Multilevel models

CMED6040 – Session 5

Tim Tsang (matklab@hku.hk)

School of Public Health
The University of Hong Kong

13 June 2023

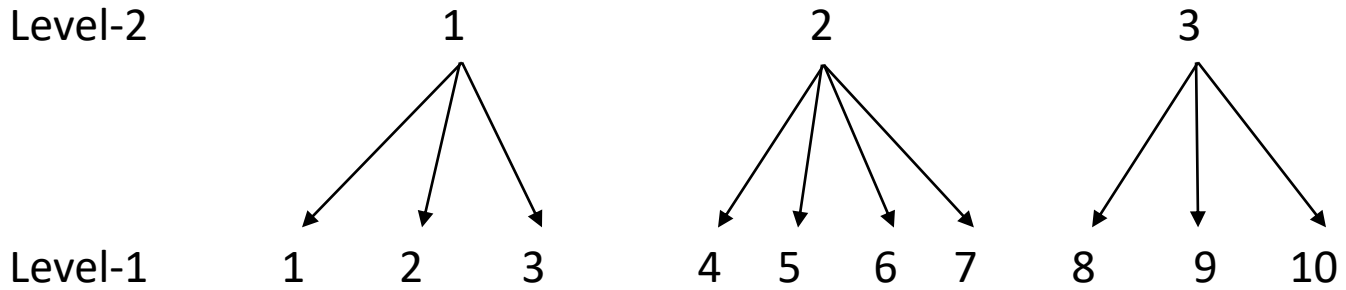
Session 5 learning objectives

After this session, students should be able to

- Recognize situations where multilevel models are needed to describe the data with a hierarchical structure
- Estimate and interpret effects under hierarchical linear model
- Perform model selection for hierarchical linear model

Multilevel models

- Multilevel models is a class of models which can handle grouped or clustered data
- Also called hierarchical models
- Handle data with nested structure
- Handle observations which are not independent (within group/cluster)



Examples of units for different levels

Level-1 unit	Level-2 unit
Students	Schools
Household members	Households
Families	Neighborhoods
Patients	Wards
Cities	Countries
Blood pressure	***Patients***

Different names of multilevel models

- Random effects (intercept/slope) model
- Mixed-effects model (fixed + random effects)
- Multilievel models (MLM)
- Hierarchical linear model (HLM)

Why multilevel models

- Research questions have a multi-level structure
  - e.g. household – household members, ward – patients
- Observations not independent, correlated within units
  - Violated linear regression assumptions
  - Biased estimate of standard error / uncertainty
  - Need to allow for dependence for correct inference
- Interested in both contextual and individual effects

Why multilevel models

If ignoring multi-level structures:

- Ecologic fallacy
  - Implications not always applicable to individual level
- Atomistic fallacy
  - May miss important group level effects
- Inflated type I error
  - Ordinary regression assumed N samples are all independent
  - However, “effective sample size” <N when observations are correlated
  - Leads to underestimation of the uncertainty or standard error (p-values too small)
- Loss of statistical power
  - Neglecting clustering effect in the intercept may reduce signal-to-noise ratio

Fixed vs random effects

- Fixed effects:
  - Interested in group-level estimates, especially when the number of groups is small
  - Need to adjust for the confounding group-level effects
  - Less efficient as more parameters are needed
  - Assumed conditional independence
- Random effects:
  - Groups are regarded as a random sample from a greater population, especially when the number of groups is large
  - More efficient but vulnerable to bias. The random effects need to be uncorrelated to the other covariates of the model.

Hierarchical linear model

- Dependent variable – variable at the lowest level
- Explanatory variables can be from any level
- If Y_{ij} represent the dependent variable of individual *i* in group *j*

$$Y_{ij} = \beta_{0j} + \varepsilon_{ij} \quad \text{(intercept only)}$$

where β has its random part:

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$\varepsilon_{ij} \sim N(0, \sigma^2), U_{0j} \sim N(0, \tau_0^2)$, ε_{ij}, U_{0j} independent

- Note that the statistical parameter for the random effect is τ_0^2

Hierarchical linear model

- The model can be rewritten as

$$Y_{ij} = \gamma_{00} + U_{0j} + \varepsilon_{ij}$$

$\varepsilon_{ij} \sim N(0, \sigma^2), U_{0j} \sim N(0, \tau_0^2)$

- Then the variance can be decomposed into:

$$\mathrm{var}(Y_{ij}) = \mathrm{var}(U_{0j}) + \mathrm{var}(\varepsilon_{ij}) = \tau_0^2 + \sigma^2$$

- Covariance between individuals in the same group:

$$\mathrm{cov}\left(Y_{ij}, Y_{i'j}\right) = \mathrm{cov}\left(U_{0j}, U_{0j}\right) + \mathrm{cov}\left(\varepsilon_{ij}, \varepsilon_{i'j}\right) = \tau_0^2 + 0 = \tau_0^2$$

- The (intraclass) correlation is therefore

$$\rho(Y_{ij}, Y_{i'j}) = \mathrm{cov}(Y_{ij}, Y_{i'j}) / \sqrt{\mathrm{var}(Y_{ij})\mathrm{var}(Y_{i'j})} = \frac{\tau_0^2}{\tau_0^2 + \sigma^2}$$

Hierarchical linear model

- HLM with random effect predictors (X_{ij})
- If Y_{ij} represent the dependent variable of individual *i* in group *j*

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + \varepsilon_{ij}$$ (intercept + 1 predictor)

where the βs have their random part:

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

$$\beta_{1j} = \gamma_{10} + U_{1j}$$

$$\mathrm{var}(\varepsilon_{ij}) = \sigma^2, \mathrm{var}(U_{ij}) = \tau_i^2$$

Hierarchical linear model

- The model can be rewritten as

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{10}X_{ij}}_{\text{Fixed effects}} + \underbrace{U_{0j} + U_{1j}X_{ij} + \varepsilon_{ij}}_{\text{random effects}}$$

$$\varepsilon_{ij} \sim N(0, \sigma^2), U_{ij} \sim N(0, \tau_i^2)$$

- Also called random intercept and slope model
- This can be generalized to multiple predictors

