No More “Excuses” for UNC:
A Hierarchical Model Evaluating Pediatric Heart Surgical Programs

Group: Yunran Chen, Yiwei Gong, Lingxi Song, Siqi Fu

Motivation

The New York Times exposed the **extremely high mortality rate** of surgical program of UNC Children's Hospital, **especially for complex surgeries**.

UNC argued **larger programs generally have lower mortality rate** and **judging based on mortality rate is unfair to small programs**.

- Does UNC program really behave badly?
- What about other programs?
- How can we evaluate the complex surgical programs in a scientific way based on these mortality rates?

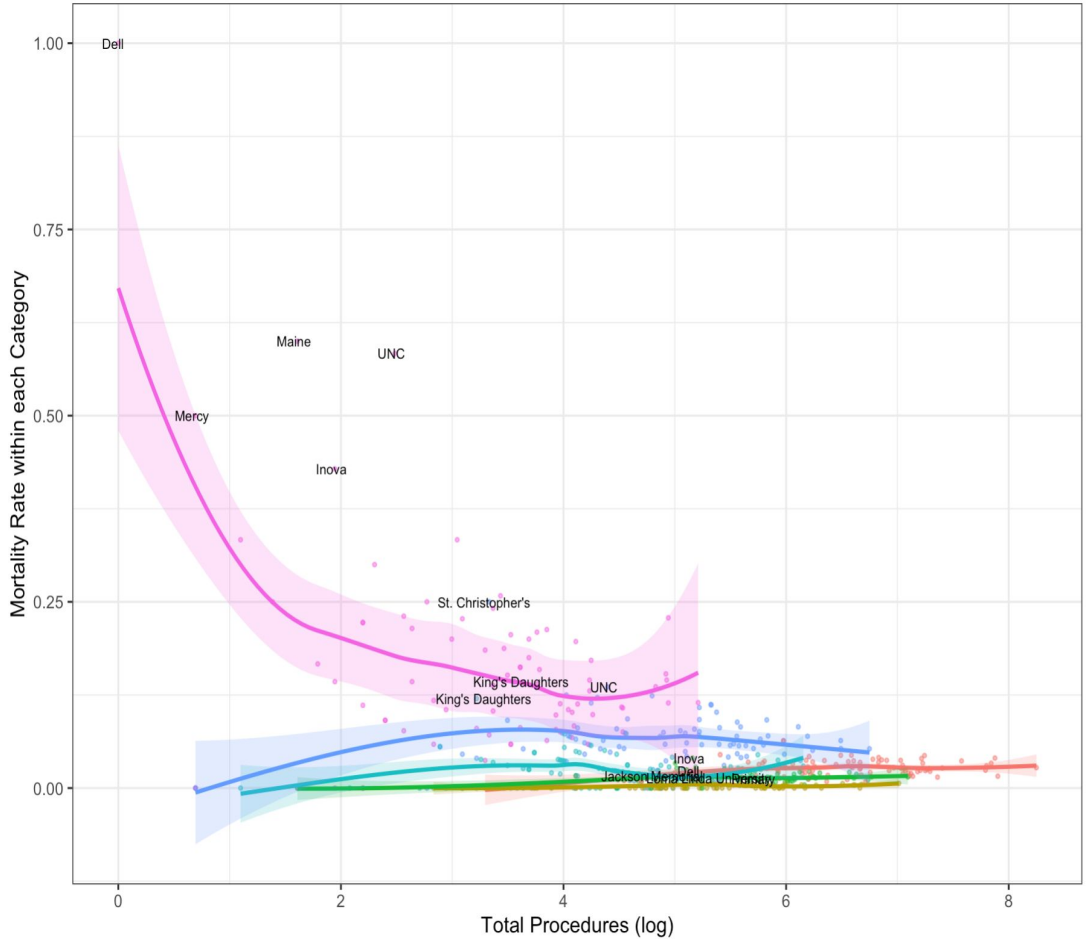
Introduction

- **Unavailable influencing factors**
- **Volume-based performance of each hospital**
  - Large variance of low-volume hospitals.
    - 64% hospitals do < 250 procedures
    - decrease the power of statistical results
  - Lower mortality rate of high-volume hospitals
    - more resources allocation
    - valuable experience of teams
    - high reputation attract patients)
- **Case-mix pattern of each hospital**
  - Distinct mortality rate of different procedure complexity
  - Proportion of procedure type varies across hospitals, related to volume

