Impact of Small Group Size on Neighborhood Influences in Multilevel Models

Katherine P. Theall[1], **Richard Scribner**[1], **Stephanie Broyles**[2], **Qingzhao Yu**[1], **Jigar Chotalia**[1], **Neal Simonsen**[1], **Matthias Schonlau**[3], and **Bradley P. Carlin**[4]

[1]Louisiana State University Health Sciences Center School of Public Health, New Orleans, Louisiana, United States

[2]Pennington Biomedical Center, Population Sciences, Baton Rouge, Louisiana, United States

[3]RAND Corporation, Center for Population Health and Health Disparities, Santa Monica, California, United States

[4]University of Minnesota School of Public Health, Biostatistics, Minneapolis, Minnesota, United States

Abstract

Background—Given the growing availability of multilevel data from national surveys, researchers interested in contextual effects may find themselves with a small number of individuals per group. Although there is a growing body of literature on sample size in multilevel modeling, few have explored the impact of group size < 5.

Methods—In a simulated analysis of real data, we examined the impact of group size < 5 on both a continuous and dichotomous outcome in a simple two-level multilevel model. Models with group sizes 1 to 5 were compared to models with complete data. Four different linear and logistic models were examined: empty models, models with a group-level covariate, models with an individual-level covariate, and models with an aggregated group-level covariate. We further evaluated whether the impact of small group size differed depending on the total number of groups.

Results—When the number of groups was large (N=459), neither fixed nor random components were affected by small group size, even when 90% of tracts had only 1 individual per tract and even when an aggregated group -level covariate was examined. As the number of groups decreased, the standard error estimates of both fixed and random effects were inflated. Furthermore, group-level variance estimates were more affected than were fixed components.

Conclusions—Datasets where there are a small to moderate number of groups with the majority very small group size (n < 5) size may fail to find or even consider a group-level effect when one may exist and also may be under-powered to detect fixed effects.

Keywords

Multilevel; Neighborhood; Body Weight; Obesity; Sample Size

Corresponding Author and Reprints: Katherine P. Theall, PhD, Louisiana State University Health Sciences Center, School of Public Health, Epidemiology Program, 1615 Poydras St., Suite 1400, New Orleans, LA 70112, kthea1@lsuhsc.edu.

INTRODUCTION

Numerous national population-based health surveys in the U.S. have now been georeferenced, permitting investigators to examine health outcomes from a multilevel framework. These include, for example, the National Survey of Families and Households, the Centers for Disease Control and Prevention's (CDC) National Center for Health Statistics data (e.g., National Survey on Family Growth, National Health and Nutrition Examination Survey), the Fragile Families and Child Well being Survey, the National Longitudinal Study of Adolescent Health, among others. With the growing availability of linked individual- and community-level data, investigators may find themselves with many places sampled throughout a country but few respondents per contextual group (be it neighborhood or other grouping) [1]. The degree to which this occurs will of course depend on the level of grouping and survey design, but will be most pronounced when the census tract or block group is utilized.

In general, it has been shown that for higher-level, contextual effects (i.e., that contribute to between-group variance), the number of groups appears to be more important for unbiased estimates (including appropriate standard errors) and model performance in a multilevel analysis than the group size [2, 3]. Although there is a growing body of literature on the effects of group size (i.e., number per group) and on the number of groups in multilevel or hierarchical modeling [4–8], to our knowledge, only one other study has examined the impact of group sizes less than five [1]. Such situations are increasing due to the availability of geographically-referenced national survey data and multilevel studies that will follow. While the purpose of Clarke and colleague's study was to evaluate the impact of single-level (disaggregating) vs. multilevel linear and discrete models on fixed and random components, they found, using Monte Carlo simulated data, that two-level multilevel models can be reliably estimated with small group sizes of an average of only five observations per group. The authors also found that, with extremely small group sizes, group-level variances may be overestimated, leading to Type II error. However, disaggregating the multilevel design, as they also demonstrated, may increase the risk of Type I error. This study, however, did not examine the impact of small group size on varying number of groups or the impact of small group size on individual-level, group-level, and aggregated group-level covariates. Given that small group size may lead to group-level variances being overestimated, leading to Type II error, it is important to examine the impact of small group sizes on outcomes other than model estimation.