Estimation procedure

- Need to estimate γ, σ^2 and τ_i^2
- Maximum likelihood (ML)
  - $\hat{\gamma}$ as a function of σ^2 and τ_i^2
  - Numerical maximization is needed
  - Used when likelihood ratio test is carried out
- Restricted maximum likelihood (REML)
  - ML estimates for the variances are biased (underestimated)
  - REML corrects for the loss in degrees of freedom (using n – p as denominator) and produces unbiased variance estimates
  - Preferred when sample size is small

Intraclass correlation (ICC)

- Measure the proportion of variance explained by the between group variance

$$\rho = \frac{s_b^2}{s_b^2 + s_w^2}$$

where s_b^2, s_w^2 are the between and within clusters variance respectively

- If $\rho = 1$, $s_w^2 = 0$, all units within clusters are identical
- If $\rho = 0$, $s_b^2 = 0$, groupings not informative
- In general, large ρ → higher variability (lower similarity) across clusters, lower variability (higher similarity) within clusters
- Design effect = 1 + $\rho(m-1)$, m = mean group/cluster size

Design effect: variance relative to that obtained from simple random sampling

Simple example – ICC

```
require(ICC)
n.group <- 10
y <- 1:1000

group <- rep(1:n.group, each=length(y)/n.group)
rand.y <- sample(y, length(y), replace=F)
```

- groupings by sorted values:

```
ICCest(as.factor(group), y)
```

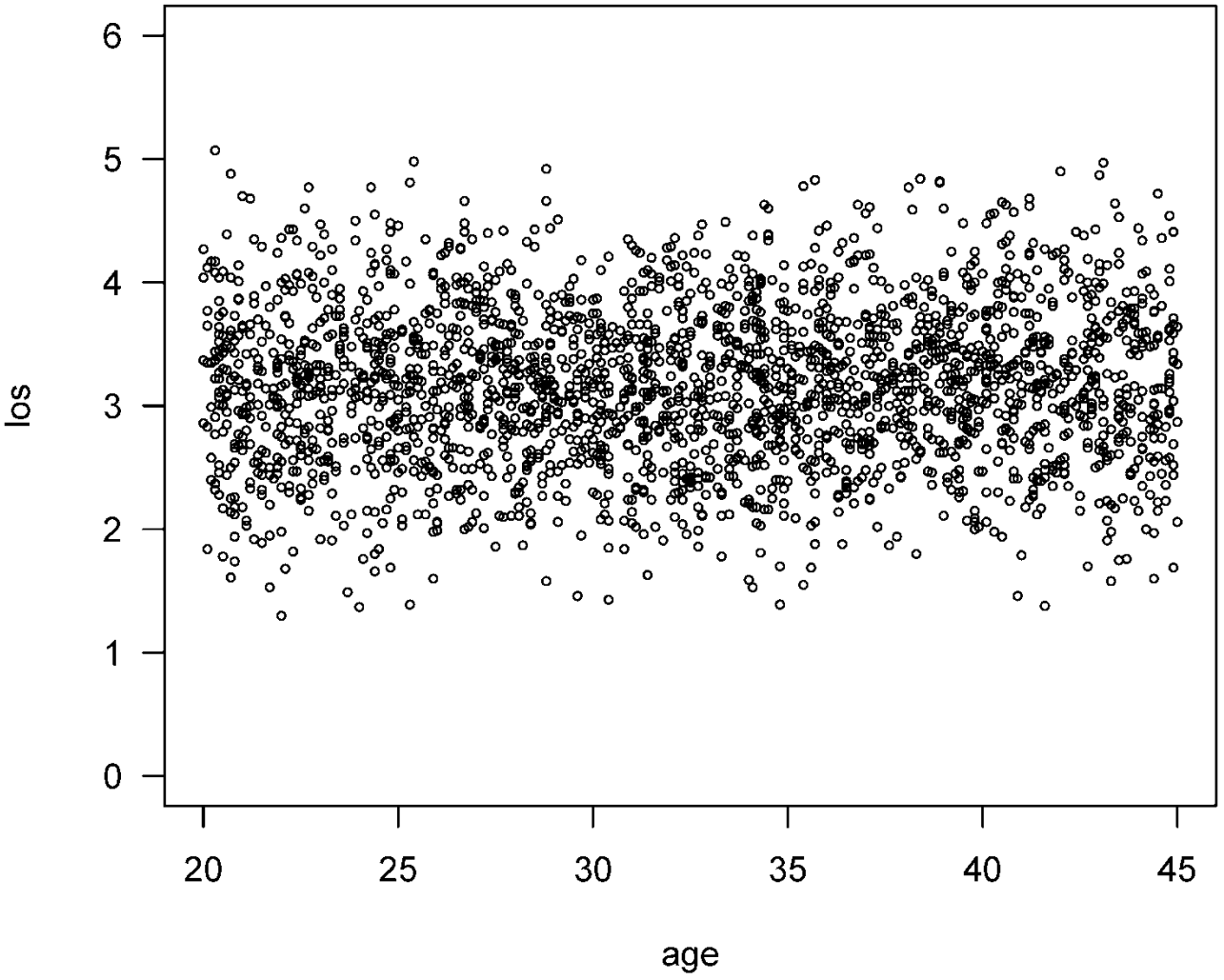
- randomly assigned groupings:

```
ICCest(as.factor(group), rand.y)
```

Example – Maternity length of stay

- Suppose we are interested in the association between age and maternity length of stay (LOS) for mothers after cesarean-section delivery
- There may also be hospital-level factors (e.g. hospital size) affecting mlos
- Data were collected from 200 mothers from each of the 12 hospitals
- Age: 20-45y
- Hospital size: small, medium, large
- Data saved in examplemlos.csv

Maternity LOS – age effect



- Any association between age and los?

```
plot(mlos$age,mlos$los,ylim=c(0,6))
```