Goal

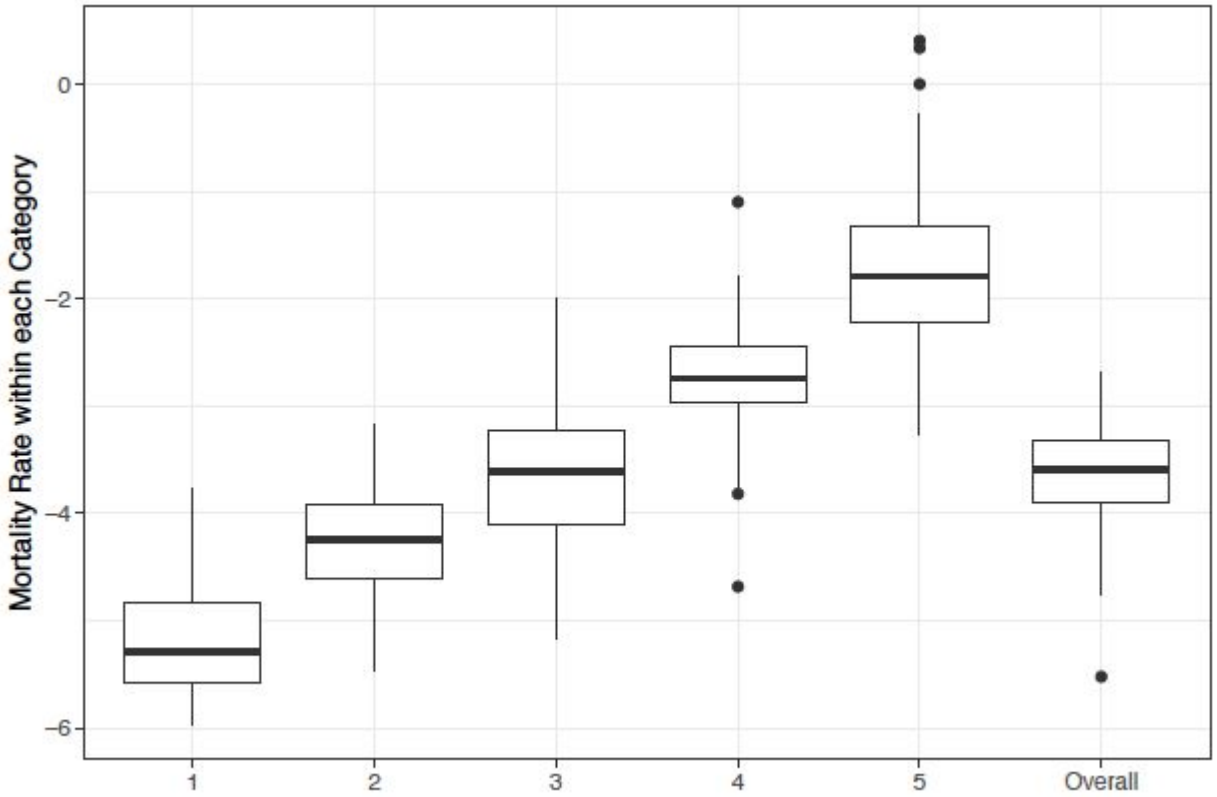
- **Build a hierarchical model** allowing information sharing, but share it **wisely**
- **Estimate** expected mortality rate **addressing the aforementioned characteristics**
- **Evaluate heart surgery program** based on O/E, benefiting patients' decision making and supervising the quality of programs
- **Evaluate** the performance of **UNC**