Our objective was to examine the impact of a group size less than five on both fixed and random components in a simple two-level multilevel model, with a sufficient number of groups to test random slope variances[3, 9–11]. While like Clarke [1]we also use continuous and dichotomous outcomes and examine extremely small group sizes, we focus on both the fixed as well as random components and expand analyses to examine (*using real data*) not only empty models, but models with a group-level covariate, models with an individual-level covariate, and models with an aggregated (from individual group value) group-level covariate. Additionally, we explore whether the impact of small group size on fixed and random effects differs depending on the total number of groups—a parameter held constant (at 200) in Clarke's study.

MATERIALS AND METHODS

We utilize data from a multilevel study conducted in the United States that examined influences on body mass index (BMI). Data were obtained from Louisiana Department of Motor Vehicles (DMV) driver's license records from 1997 and included 223,747 individuals nested in 459 census tracts (taken as the "neighborhood" unit). This study was approved by

the Institutional Review Board of Louisiana State University Health Sciences Center, New Orleans, Louisiana.

Outcomes

BMI was determined using reported heights and weights and calculated using the Centers for Disease Control (CDC) formula[1] (i.e., (weight in pounds / (height in inches)^2) × 703). BMI was examined as a continuous outcome of interest, while overweight or obesity was examined as a dichotomous outcome. Overweight or obesity was defined as a BMI of 25 or greater. According to CDC definition, an adult who has a BMI between 25 and 29.9 is considered overweight; a BMI of 30 or higher is considered obese. The prevalence of obesity or overweight in the study sample was 48.1% and mean BMI (SD, one standard deviation) was 25.5 (5.1).

BMI and overweight/obesity were chosen as outcomes given their universal measurement and the fact that BMI has been shown to have a moderately high (e.g., > 4.0%) intraclass correlation coefficient (ICC) in our and other previous research [12, 13]. Group-level influences on individual outcomes are often expressed as the ICC, calculated for our linear model as:

$$\frac{V_{\text{neighborhood}}}{V_{\text{neighborhood}} + V_{\text{individual}}} \times 100\%$$

where $V_{\text{neighborhood}}$ = variance between census tracts or neighborhoods and $V_{\text{individual}}$ = variance among individuals within neighborhoods. An ICC at or above 2% is suggestive of a potential higher level effect (e.g., neighborhood) and worth examining in a multilevel framework [14].

Independent Variables

While the primary models examined were empty random coefficient models (i.e., no independent variables included) to determine the impact on basic random and fixed components, we also ran models with (a) one group-level covariate, socioeconomic status (SES); (b) one individual level covariate, individual age in years; and (c) one group-level covariate calculated using the aggregate of individual level data, aggregate age. SES and age were chosen given their association with BMI [15, 16]. The socioeconomic index was calculated for all tracts as the sum of z-scores of three factors in the U.S. Census: % with less than high school education, % living in poverty, and % of males not in the labor force. Aggregate age was defined at the group level as the average age of sampled individuals per census tract. An aggregate group-level covariate was examined to explore the potential impact of group size on aggregated compositional factors, which are common in many multilevel studies.

Simulation Models and Procedures

The proportion of tracts with five or fewer individuals —10, 25, 50, 75, and 90 %— as well as the number of individuals in the "low n" tract (2, 3, 4, or 5) were varied. One hundred (100) datasets were generated for each condition based on random sampling with replacement from the original dataset which included 223,747 individuals nested in 459 census tracts: the tracts contributing small n and individuals within tracts were randomly sampled in R according to the various simulated conditions (R version 2.8®). Analyses of

---

[1]BMI defined by CDC in: http://www.cdc.gov/nccdphp/dnpa/bmi/index.htm,

the simulated datasets were performed using two different models—one a linear random coefficients model (BMI as outcome) and one a logistic random coefficients model (overweight/obesity as outcome). All analyses were conducted in SAS version 9, with PROC MIXED used for continuous and PROC GLIMMIX for dichotomous outcomes. Restricted maximum likelihood (REML) estimation was employed for all models. Finally, each dataset was analyzed according to four different models – 1) the empty linear and logistic models, 2) linear and logistic models with a group-level covariate (SES), 3) linear and logistic models with an individual-level covariate (age), and 4) linear and logistic models with a group-level covariate (mean age) based on aggregated individual data – and the parameter estimates were summarized over the 100 simulated datasets for each condition.