Ordinary linear regression – age effect

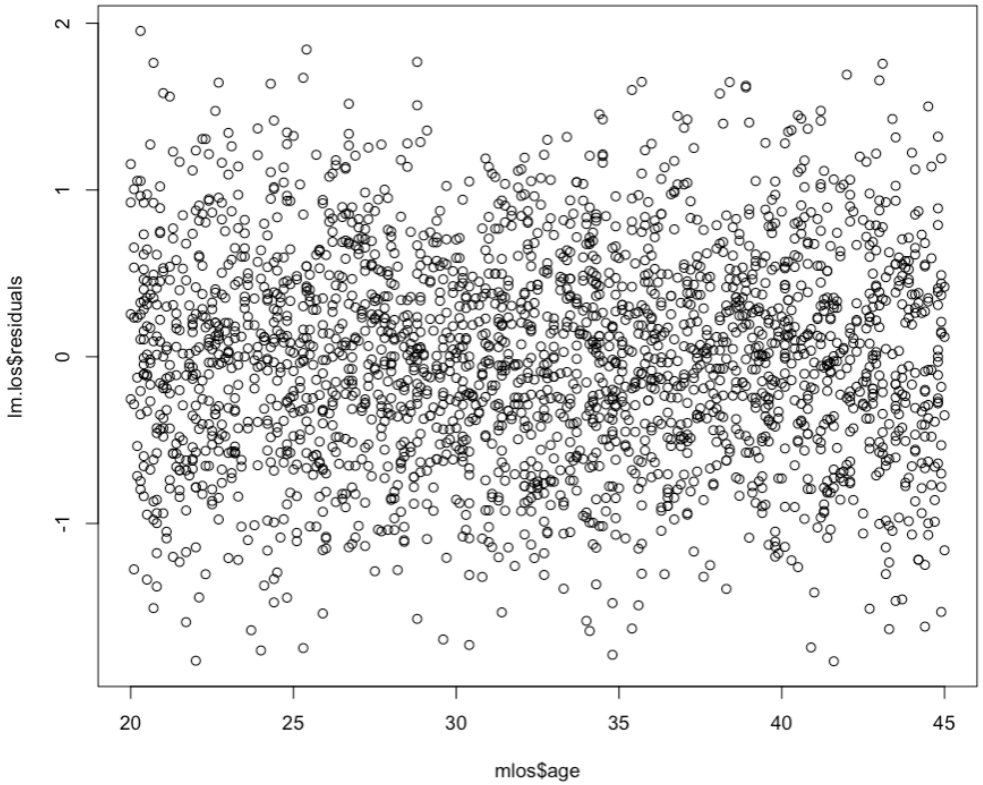
- Fit a linear regression for maternity LOS on age

```
mlos <- read.csv('examplemlos.csv')
lm.los <- lm(los~age, data=mlos)
summary(lm.los)
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.029551   0.060594  49.997   <2e-16 ***
age         0.004254   0.001824   2.332   0.0198 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Age is positively associated with maternity LOS from the linear regression model without consider potential within-hospital effect

Residual plots

- Perform some model diagnostics (e.g. residual plot) for the linear regression model

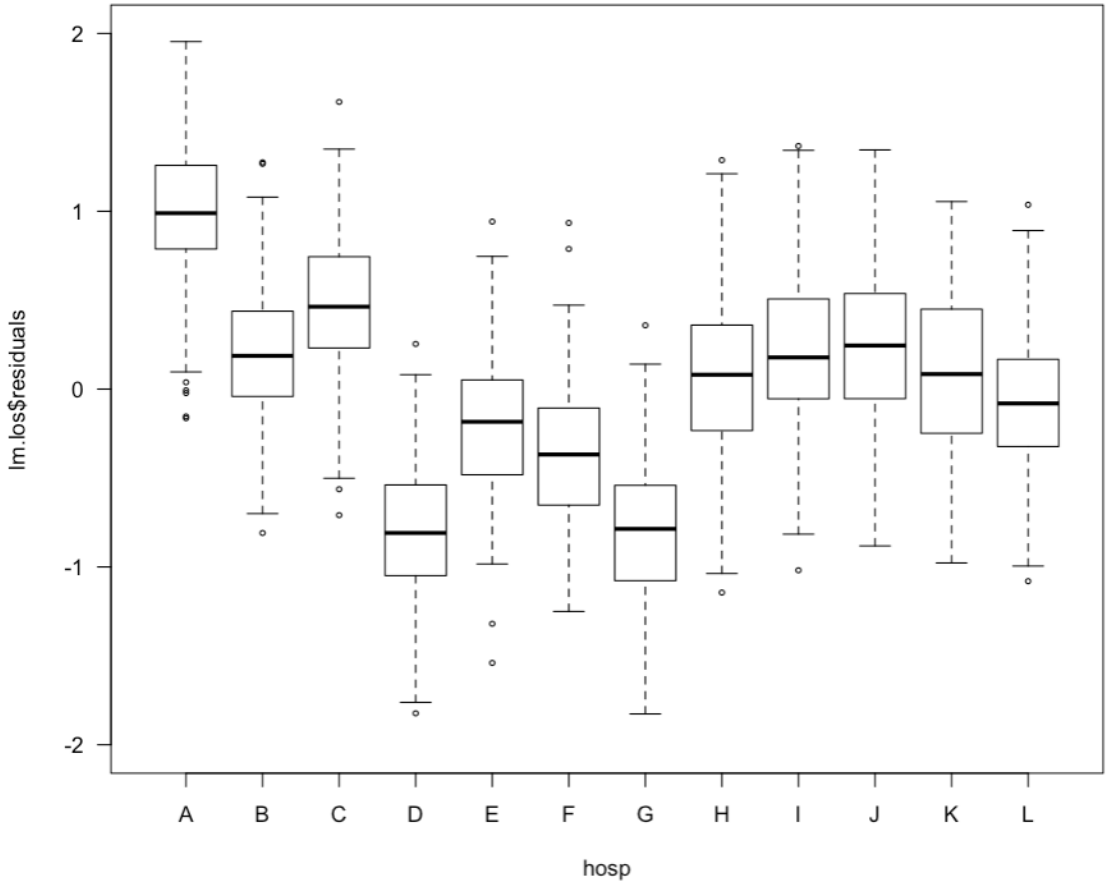


- No obvious pattern on age
- Centre around 0