EDA: **Volume-based performance of each hospital**



- Mortality rate **decreases** significantly with **increase** of total procedures for **category 5.**
- **Variance decreases** with increase of total procedures.
- UNC has higher mortality rate especially for category 4 and 5.
- Need a **volume-based shrinkage**

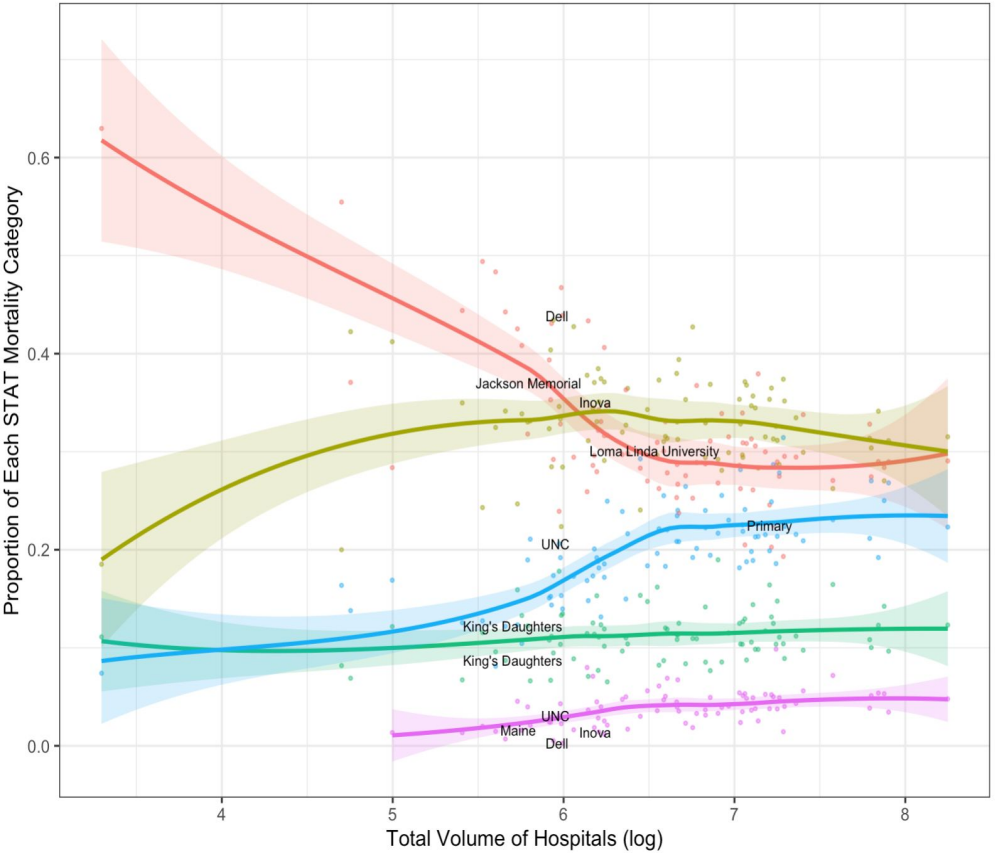
EDA: **Ordinal Effects of Procedure Complexity**



- More complex procedure has a **higher mortality rate** and a **longer tail**.
- **Need fix effects + heavy-tail prior**

-

EDA: Case mix of each hospital



- As the volume increases,
  - Category 1: a significant decrease
  - Category 2+4: a significant increase
- Need a shrinkage based on **volume of procedure type**

