SingHealth DukeNUS ACADEMIC MEDICAL CENTRE

IADR GENERAL SESSION
SAN DIEGO, CA, USA · MARCH 25-28, 2026
104TH GENERAL SESSION & EXHIBITION OF THE IADR
55TH ANNUAL MEETING OF THE AADOCR
50TH ANNUAL MEETING OF THE CADR

Academic Medicine
improving patients' lives

Modelling Contextual Effects in Oral Health Disparities

Hands-On workshop: Multilevel Models in R

Huihua Li
National Dental Centre Singapore,
Duke-NUS Medical School, Singapore
27 March 2026

Conflict of interest

No disclosure

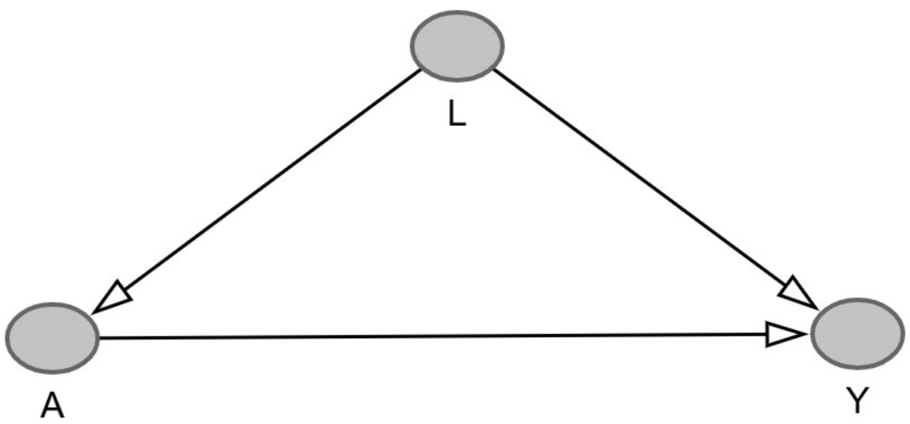
National Dental Centre Singapore
SingHealth

Restricted, Sensitive (Normal)

SingHealth DukeNUS ACADEMIC MEDICAL CENTRE

Directed Acyclic Graph

Question: Is high dental pain prevalence among school kids simply due to high sugar consumption of individual kids?



Exposure (A): Dichotomized sugar consumption
Outcome (Y): Dental pain (Yes/No)
Level 1 confounders (L1): Sex, Age, family income
Level 2 confounder (L2): School type (Private/Public)
Cluster: School

Contextual effects estimation: Hands-on in R

1. **Core Concepts**
   – General Contextual Effects (GCE) & Specific Contextual Effects (SCE)
2. **Key Metrics**
   – Variance Partition Coefficient (VPC)/Intraclass Correlation Coefficient (ICC)
   – Median Odds Ratio (MOR)
3. **Multilevel Modelling Strategy**
4. **Hands-on Demonstration**
   – Contextual effects estimation using simulated data in R

Contextual Effects

Aspect	General Contextual Effects (GCE)	Specific Contextual Effects (SCE)
Core meaning	Overall influence of the place/context itself (often unobserved)	Place adds effect via measured factors (e.g. school type)
Oral health example	Different dental pain prevalence in different schools even after individual factors adjustment	Higher dental pain prevalence in public schools persists after full adjustment — due to poor fluoride access, fewer dentists, food environment etc.
Estimation	Variance Partition Coefficient (VPC) / Intraclass Correlation Coefficient (ICC)/ Median Odds Ratio (MOR) for binary outcome	Fixed OR for higher level variable

Diez Roux (2002); Merlo, et al. (2018)

Variance Partition Coefficient (VPC)

- The proportion of the total variance that is attributable to a particular level

- VPC attributable to the higher level: often used interchangeably as Intraclass Correlation Coefficient (ICC)

$$VPC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

  – $\sigma_e^2 = \frac{\pi^2}{3} = 3.29$ when outcome is binary
  – VPC varies with covariate values if random slopes present

Snijders and Bosker RJ (2012)

Median Odds Ratio (MOR)

- $MOR = e^{\phi^{-1}(0.75)\sqrt{2\sigma_u^2}}$

  σ_u^2 = estimated variance of the random intercepts (between-group variance) from the multilevel logistic model
  $\phi^{-1}(0.75) \approx 0.6745$ (the 75th percentile of the standard normal distribution)

- How much the odds of the outcome would (typically) increase (in the median case) if a person moved from a lower-risk cluster to a higher-risk cluster — everything else being equal.