```
plot(mlos$age,lm.los$residuals)
```

Residual plots

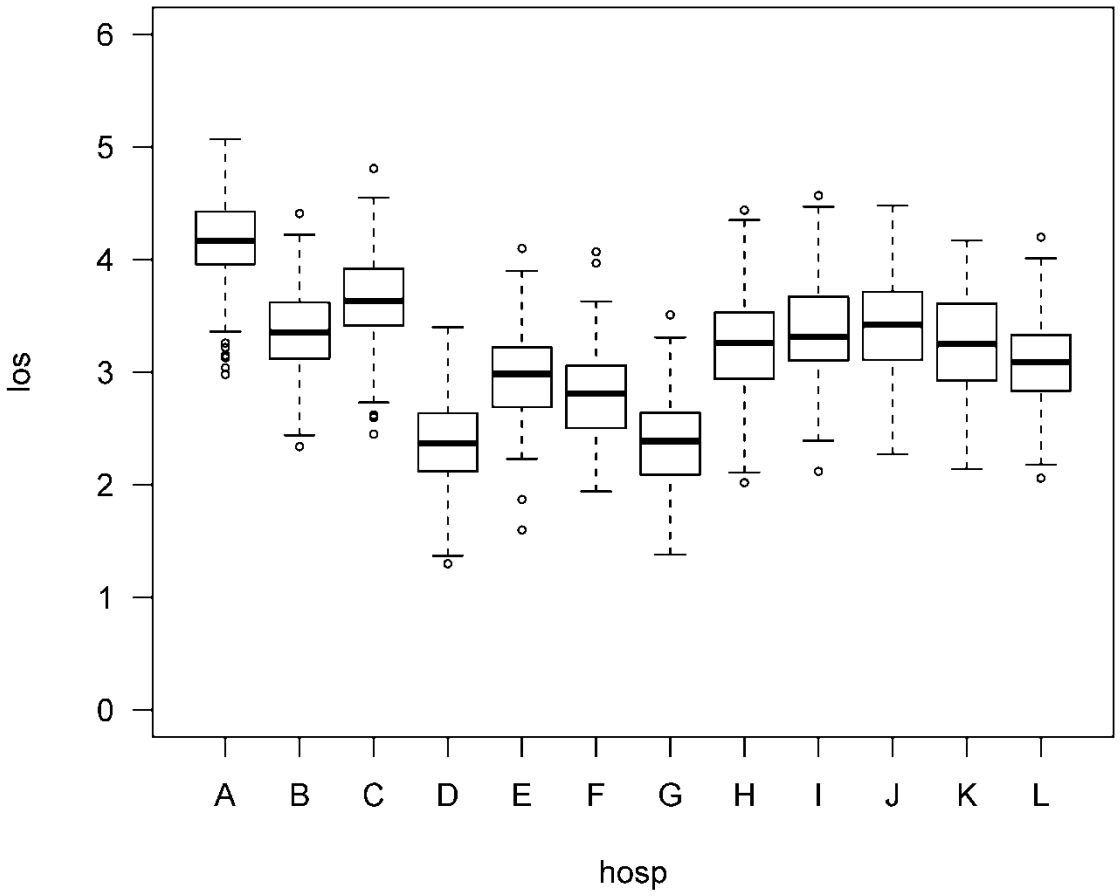
- Perform some model diagnostics (e.g. residual plot) for the linear regression model



- Obvious clustering of residuals by hospitals
- Violated the independence assumption

```
with(mlos,
plot(lm.los$residuals~
hosp, ylim=c(-2,2),
cex=0.5, las=1))
```

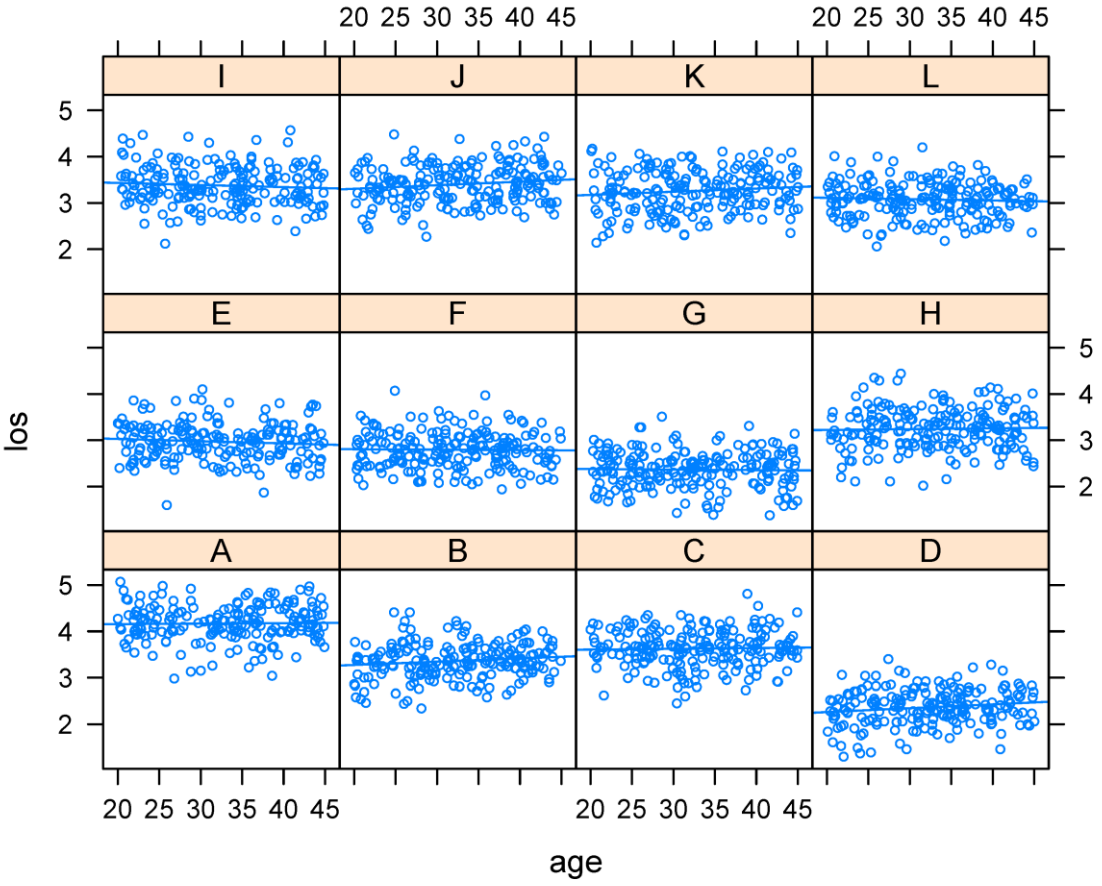
Variation of maternity LOS between hospitals