Model: Bayesian hierarchical model

$$logit(P(Y_{hi}=1)) = \beta_0 + \beta_1 I(Type2)_{hi} + \beta_2 I(Type3)_{hi} + \beta_3 I(Type4)_{hi} + \beta_4 I(Type5)_{hi}$$
$$+ \beta_5 \log(TypeVol)_h + b_{0h} + b_{1h} \log(TypeVol)_{hi}$$

i: individual procedure i

h: hospital h

$Typek$: STAT Mortality Category k

$Y = 1$: failure of procedure

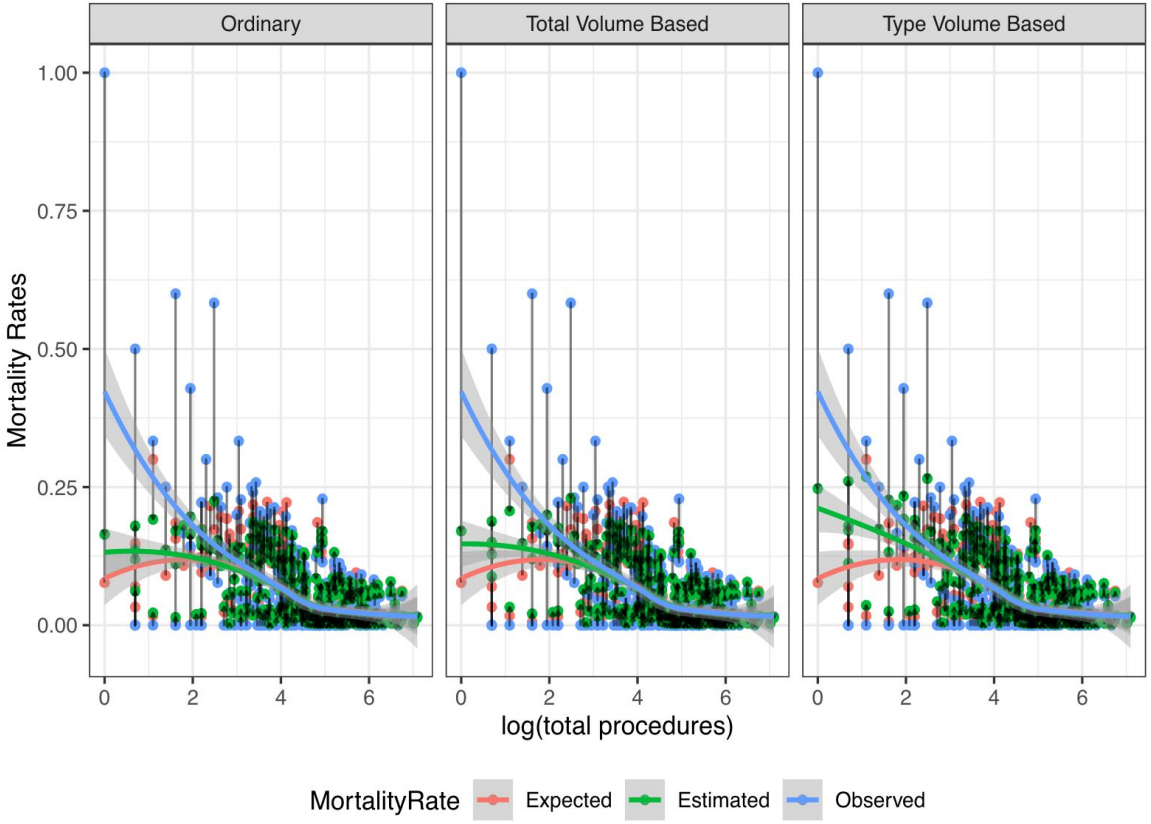
Priors:

$$\beta_0 \perp \beta_1 \perp \beta_2 \perp \beta_3 \perp \beta_4 \perp \beta_5 \sim T_3(0, 10)$$

$$\begin{pmatrix} b_{0h} \\ b_{1h} \end{pmatrix} \sim N(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{12} & \tau_{22} \end{pmatrix})$$

$$\tau_{11}, \tau_{22} \sim HC(0, 5)), \text{ and } \tau_{11}, \tau_{22} \sim LKJ(1)$$