- It directly reflects the impact of the clustering/context on the individual outcome.

Merlo, et al. (2006)

Multilevel Modelling Strategy

Model	Predictors added	Purpose
M0	None (empty random intercept)	Baseline clustering (VPC/ICC)—how much variation is between clusters
M1	Main exposure only	Crude association between exposure and outcome, accounting for clustering
M2	Main exposure + compositional confounders	Compositional effects—how much of the exposure and outcome difference is due to who lives/attends there
M3	Main exposure + compositional confounders+ contextual confounders	Best estimate of true contextual/place-based effect

Multilevel Modelling

```
##Model 0: Two-level variance-components binomial model
M0 <- glmer(dental_pain ~ 1 + (1|school_id), data = Pain,
            family = binomial(link="logit"))

# Model 1: Two-level Random Intercept binomial model with exposure only
M1 <- update(M0, . ~ . + high_sugar)

# Model 2: Two-level Random Intercept binomial model adjusted
##by level 1 confounders
M2 <- update(M1, . ~ . + age_c + sex + family_income)

# Model 3: Two-level Random Intercept binomial model adjusted by
##all confounders
M3 <- update(M2, . ~ . + school_type)
```

Results: M0

```
## Results of M0
summary(M0)

## probability of
## having pain
plogis(fixef(M0))
#Observed prevalence
mean(Pain$dental_pain)
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace Approximation) ['glmerMod']
 Family: binomial  ( logit )
Formula: dental_pain ~ 1 + (1 | school_id)
   Data: Pain

     AIC       BIC    logLik -2*log(L)  df.resid
  6301.1    6314.2   -3148.6    6297.1      4998

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.8974 -0.7440 -0.5083  0.9823  2.5854

Random effects:
 Groups    Name        Variance Std.Dev.
 school_id (Intercept) 0.6977   0.8353
Number of obs: 5000, groups:  school_id, 300

Fixed effects:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.59988    0.05836  -10.28   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

```
> plogis(fixef(M0))
(Intercept)
  0.3543713
> #Observed prevalence
> mean(Pain$dental_pain)
[1] 0.374
```

Results: M0

```r
# manual VPC in a tidy way
var_u <- VarCorr(M0)$school_id[1,1]
vpc <- var_u / (var_u + pi^2 / 3)
var_u
vpc
```