```
with(mlos, plot(los~hosp,
ylim=c(0,6), cex=0.5,
las=1))
```

- Large variation in maternity LOS across hospitals
- Multilevel model should describe the data better

xyplot() to show data



- In package “lattice”

```
require(lattice)
xyplot(los~age|hosp,
data=mlos, type=c('p','r'))
```

- ‘p’ for points
- ‘r’ for regression line

Linear mixed model (LMM) / HLM in R

- Package: lme4
- lmer(*formula*, *REML=TRUE, offset, subset, data*)
  - *formula:* fixed effects specification similar to lm(), random effects: (*RE_var*|*group*)
  - *REML=TRUE* for restricted maximum likelihood, otherwise maximum likelihood method will be used for the estimation
  - *offset, subset, data* as usual
- Can include random intercept, random slope, group level predictors, etc

- glmer() for generalized linear mixed-effects model (GLMM)

Formula specification in lmer()

Formula	Alternative	Meaning
`(1|g)`	`1 + (1|g)`	Random intercept with fixed mean
`0 + offset(o) + (1|g)`	`-1 + offset(o) + (1|g)`	Random intercept with a priori means
`(1|g1) + (1|g2)`	`1 + (1|g1) + (1|g2)`	Intercept varying among g1 and g2
`x + (x|g2)`	`1 + x + (1+x|g2)`	Correlated random intercept and slope

(Bates et al., J Stat Softw 2015)

Maternity LOS – random intercept model

- Suppose we fit a random intercept model for maternity LOS (did not include age)

```
require(lme4)
hlm.los1 <- lmer(los ~ 1 + (1|hosp), data=mlos)
summary(hlm.los1)

Fixed effects:
            Estimate Std. Error t value
(Intercept)   3.1676     0.1468   21.58
```

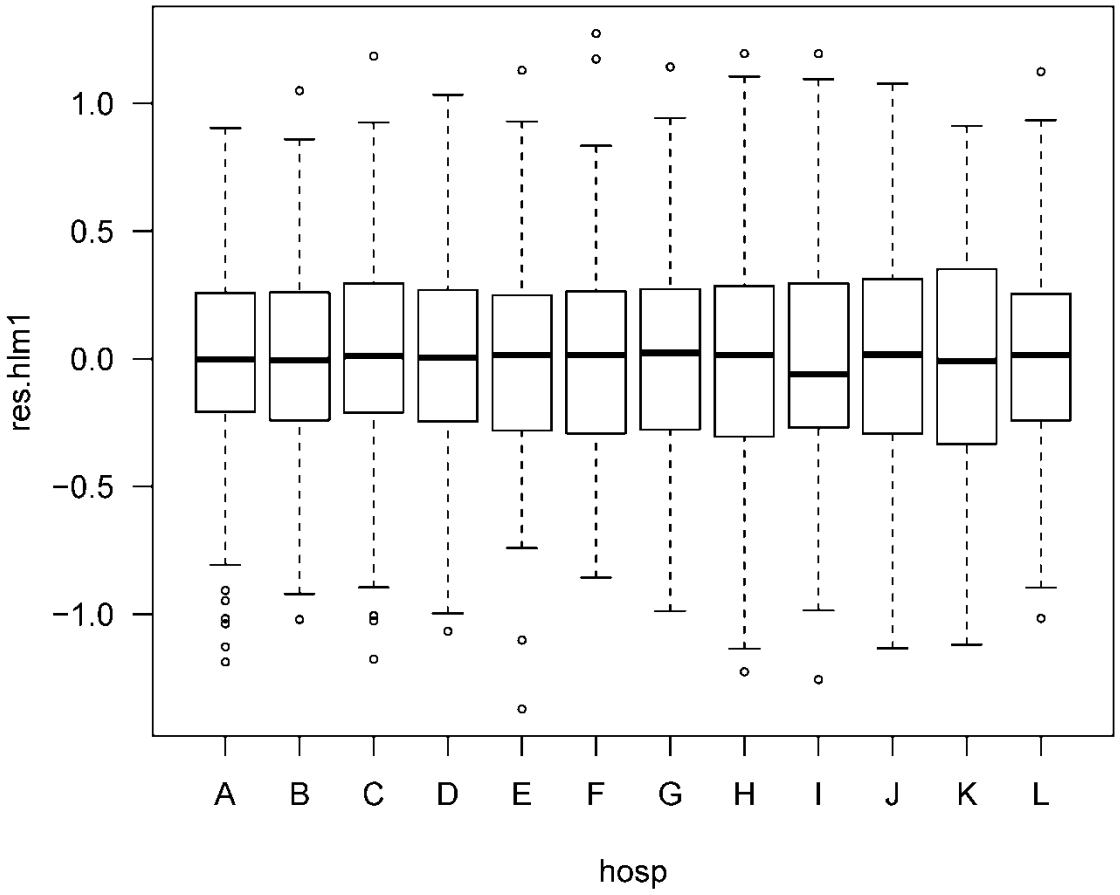
- The estimated overall mean of maternity LOS is 3.2 days

Maternity LOS – random intercept model

```
Random effects:
 Groups   Name        Variance Std.Dev.
 hosp     (Intercept) 0.2576   0.5076
 Residual             0.1659   0.4073
Number of obs: 2400, groups:  hosp, 12
```

- ICC = 0.2576 / (0.2576+0.1659) = 0.61
- Large variation between hospitals

Residual plot



- No obvious clustering of residuals by hospitals

```
with(mlos,
plot(summary(hlm.los
1)$residuals ~ hosp,
ylim=c(-3,3),
cex=0.5, las=1))
```

Maternity LOS – random intercept model with age

- Suppose we fit a random intercept model for maternity LOS and include age as a fixed effect

```
hlm.los2 <- lmer(los ~ age + (1|hosp), data=mlos)
summary(hlm.los2)

