients	No. dead	Hospital location	Academic (Y/N)	Hospital 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 ≥ 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.

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%

Discussion

- 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

Discussion

- 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
  - Provide a flexible framework for analyzing a variety of different types of response variables and for incorporating covariates at different levels of hierarchal structure

Discussion

- In addition, multi-level models can be used to address some key technical concerns in profiling analysis including:
  - permitting the impact of patient severity on outcome to vary by provider
  - adjusting for within-provider correlations
  - accounting for differential sample size across providers
- The multi-level regression framework permits risk adjustment using patient-level data and incorporation of provider characteristics into the analysis

Discussion

- The consideration of provider characteristics as possible covariates in the second level of the hierarchical model is dictated by the need to explain as large a fraction as possible of the variability in the observed data
- In this case, more accurate estimates of hospital-specific adjusted outcomes will be obtained with the inclusion of hospital specific characteristics into the model

Key words

- Profiling
- Case-mix adjustment
- Borrowing strength
- Hierarchical logistic regression model
- Bayesian estimation and Monte Carlo Markov Chain
- Ranking probabilities