```
> var_u
[1] 0.6977395
> vpc
[1] 0.174977
```

```r
# VPC with performance
icc(M0)
```

```
# Intraclass Correlation Coefficient

    Adjusted ICC: 0.175
  Unadjusted ICC: 0.175
```

National Dental Centre Singapore
SingHealth

Restricted, Sensitive (Normal)

SingHealth DukeNUS ACADEMIC MEDICAL CENTRE

VPC/ICC

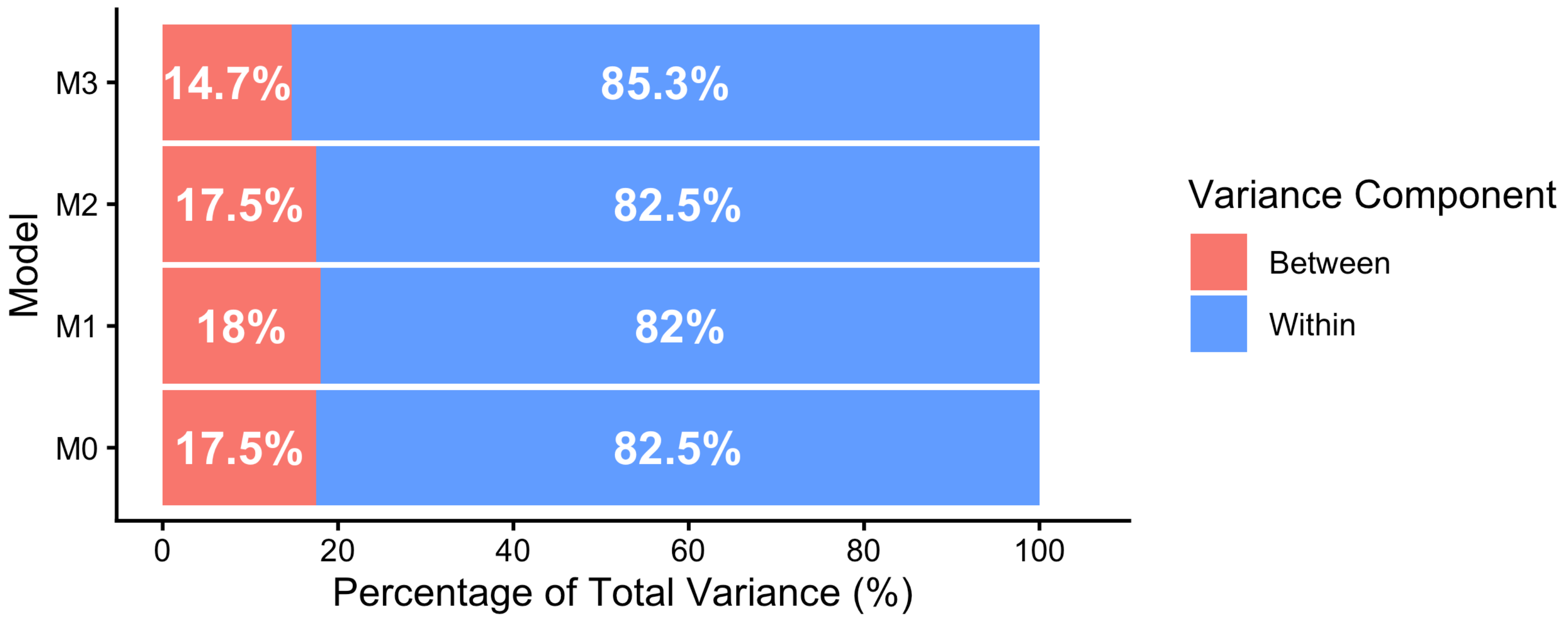
```r
models <- list(M0 = M0, M1 = M1, M2 = M2, M3 = M3)

vpc_table <- map(names(models), ~ tibble(
  Model       = .x,
  VPC_percent = round(icc(models[[.x]])$ICC_adjusted[1] * 100, 1),
  School_Var  = VarCorr(models[[.x]])$school_id[1, 1]
)) %>% list_rbind()

vpc_table %>% print()
```

	Model	VPC_percent	School_Var
	<chr>	<dbl>	<dbl>
1	M0	17.5	0.698
2	M1	18	0.721
3	M2	17.5	0.700
4	M3	14.7	0.565

VPC/ICC



M3
M2
M1
M0
Model
14.7%
85.3%
17.5%
82.5%
18%
82%
17.5%
82.5%
0
20
40
60
80
100
Percentage of Total Variance (%)
Variance Component
Between
Within

Final model: M3

```
tidy(M3, conf.int = TRUE, exponentiate = TRUE) %>%
  filter(effect == "fixed") %>%
  dplyr::select(term, estimate, std.error, statistic, p.value, conf.low, conf.high) %>%
  mutate(across(where(is.numeric), ~ round(., 3)))
```