Fixed effects:
            Estimate Std. Error t value
(Intercept) 3.115889   0.151502  20.567
age         0.001594   0.001174   1.357
```

- The estimated age effect on maternity LOS is small

Maternity LOS – random intercept model with age

```
Random effects:
 Groups   Name        Variance Std.Dev.
 hosp     (Intercept) 0.2572   0.5071
 Residual             0.1659   0.4073
Number of obs: 2400, groups:  hosp, 12
```

- No change in the variance estimates for random effects
- Low explanatory power by age

Testing of the fixed effects

- t-test can be used
  - Need to calculate the appropriate degree of freedom
- Likelihood ratio test (preferred)
  - Need to be nested models
  - Use likelihood from maximum likelihood estimation
- Simulation

t-test for the fixed effect

- Available in package “lmerTest” (will replace lmer() in the package lme4)

```
require(lmerTest)
hlm.los2 <- lmer(los ~ age + (1|hosp), data=mlos)
summary(hlm.los2)
Fixed effects:
             Estimate Std. Error         df t value Pr(>|t|)
(Intercept) 3.116e+00  1.515e-01 1.250e+01  20.567 4.92e-11 ***
age         1.594e-03  1.174e-03 2.387e+03   1.357    0.175
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Age is not significant
- unload the package: `detach("package:lmerTest", unload=TRUE)`

LRT for the fixed effect

- Recall that the likelihood ratio statistics

$$2|\log l(\theta_0) - \log l(\theta_1)| \sim \chi_p^2$$

  where θ_0, θ_1 are the estimated parameter for the restricted and full models respectively, *p* is the difference in the number of parameters in the models
- Need to use maximum likelihood estimation (REML=FALSE in lmer())
- Can use anova() to perform LRT

LRT for the fixed effect

- H_0: γ_{10}= 0
- Fit the restricted and full model using ML

```
hlm.los2.ml <- lmer(los ~ age + (1|hosp), data=mlos, REML=FALSE)
hlm.los20.ml <- lmer(los ~ 1 + (1|hosp), data=mlos, REML=FALSE)
anova(hlm.los20.ml, hlm.los2.ml)

object: los ~ 1 + (1 | hosp)
..1: los ~ age + (1 | hosp)
       Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
object  3 2573.8 2591.2 -1283.9   2567.8
..1     4 2574.0 2597.1 -1283.0   2566.0 1.8445      1     0.1744
```

- Age is not significant

LRT for the random intercept effect

- H_0: $\tau_i^2 = 0$ vs H_1: $\tau_i^2 > 0$
- Hint: use logLik() to obtain the log-likelihood

```
lm.los20 <- lm(los ~ age, data=mlos)
lrt <- as.numeric(2*abs(logLik(lm.los20)-logLik(hlm.los2.ml)))
pchisq(lrt, df=1, lower.tail=F)
```

- p-value ≈ 0, rejected the null hypothesis
- Random intercept model is selected

Maternity LOS – random intercept and slope model

```
hlm.los3 <- lmer(los ~ age + (age|hosp), data=mlos, REML=FALSE)
summary(hlm.los3)

Fixed effects:
             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  3.115048   0.164367  8.185676  18.952 4.73e-08 ***
age          0.001616   0.001340 11.444577   1.207    0.252
```

- Age effect is not statistically significant at 5% significance level

Maternity LOS – effect of hospital size

```
hlm.los4 <- lmer(los ~ age + size + (age|hosp), data=mlos,
REML=FALSE)
summary(hlm.los4)
Fixed effects:
             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  3.444049   0.241827 10.431260  14.242  3.6e-08 ***
age          0.001631   0.001381 10.291301   1.181   0.2642
sizemedium  -0.777204   0.311014  9.647111  -2.499   0.0323 *
sizesmall   -0.210188   0.311018  9.646373  -0.676   0.5150
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

- Medium sized hospital associated with shorter maternity LOS

Maternity LOS – effect of hospital size

```
anova(hlm.los4, hlm.los3)
Models:
hlm.los3: los ~ age + (age | hosp)
hlm.los4: los ~ age + size + (age | hosp)
          Df    AIC    BIC  logLik deviance  Chisq Chi Df Pr(>Chisq)
hlm.los3   6 2577.9 2612.6 -1283.0   2565.9
hlm.los4   8 2577.9 2624.2 -1280.9   2561.9 4.0494      2      0.132
```

- Effect of hospital size is not significant
- Note that the df for the LRT is 2
  - 2 parameters used for the 3 levels of hospital size

“Estimated” coefficients

- It is possible to predict the β
  - β is random
- Posterior mode can be obtained using coef()

```
coef(hlm.los2)

$hosp

  (Intercept)         age
A    4.114082 0.001593742
B    3.308768 0.001593742
C    3.572766 0.001593742
D    2.315236 0.001593742
E    2.919731 0.001593742
```

…

Model comparison

- No straightforward statistics comparable to R^2 in linear regression
  - difficult to interpret with random effects
- May use AIC, BIC, deviance, etc
- Only valid for models using ML (REML=FALSE)