Model Evaluation



- **Shrink** the observed mortality rates towards the expected mortality rates, **more shrinkage on low-volume hospitals**
- Our model shrinks towards more **reasonable direction**, **matching the fact** mortality rates decrease as volumes increase
- Out model performance slightly **better** in terms of **waic**. (O:1580>V:1578>TV:1570)

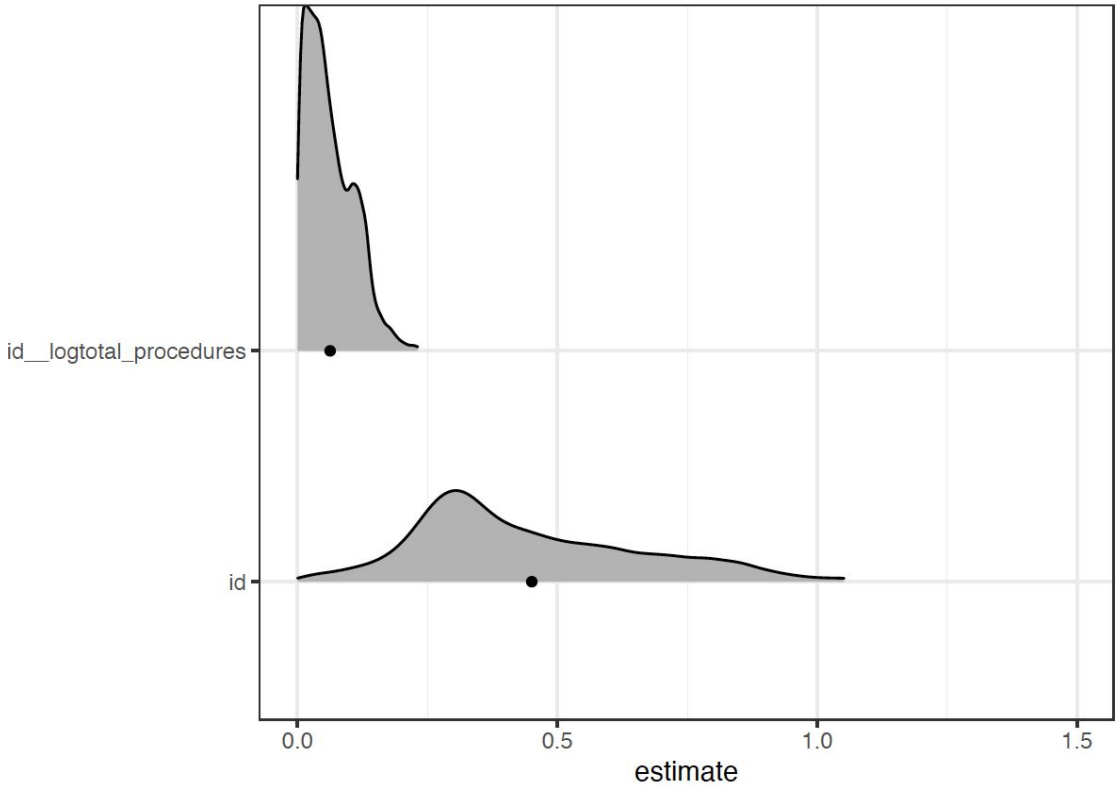
Estimation

Table 1: Point Estimate and CI for Random and Fixed Effect

	Estimate	Est.Error	l.95..CI	u.95..CI
Intercept	-4.87	0.31	-5.44	-4.25
procedure_typeSTATMortalityCategory2	1.46	0.12	1.22	1.70
procedure_typeSTATMortalityCategory3	1.76	0.14	1.47	2.03
procedure_typeSTATMortalityCategory4	2.93	0.12	2.69	3.15
procedure_typeSTATMortalityCategory5	3.56	0.16	3.25	3.85
logtotal_procedures	-0.15	0.05	-0.26	-0.05
sd(Intercept)	0.45	0.23	0.10	0.95
sd(logtotal_procedures)	0.06	0.05	0.00	0.17
cor(Intercept,logtotal_procedures)	-0.50	0.52	-0.97	0.82