In all models particular attention was paid to the estimates and standard errors for both the fixed and random components, including measures of the group- or tract-level influence or the variance in individual outcomes that can be attributed to differences between census tracts or groups [3, 14]. These measures are particularly useful when examining group-level influences on health or other outcomes.

For the logistic models, the ICC was calculated by following the linear threshold model or latent variable model method formula of Snijders [3] based on an underlying continuous variable with $V_{individual} = \Pi^2 / 3$ (i.e., 3.29). This assumes that the unobserved individual variance follows a logistic distribution, so that the variance of a standard logistic distribution is $\Pi^2 / 3$. However, the pseudo ICC for non-linear models may be difficult to understand in epidemiological terms and therefore we also examined the Median Odds Ratio (MOR) as described by Merlo and colleagues [17]. The MOR, like the ICC calculation using the linear threshold model method, is independent of the prevalence of the outcome. It represents the median value of the odds ratio for all possible comparisons of individuals from a lower to higher risk area. High group-level variation in the risk (i.e., greater group-level influence) would result in higher MOR values, while low group-level variation in risk would result in lower MOR values (i.e., close to 1.0). The MOR was calculated as:

$$\exp[\,0.95(\sqrt{V_{neighborhood}}\,)]$$

where $V_{neighborhood}$ = variance between neighborhoods.

While the impact of group size was of primary interest, we also varied the number of groups(N=459, 100, 50, and 30) for a small group size scenario of 2 individuals per tract (i.e., 90% of the group shaving 2 per tract). One hundred (100) datasets were generated for each condition (N) based on random sampling from the original dataset, sampling N tracts and then a random sample of 2 individuals per tract for 90% of the selected tracts. PROC MIXED and GLIMMIX were used to run multi-level models and the average estimates (out of 100) were examined. These simulations were run for both linear and logistic models and for models with no covariates (empty model) as well as models with individual-level, group-level, and aggregate -level covariates.

RESULTS

Table 1 presents the results of the empty random effects linear (top panel) and logistic(lower panel) regression models with one individual per tract for 90 to 10% of the tracts, with the full sample results included in the first column. The average number of individuals per tract in the original dataset was 447 (range = 4–1549). With all sampled subjects per tract included, a significant amount of the variance in BMI was apportioned to the census tract or

neighborhood level, evidenced by the ICC (4.23%). As shown in the second through fifth columns, there was little change in the average estimates in both fixed and random components across the sampling schemes—90% to 10% of tracts with two per tract. Even with a group size of one in 90% of census tracts, there is little impact on average estimates and only slight differences in the magnitude of the random components and the ICC (e.g, ICC=4.23% in full model vs. 4.45% in model estimates with 90% of tracts with n=1). This held true for logistic regression models as shown in the lower panel of Table 1, with minimal changes in the ICC or MOR estimates. There was inflation of the standard errors and decreased precision across confidence intervals across the conditions, although this may be due to the reduced sample size.

Results of linear and logistic model simulations for a group size of two are presented in Table 2, demonstrating the minimal changes observed across sampling schemes with a group size of 2. Although not presented in tables, similar results were seen in simulations with group size of 3, 4 and 5.

Shown in Table 3 are the results from the linear and logistic random effects with the addition of an individual-level (age) covariate to models with two individuals per tract for 90 to 10% of tracts. As seen in the empty regression models, other than inflated standard errors, there was little change in either fixed or random components for all models, even with 90% of tracts having a group size of two. Similar trends were observed when a group-level covariate (SES) was added to the empty models, as shown Table 4 for both linear (top panel) and logistic (bottom panel) models, and for samples with group sizes of 1, 3, 4, and 5.

Because aggregation of individual-level variables to obtain a group-level factor is common in multilevel analyses, we also examined the impact of an aggregated group-level factor on both fixed and random components. Table 5 presents the results of linear (top panel) and logistic (bottom panel) models with a group size of two for 90 to 10% of tracts when aggregated age is included in the models. As was the case with other models, there was very little difference in estimates from the full model, even when 90% of tracts had only two individuals per tract. Although not shown, results were similar for samples with group sizes of 1, 3, 4, and 5.