```
AIC(hlm.los1.ml, hlm.los2.ml, hlm.los3, hlm.los4)
             df      AIC
hlm.los1.ml   3 2573.852
hlm.los2.ml   4 2574.008
hlm.los3      6 2577.936
hlm.los4      8 2577.887
```

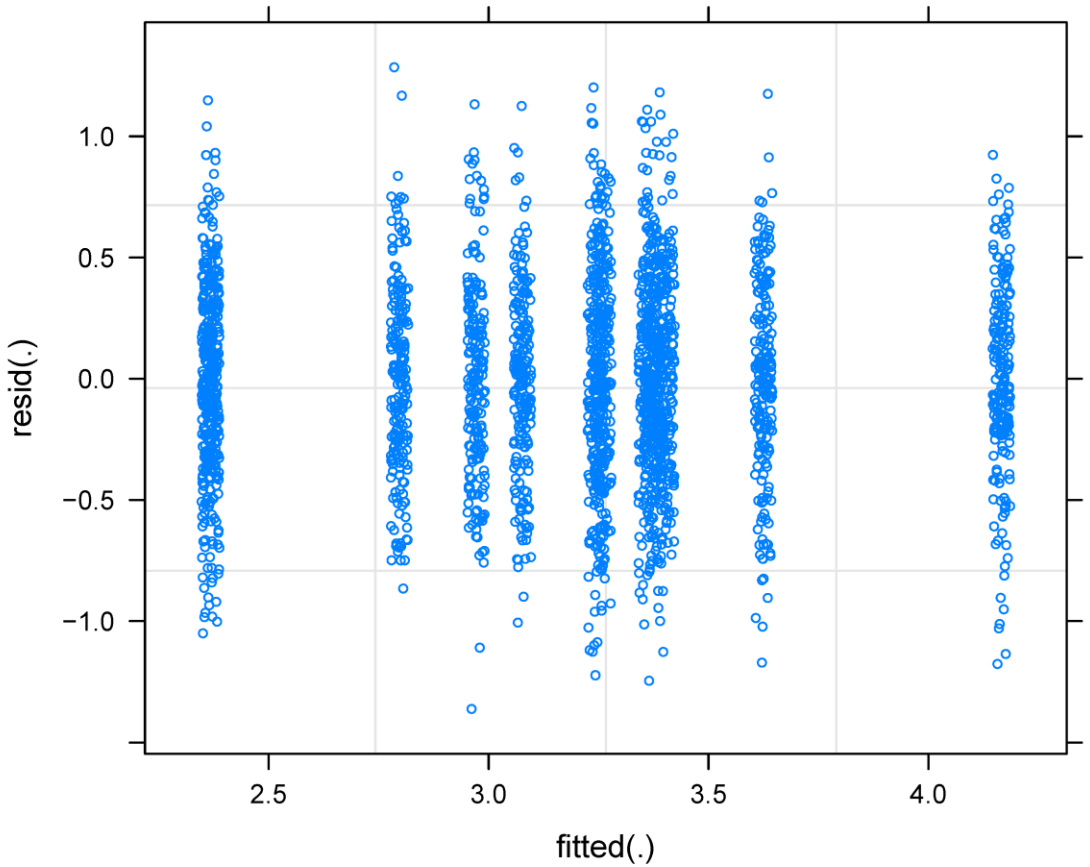
- Model with random intercept and is selected

Model diagnostics

- Compare fitted and observed values
- Residual plot:
  - Residuals by predictors (to assess incorrect specification)
  - Residuals by predicted value (to assess potential heteroskedasticity)
- qqplot for normality of the random effects
- qqmath()
- Level-2 (group) diagnostics not covered

Residuals by fitted values

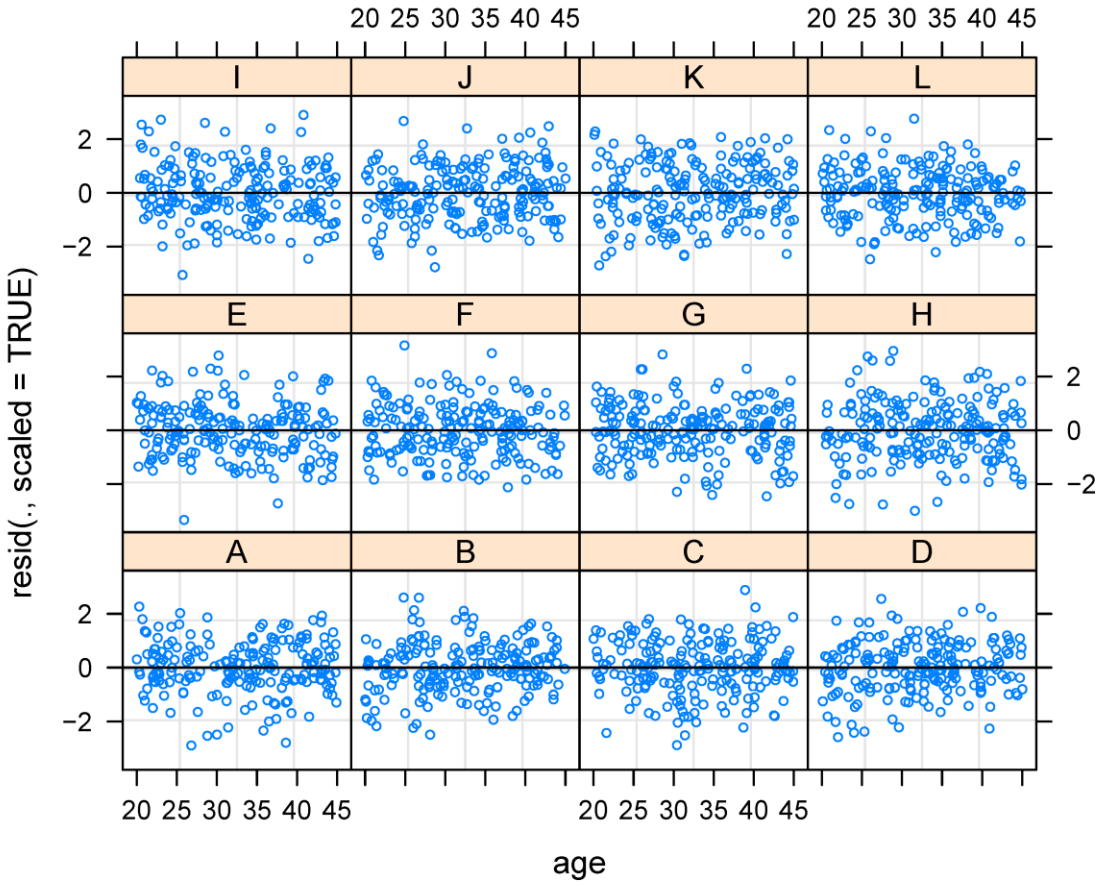
```
plot(hlm.los2, resid(.)~fitted(.))
```



- No obvious heteroskedasticity

Residuals by age

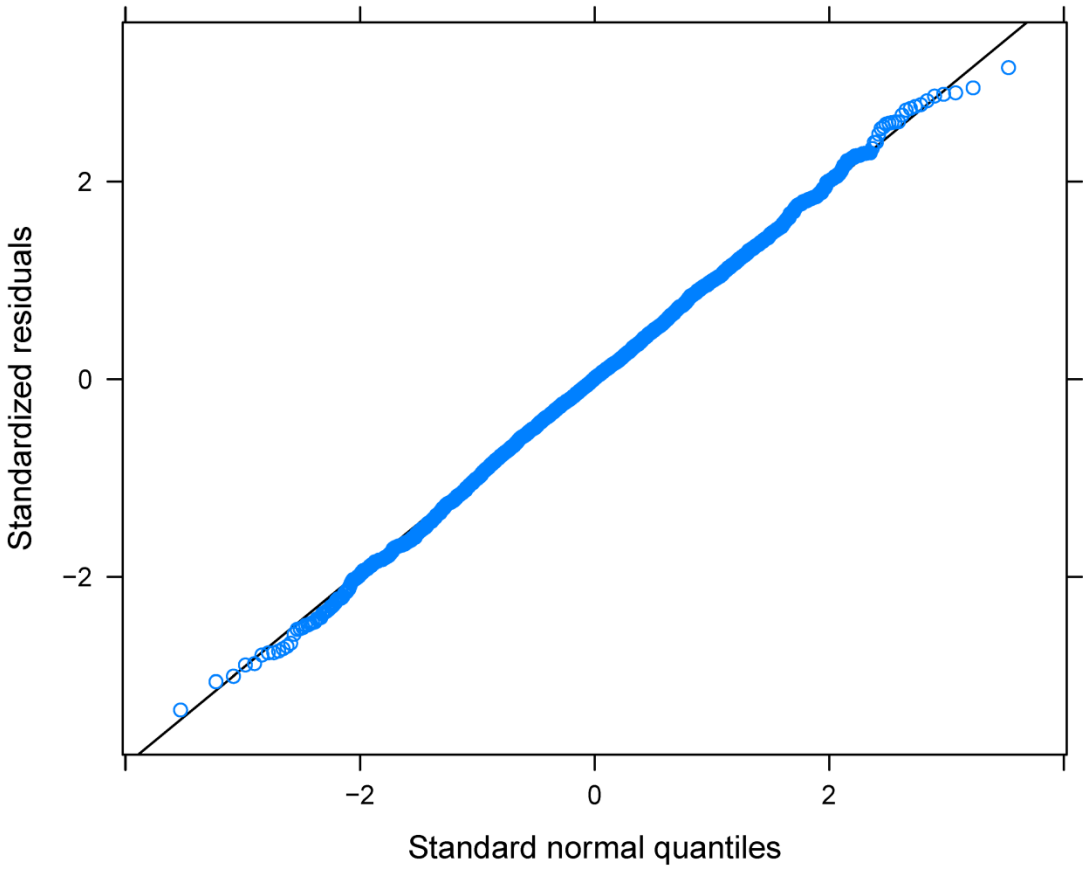
```
plot(hlm.los2, resid(., scaled=TRUE) ~ age | hosp, abline = 0)
```



- No obvious nonlinearity

qqplot for the random effects