- **All the terms included are significant.**
- **The odds of the mortality rate of type 2,3,4,5 procedure is 3.4 to 5.5, 4.4 to 7.6, 14.7 to 23.4 times of that of type 1 procedure.**
- **increasing # of a certain type of procedure by 10% will decrease the odds of mortality rate by 0.5% to 2.4%**

Estimation



- Clear heterogeneity across hospitals
- A large number of hospitals have higher mortality rate compared to a typical hospital
- Less heterogeneity across potential effect of type volume of hospitals -- may suggest similar learning rate from experience

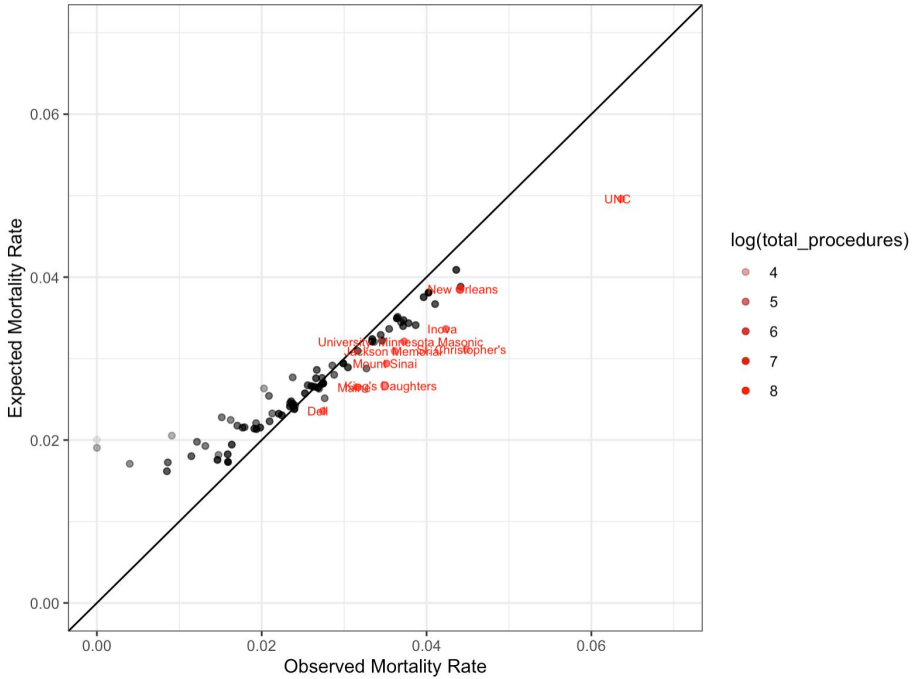
Inference

Table 2: Top Ten Best Hospital

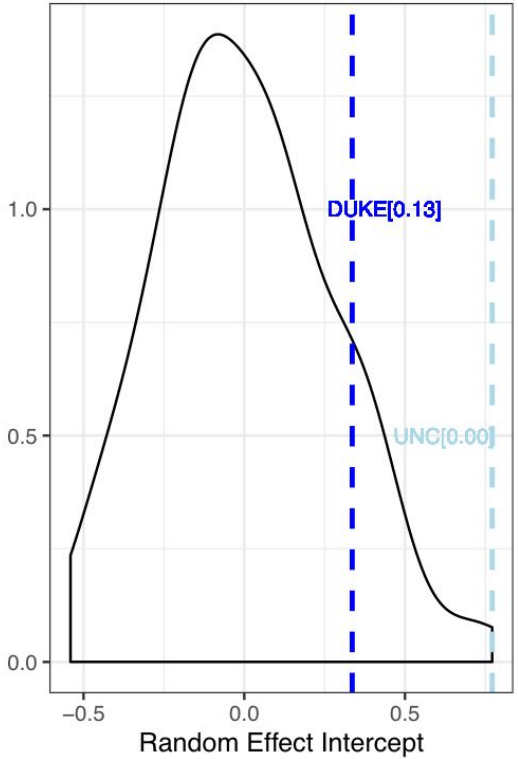
hospital_name	OE	OE_l	OE_u	star
Geisinger Medical Center	0.000	0.000	0.000	3
University of Kentucky Healthcare	0.000	0.000	0.000	3
Connecticut Children's Medical Center	0.233	0.144	0.450	3
Nemours Children's Hospital	0.442	0.260	0.940	3
Penn State Children's Hospital	0.499	0.334	0.846	3
UF Health Shands Children's Hospital	0.525	0.353	0.880	3
University of Maryland Children's Hospital	0.614	0.416	1.016	2
Helen DeVos Children's Hospital	0.636	0.433	1.014	2
University of Wisconsin Hospitals and Clinics	0.665	0.434	1.103	2
Mercy Medical Center	0.683	0.440	1.136	2