Because results thus far have been based on a relatively large total number of groups (N=459), we also examined the impact of small group size for varying number of groups. Table 6 presents the results of these simulations for linear models (empty, individual-level, group-level, and aggregate-level covariates) across N's of 459 (original), 100, 50 and 30 and with 90% of the groups having 2 per group. We limited the number of groups to 30 based on the general 30/30 rule[9], assuming that 30 would be an average minimal number of groups for any two-level multilevel analysis. As shown in Table 6, as the number of groups decreased the standard error estimates of both fixed and random effects were inflated. Furthermore, group-level variance estimates were inflated as the number of groups became smaller.

With respect to inflated standard errors, results were similar for logistic regression models. However, we observed no substantial difference in the magnitude of the group-level variance estimate (and therefore ICC and MOR) across varying number of groups. With a small group size (90% of tracts having 2 per tract) and N=30, e.g., the average group-level variance for the empty logistic model was 0.008066 (ICC=0.24, MOR=1.08) while for N=459 this average estimate was 0.007135 (ICC=0.23, MOR=1.08).

DISCUSSION

In general, neither the fixed nor the random effects parameter estimates were affected by small group size when the number of groups was large. This is true for empty models as well as models that include individual-or group -level covariates. While the addition of an individual-and group -level covariate had little impact on the simulated models, we thought this may not be the case for an aggregated group -level factor but again, there was little change from full sample model results. Inflation of the standard errors was observed but likely due to the decreased sample size.

When the number of groups was also varied, however, inflation of the standard errors of both fixed and random components were substantial. Furthermore, random components were impacted by small group size when the number of groups was also smaller (e.g., 50 or 30), with upward bias in the random between-group variance component estimates observed in this study. Although the standard errors are not shown in Table 6 for random components, the between -tract random variance component remained insignificant even with 100 groups(albeit 90% with n=2), at a relatively high original (full data) ICC (4.23%). While somewhat expected due to decreasing sample size, given the lack of a generally applicable formula for the standard error with REML estimators [2], it is difficult to tell whether the inflation is beyond what would be expected by increasing sample error. As the number of groups increased, the standard errors (for both fixed and random components) and random variance estimates begin to approach those seen in the full data set(N=459) with a small group size (90% with n=2).

Results are similar to Clarke and colleagues [1], and build on a small group size's impact on random components and covariate types. However, we find differences with respect to the effect on standard error estimates and to random components when the group size is small and number of groups is 100 or less. Even with a sufficient number of groups based on published rules of thumb [9], there may be bias with extremely small (i.e., n=2) group size.

While this study is not without its limitations, our findings have implications for research into not only group-level effects on individual outcomes, but also on individual-level factors. With respect to group-level effects, if the ICC or MOR or equivalent measure is used as the primary judge of the relative importance of a neighborhood-level risk factor, then conclusions will depend on the type of outcome and regression model. When all (or nearly all) groups have a small group size and the number of groups is also small (e.g., 30 or 50), the group-level variance and ICC calculation is biased upward. This noise could be due to the use of real data. The ICC or MOR does not appear to be as impacted by small group size and smaller number of groups in the case of the logistic model with the latent variable approach [18]for ICC calculation in a logistic model.

Beyond the ICC or MOR estimates, however, is the case of even considering group-level effects. If a substantial proportion of groups have a small number per group and the number of groups is also small so that the standard error of the random between-group variance is inflated to the degree that an insignificant between-group variance is observed, then researchers may conclude (despite the value of the ICC or other measure of clustering) that there is no group-level effect or that there is no need to consider a group-level factor (or multilevel analysis) when, in fact, there may be. This would have implications for the type of analyses chosen as well as conclusions drawn and may be even more important when the number of groups is small. Such conclusions would lead one to perhaps disaggregate into traditional ordinary least squares or logistic regression, which would result in increased risk of Type I error [1].

Results suggest that with a small group size (n=2) and small number of groups, the between-tract random variance component may fail to reach statistical significance (Type II error), even for a relatively high ICC or when one may expect a group-level effect. Such a situation could very well occur once data is stratified or a particular subgroup of the population is singled out, e.g. black female adolescents.

If the significance of a fixed component parameter estimate is used to judge the importance of a group-level effect, our findings suggest that such inferences may be under-powered with small group size and small number of groups. The same would hold true for individual-level fixed effects parameters, given the inflated standard errors of fixed effects components. While the number of groups remains important when investigating group-level or contextual effects, the group size should also be taken into account. Researchers working with multilevel study designs should remain aware of small group size and the number of groups when conducting research on full data or a subset of the data (e.g., among one age group). Furthermore, additional simulations are warranted to examine more fully the threshold at which very small group sizes may have an impact on fixed and random components, as well as the impact of group size on the number of covariates placed in the model and differential small group sizes based on exposure or outcomes of interest.