```
qqmath(hlm.los2)
```



- Normality assumption looks valid

HLM example – Home-based hospice care

- Chen et al., Medicine 2015
- Objective: to assess the impact of home-based hospice care on end-of-life medical expenditure
- Subject: 78,613 cancer decedents in Taiwan
- HLM used to account for potential clustering effect by hospitals

Medicine®

OBSERVATIONAL STUDY (STROBE COMPLIANT)

OPEN

Home-Based Hospice Care Reduces End-of-Life Expenditure in Taiwan

A Population-Based Study

Li-Fu Chen, MD, Chun-Ming Chang, MD, and Chih-Yuan Huang, MD

HLM example – Home-based hospice care

Statistical Methods

The SPSS version 15 (SPSS, Inc., Chicago, IL) was used to analyze the data. We used χ^2 test for categorical variables and 1-way ANOVA test for continuous variables in Table 1. We performed multilevel analysis (hierarchical linear regression) using hospital as a random-intercept model to compare the EOL medical expenditure in their last 6 months of life between cancer decedents with and without home-based hospice care in Table 2. We also used multilevel analysis using a random-intercept model to compare the EOL medical expenditure in their last 6 months of life between cancer decedents with and without home-based hospice care by different hospital spending intensities in Table 3. The model was chosen because of the potential clustering effect exerted by a hospital. Hospital-level policies, procedures, or physician compensation mechanisms may produce differential EOL expenditure patterns in each hospital. A value of $P < 0.05$ was used to determine statistical significance.

- Dependent variable: End-of-life medical expenditure
- Performed multilevel analysis (HLM)
- Random-intercept model
- To account for hospital-level effect

HLM example – Home-based hospice care

TABLE 2. Medical Cost for Taiwanese Cancer Decedents From 2009 to 2011 by Multivariate Analysis Using a Random-Intercept Model (n = 78,613)

Parameters	Estimate	*P* Value
Home-based hospice care		
No use hospice	Reference	
Use hospice	−2452	<0.001

Medical cost of aggressive in the last 6 months of life US$12,397 ± 10,754. Adjust for the patients' socioeconomic status, gender, age group, Charlson Comorbidity Index Score, cancer group, primary physician's specialty, hospital characteristics, caseload group, urbanization, geographic region, and year.

TABLE 3. Medical Cost for Taiwanese Cancer Decedents From 2009 to 2011 by Multivariate Analysis Using a Random-Intercept Model According to Hospital Spending Intensity

Parameters	Estimate	*P* Value
High-hospital spending index (mean = 13,776)		
Home-based hospice care		
No use hospice	Reference	
Use hospice	−2811	<0.001
Medium-hospital spending index (mean = 12,655)		
Home-based hospice care		
No use hospice	Reference	
Use hospice	−2677	<0.001
Low-hospital spending index (mean = 10,382)		
Home-based hospice care		
No use hospice	Reference	
Use hospice	−1682	<0.001

Adjusted for the patients' socioeconomic status, gender, age group, Charlson Comorbidity Index Score, cancer group, primary physician's specialty, hospital characteristics, caseload group, urbanization, geographic region, and year.

- Home-based hospice care associated with a lower EOF medical expenditure
- Hospital effects are 'nuisance' parameters

Review of multilevel models

- Multilevel models can describe the dependence between observations with hierarchical structure
- Provide unbiased estimates for the standard error for correct inference
- Can use fixed or random effects for the group-level effects
- Selection of models and testing of parameters can be carried out using likelihood ratio test