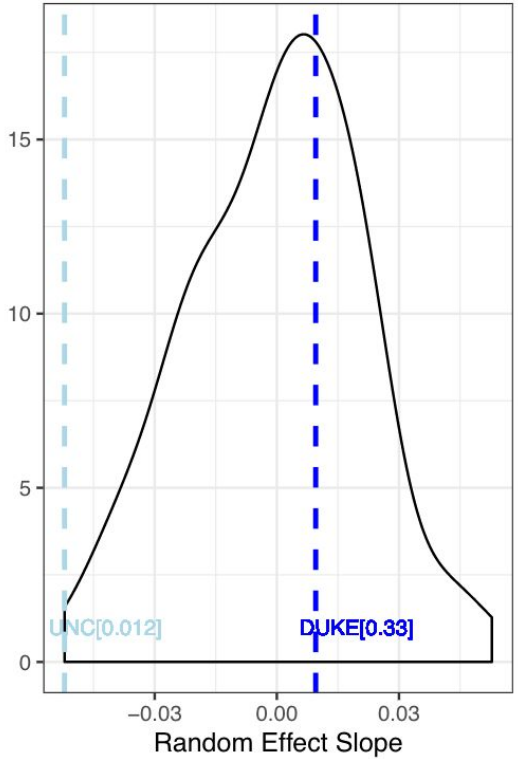
Table 3: Top Ten 'Killing' Hospital

hospital_name	OE	OE_l	OE_u	star
St. Christopher's Hospital for Children	1.440	0.923	2.422	2
Children's Hospital of the King's Daughters	1.305	0.794	2.165	2
UNC Children's Hospital	1.282	0.903	1.930	2
Inova Children's Hospital	1.259	0.872	1.916	2
Mount Sinai Hospital	1.196	0.832	1.810	2
Maine Medical Center	1.196	0.732	2.064	2
Jackson Memorial Hospital	1.169	0.786	1.842	2
Dell Children's Medical Center	1.164	0.777	1.878	2
University of Minnesota Masonic Children's Hospital	1.161	0.807	1.744	2
Children's Hospital New Orleans	1.142	0.832	1.609	2



UNC Program Evaluation: **No More "Excuses" for UNC**





- UNC is **ranked as bottom 3** based on O/E score
- The **probability** of UNC **ranked in the bottom 5** based on O/E scores is **0.86** (Duke is 0.23)
- **Extreme value** of random effect and random intercept
  - more likely to fail a procedure than a typical hospital
  - less likely to learn from experience than a typical hospital.

Conclusion

- **Excellent performance** on addressing the **volume-based shrinkage** and considering **case mix** problem.
- **Fits the data well** as shown by model diagnosis and comparison among alternative model.
- Valuable for benefiting patients' decision making and supervising the quality of each program.
- **UNC should suspend complex heart procedures**
- Suggest:
  - Merge competing programs in a nearby neighborhood
  - Enhance the patient referrals to nearby better programs