Acknowledgments

References

1. Clarke P. When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. Journal of Epidemiology and Community Health. Aug 1; 2008 62(8):752–8. [PubMed: 18621963]

2. Bickel, R. Multilevel Analysis for Applied Research: It's Just Regression!. New York, NY: Guilford Press; 2007. p. 428

3. Snijders, T.; Boskers, R. Multilevel analysis: An introduction to basic and advanced multilevel modeling. London: Sage; 1999.

4. Maas CJM, Hox JJ. Sample sizes for multilevel modeling. Am J Public Health. 1999; 89:1181–6. [PubMed: 10432903]

5. Maas CJM, Hox JJ. Robustness issues in multilevel regression analysis. Statistica Neerlandica. 2004; 58(2):127–37.

6. Raudenbush S, Xiao-Feng L. Effects of study duration, frequency of observation, and sample size on power in studies of group differences in polynomial change. Psychol Methods. 2001; 6(4):387–401. [PubMed: 11778679]

7. Moineddin R, Matheson FI, Glazier RH. A simulation study of sample size for multilevel logistic regression models. BioMed Central. 2007:34.

8. Snijders TAB, Bosker RJ. Standard Errors and Sample Sizes for Two-Level Research. 1993; 18(3): 237–59.

9. Kreft, IGG.; de Leeuw, J. Introducing multilevel modeling. Newbury Park, CA: Sage Publications; 1998.

10. Heck, RH.; Thomas, SL. An Introduction to Multilevel Modeling Techniques. Lawrence Erlbaum Associates; 2000.

11. Hox, JJ. Multilevel Analysis: Techniques and Applications. Lawrence Erlbaum Associates; 2002.

12. Holsten JE. Obesity and the community food environment: a systematic review. Public Health Nutr. 2009; 12(3):397–405. [PubMed: 18477414]

13. Scribner R, Mason K, Simonsen N, Su J, Theall K. Obese Neighborhoods: A multilevel, spatial analysis of environmental predictors of BMI at the neighborhood level. Obesity Research. under review.

14. Bryk, AS.; Raudenbush, SW. Hierarchical Linear Models: Applications and Data Analysis Methods. 1992.

15. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and Trends in Obesity AmongUS Adults, 1999–2000. JAMA. Oct 9; 2002 288(14):1723–7. [PubMed: 12365955]

16. Sundquist J, Malmstrom M, Johansson SE. Cardiovascular risk factors and the neighbourhood environment: a multilevel analysis. Int J Epidemiol. Oct 1; 1999 28(5):841–5. [PubMed: 10597980]

17. Merlo J, Chaix B, Ohlsson H, Beckman A, Johnell K, Hjerpe P, et al. 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. Apr 1; 2006 60(4):290–7. [PubMed: 16537344]

18. Snijders, T.; Boskers, R. Multilevel analysis An introduction to basic and advanced multilevel modeling. London: Sage; 1999.

Table 1

Impact of Group Size = 1 on Empty Model for Body Mass Index from DMV Study Data – Linear Random Effects Model

n=4 to 1549, average=447 [a]	% of Tracts with 1 per Tract				
	90	75	50	25	10
Intercept	25.61 (25.35–25.27)	25.59 (25.46–25.72)	25.61 (25.52–25.69)	25.59 (25.53–25.65)	25.60 (25.56–25.63)
25.5985					
Standard error	0.14 (0.12–0.15)	0.09 (0.08–0.10)	0.07 (0.08–0.07)	0.06 (0.056–0.062)	0.05 (0.053–0.055)
0.05143					
Random Effects					
Variance between tract intercepts	1.15 (0.76–1.53)	1.16 (0.92–1.40)	1.15 (1.03–1.27)	1.13 (1.06–1.21)	1.14 (1.10–1.18)
1.1357					
Variance within tracts	24.63 (22.82–26.45)	24.49 (23.59–25.38)	24.55 (23.97–25.14)	24.5 (24.15–24.88)	24.50 (24.29–24.72)
24.515					
ICC%	4.45 (3.00–5.91)	4.51 (3.56–5.47)	4.47 (4.00–4.95)	4.42 (4.12–4.72)	4.44 (4.29–4.59)
4.23					