```
# A tibble: 6 × 7
  term                estimate std.error statistic p.value conf.low conf.high
  <chr>                  <dbl>     <dbl>     <dbl>   <dbl>    <dbl>     <dbl>
1 (Intercept)            0.144     0.02     -13.8    0        0.109     0.189
2 high_sugarhigh         2.28      0.156     12.0    0        1.99      2.61
3 age_c                  1.05      0.035      1.58   0.114    0.987     1.12
4 sexfemale              1.30      0.085      4.10   0        1.15      1.48
5 family_incomelow       1.16      0.082      2.08   0.037    1.01      1.33
6 school_typepublic      2.78      0.405      7.00   0        2.09      3.70
```

Specific Contextual Effect: OR=2.78 (2.09, 3.70)

Final model: M3

General Contextual Effect

```
##General contextual effect
# MOR: for use with the multilevel logistic regression model
# var_u denote the estimate variance of the random effects.
var_u=VarCorr(M3)$school_id[1,1]
MOR <- exp(((2*var_u)^0.5)*qnorm(0.75))
MOR
```

```
> MOR
[1] 2.048593
```

National Dental Centre Singapore
SingHealth

Restricted, Sensitive (Normal)

SingHealth DukeNUS ACADEMIC MEDICAL CENTRE

Single level model

```r
M_SL <- glm(dental_pain ~ high_sugar + age_c + sex + family_income + school_type,
            data = Pain, family = binomial(link="logit"))
tidy(M_SL, conf.int = TRUE, exponentiate = TRUE) %>%
  dplyr::select(term, estimate, std.error) %>%
  mutate(across(where(is.numeric), ~ round(., 3)))
```

```
# A tibble: 6 × 7
  term               estimate std.error
  <chr>                 <dbl>     <dbl>
1 (Intercept)           0.176     0.091
2 high_sugarhigh        2.11      0.063
3 age_c                 1.05      0.031
4 sexfemale             1.31      0.06
5 family_incomelow      1.13      0.066
6 school_typepublic     2.49      0.088
```

```
# A tibble: 6 × 7
  term               estimate std.error
  <chr>                 <dbl>     <dbl>
1 (Intercept)           0.144     0.02
2 high_sugarhigh        2.28      0.156
3 age_c                 1.05      0.035
4 sexfemale             1.30      0.085
5 family_incomelow      1.16      0.082
6 school_typepublic     2.78      0.405
```

Summary

- Contextual effects should be considered when investigating associations between exposures and oral health outcomes — ignoring them risks incomplete or biased explanations of disparities.
- Both General Contextual Effects (GCE: ICC/MOR/residual clustering) and Specific Contextual Effects (SCE: fixed ORs for measured high level factors) provide complementary insights.
- Multilevel regression is the recommended methodology to properly quantify, distinguish, and interpret these effects in oral health disparities research.

References

- Diez Roux AV. A glossary for multilevel analysis. J Epidemiol Community Health. 2002 Aug;56(8):588-94. doi: 10.1136/jech.56.8.588. PMID: 12118049; PMCID: PMC1732212.
- Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, Råstam L, Larsen K. A brief conceptual tutorial of multilevel analysis in social epidemiology: using measures of clustering in multilevel logistic regression to investigate contextual phenomena. J Epidemiol Community Health. 2006 Apr;60(4):290-7. doi: 10.1136/jech.2004.029454. PMID: 16537344; PMCID: PMC2566165.
- Merlo J, Wagner P, Austin PC, Subramanian SV, Leckie G. General and specific contextual effects in multilevel regression analyses and their paradoxical relationship: A conceptual tutorial. SSM Popul Health. 2018 May 19;5:33-37. doi: 10.1016/j.ssmph.2018.05.006. PMID: 29892693; PMCID: PMC5993177.
- Snijders TAB, Bosker RJ (2012) Multilevel analysis: an introduction to basic and advanced multilevel modeling. Sage, Los Angeles

SingHealth DukeNUS
ACADEMIC MEDICAL CENTRE

Academic Medicine
improving patients' lives

Thank you!

gmslh@nus.edu.sg