Impact of Group Size = 1 on Empty Model for Obesity or Overweight in DMV Study Data– Logistic Random Effects Model

n=4 to 1549, average=447 [a]	% of Tracts with 1 per Tract				
	90	75	50	25	10
Intercept	0.48 (0.47–0.50)	0.48 (0.47–0.49)	0.48 (0.48–0.49)	0.48 (0.48–0.49)	0.48 (0.48–0.49)
0.4851					
Standard error	0.011 (0.010–0.012)	0.0076 (0.0070–0.0082)	0.0057 (0.0054–0.0061)	0.0047 (0.0045–0.0050)	0.0043 (0.0043–0.044)
0.004124					
Random Effects					
Variance between tract intercepts	0.007 (0.0046–0.0093)	0.0071 (0.0056–0.0087)	0.0071 (0.0062–0.0080)	0.0070 (0.0065–0.0076)	0.0071 (0.0068–0.0073)
0.007068					
Variance within tracts	0.24 (0.241–0.246)	0.24 (0.241–0.245)	0.24 (0.242–0.244)	0.24 (0.242–0.244)	0.24 (0.0242–0.243)
0.2431					
ICC [b]	0.21 (0.14–0.28)	0.21 (0.17–0.26)	0.22 (0.19–0.24)	0.21 (0.20–0.23)	0.21 (0.21–0.22)
0.21					
Median Odds Ratio (MOR) [b]	1.08 (1.07–1.10)	1.08 (1.07–1.09)	1.08 (1.08–1.09)	1.08 (1.08–1.09)	1.08 (1.081–1.084)
1.08					

[a] Based on full sample of individuals included in study, range and average per tract. Other individuals per tract (group size) randomly selected and columns represent average estimates over 100 samples for each condition (e.g., 10% tracts with n per tract).

[b] ICC with individual-level variance calculated using the formula of Snijders based on an underlying continuous variable with $V_{individual} = \Pi^2 / 3$ (Snijders and Bosker, 1999). Because of limitations of the ICC for non-linear outcomes, the Median Odds Ratio (MOR) (Merlo et al., 2004) was also calculated.

Table 2

Impact of Group Size = 2 on Empty Model for Body Mass Index from DMV Study Data – Linear Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	25.60	25.61 (25.39–25.84)	25.59 (25.46–25.72)	25.60 (25.51–25.69)	25.60 (25.54–25.65)	25.60 (25.56–25.63)
Standard error	0.051	0.12 (0.11–0.13)	0.092 (0.08–0.10)	0.070 (0.066–0.073)	0.058 (0.052–0.055)	0.054 (0.053–0.055)
Random Effects						
Variance between tract intercepts	1.14	1.17 (0.71–1.62)	1.16 (0.92–1.40)	1.14 (1.01–1.27)	1.13 (1.06–1.20)	1.14 (1.09–1.18)
Variance within tracts	24.51	24.47 (22.70–26.24)	24.46 (23.59–25.38)	24.56 (24.06–25.04)	24.51 (24.13–24.89)	24.52 (24.33–24.70)
ICC%	4.23	4.56 (2.78–6.35)	4.55 (3.56–5.47)	4.45 (3.93–4.96)	4.41 (4.12–4.69)	4.43 (4.26–4.59)

Impact of Group Size = 2 on Empty Model for Obesity or Overweight in DMV Study Data– Logistic Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	0.48	0.48 (0.47–0.50)	0.48 (0.47–0.50)	0.48 (0.48–0.49)	0.48 (0.48–0.49)	0.48 (0.48–0.49)
Standard error	0.0041	0.010 (0.009–0.12)	0.0076 (0.007–0.008)	0.0057 (0.005–0.006)	0.0047 (0.004–0.005)	0.0043 (0.0042–0.0044)
Random Effects						
Variance between tract intercepts	0.0071	0.0070 (0.004–0.009)	0.0071 (0.006–0.008)	0.0071 (0.006–0.008)	0.0070 (0.006–0.007)	0.0071 (0.0067–0.0073)
Variance within tracts	0.24	0.24 (0.240–0.245)	0.24 (0.241–0.245)	0.24 (0.242–0.244)	0.24 (0.243–0.244)	0.24 (0.242–0.243)
ICC [b]	0.21	0.21 (0.13–0.30)	0.22 (0.17–0.26)	0.21 (0.19–0.24)	0.21 (0.20–0.23)	0.21 (0.20–0.22)
Median Odds Ratio (MOR) [b]	1.08	1.08	1.08	1.08	1.08	1.08

[a] Based on full sample of individuals included in study, range and average per tract. Other individuals per tract (group size) randomly selected and columns represent average estimates over 100 samples for each condition (e.g., 10% tracts with n per tract).

[b] ICC with individual-level variance calculated using the formula of Snijders based on an underlying continuous variable with $V_{individual} = \Pi^2/3$ (Snijders and Bosker, 1999). Because of limitations of the ICC for non-linear outcomes, the Median Odds Ratio (MOR) (Merlo et al., 2004) was also calculated.

Table 3

Impact of Group Size = 2 on Model *with Individual-Level Covariate* – DMV Study Data – Linear Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	23.52	23.57 (23.24–23.90)	23.50 (23.30–23.70)	23.53 (23.39–23.67)	23.52 (23.45–23.59)	23.52 (23.48–23.56)
Standard error	0.061	0.15 (0.142–0.165)	0.11 (0.10–0.12)	0.08 (0.08–0.09)	0.07 (0.066–0.073)	0.06 (0.063–0.065)
Age	0.05	0.05 (0.044–0.055)	0.05 (0.047–0.053)	0.050 (0.048–0.052)	0.05 (0.049–0.051)	0.05 (0.050–0.051)
Standard error	0.0007	0.002 (0.0020–0.0024)	0.0014 (0.0013–0.0015)	0.001 (0.0009–0.0010)	0.0008 (0.00078–0.00086)	0.0007 (0.00074–0.00077)
Random Effects						
Variance between tract intercepts	1.23	1.27 (0.78–1.76)	1.26 (0.10–1.53)	1.24 (1.09–1.38)	1.22 (1.14–1.30)	1.23 (1.18–1.28)
Variance within tracts	23.97	23.94 (22.15–25.73)	23.93 (23.02–24.84)	24.00 (23.50–24.50)	23.97 (23.59–24.35)	23.97 (23.79–24.16)
ICC%	4.89	5.05 (3.08–7.02)	5.01 (3.92–6.07)	4.90 (4.34–5.46)	4.86 (4.54–5.18)	4.88 (4.70–5.07)

Impact of Group Size = 2 on Model with *Individual*-Level Covariate – DMV Study Data – Logistic Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	0.27	0.27 (0.24–0.30)	0.27 (0.25–0.29)	0.27 (0.26–0.28)	0.27 (0.26–0.29)	0.27 (0.267–0.275)
Standard error	0.0052	0.014 (0.013–0.015)	0.0098 (0.009–0.010)	0.0072 (0.0069–0.0075)	0.0060 (0.0057–0.0062)	0.0055 (0.0054–0.0056)
Age	0.0052	0.0051 (0.004–0.006)	0.0052 (0.0049–0.0054)	0.0051 (0.0050–0.0053)	0.0052 (0.0051–0.0052)	0.0052 (0.0051–0.0052)
Standard error	0.000071	0.00022 (0.00020–0.00024)	0.00014 (0.00013–0.00015)	0.0001 (9.76E-05–0.0001)	0.000082 (7.84E-05–8.55E-05)	0.000075 (7.33E-05–7.62E-05)
Random Effects						
Variance between tract intercepts	0.0078	0.0077 (0.0048–0.011)	0.0079 (0.0062–0.0095)	0.0078 (0.0068–0.0088)	0.0077 (0.0072–0.0083)	0.0079 (0.0074–0.0081)
Variance within tracts	0.2374	0.24 (0.234–0.240)	0.24 (0.236–0.239)	0.24 (0.236–0.238)	0.24 (0.237–0.238)	0.24 (0.237–0.238)
ICC% [b]	0.24	0.23 (0.234–0.240)	0.24 (0.188–0.289)	0.24 (0.206–0.266)	0.23 (0.22–0.25)	0.24 (0.23–0.25)
Median Odds Ratio (MOR) [b]	1.08	1.09 (1.07–1.10)	1.09 (1.08–1.1)	1.09 (1.08–1.09)	1.09 (1.08–1.09)	1.09 (1.08–1.09)

[a] Based on full sample of individuals included in study, range and average per tract. Other individuals per tract (group size) randomly selected and columns represent average estimates over 100 samples for each condition (e.g., 10% tracts with n per tract).

[b]ICC with individual-level variance calculated using the formula of Snijders based on an underlying continuous variable with V$_{individual}$ = Π^2 / 3 (Snijders and Bosker, 1999). Because of limitations of the ICC for non-linear outcomes, the Median Odds Ratio (MOR) (Merlo et al., 2004) was also calculated.

Table 4

Impact of Group Size = 2 on Empty Model *with Group-Level Covariate* for *Body Mass Index* from DMV Study Data – Linear Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	25.6206	25.62 (25.48–25.76)	25.61 (25.53–25.69)	25.62 (25.56–25.68)	25.62 (25.58–25.65)	25.62 (25.60–25.64)
Standard error	0.02956	0.08 (0.06–0.10)	0.06 (0.05–0.06)	0.04 (0.04–0.045)	0.03 (0.032–0.035)	0.03 (0.030–0.032)
Socioeconomic status	−0.3380	−0.34 (−0.40–−0.28)	−0.34 (−0.39–−0.29)	−0.34 (−0.37–−0.31)	−0.34 (−0.35–−0.32)	−0.34 (−0.35–−0.33)
Standard error	0.01119	0.03 (0.02–0.04)	0.02 (0.02–0.025)	0.02 (0.01–0.02)	0.01 (0.012–0.014)	0.02 (0.01–0.01)
Random Effects						
Variance between tract intercepts	0.3318	0.32 (0.10–0.54)	0.33 (0.22–0.44)	0.36 (0.26–0.41)	0.33 (0.29–0.37)	0.33 (0.31–0.35)
Variance within tracts	24.5155	24.48 (22.71–26.24)	24.49 (23.59–25.38)	24.55 (24.06–25.04)	24.51 (24.14–24.89)	24.52 (24.33–24.70)
ICC%	1.34	1.28 (0.38–1.18)	1.32 (0.88–1.76)	1.35 (1.06–1.64)	1.32 (1.18–1.47)	1.33 (1.24–1.42)

Impact of Group Size = 2 on Model *with Group-Level Covariate* for *Obesity or Overweight* in DMV Study Data– Logistic Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	75	50	25	10
Intercept	0.4873	0.49 (0.47–0.50)	0.49 (0.48–0.49)	0.49 (0.48–0.49)	0.49 (0.48–0.49)	0.49 (0.48–0.49)
Standard error	0.002629	0.0074 (0.0057–0.0090)	0.0050 (0.0044–0.0056)	0.0037 (0.0033–0.0040)	0.0030 (0.0028–0.0032)	0.0028 (0.0027–0.0028)
Socioeconomic status	−0.02563	−0.026 (−0.031–−0.02)	−0.026 (−0.029–−0.0221)	−0.026 (−0.028–−0.023)	−0.026 (−0.0266–−0.0245)	−0.026 (−0.026–−0.025)
Standard error	0.000997	0.0028 (0.002–0.0036)	0.0019 (0.0016–0.0022)	0.0014 (0.0012–0.0015)	0.0011 (0.0011–0.0012)	0.0010 (0.00101–0.00109)
Random Effects						
Variance between tract intercepts	0.002502	0.0024 (0.0007–0.0041)	0.0025 (0.0017–0.0033)	0.0025 (0.0020–0.0030)	0.0025 (0.0022–0.0028)	0.0025 (0.0023–0.0027)
Variance within tracts	0.2431	0.24 (0.240–0.246)	0.24 (0.241–0.245)	0.24 (0.242–0.244)	0.024 (0.243–0.244)	0.024 (0.243–0.244)
ICC% [b]	0.08	0.07 (0.021–0.12)	0.07 (0.051–0.099)	0.08 (0.060–0.094)	0.07 (0.066–0.838)	0.08 (0.070–0.081)
Median Odds Ratio (MOR) [b]	1.05	1.05 (1.03–1.06)	1.05 (1.04–1.05)	1.05 (1.04–1.05)	1.05 (1.045–1.051)	1.05 (1.047–1.050)

[a] Based on full sample of individuals included in study, range and average per tract. Other individuals per tract (group size) randomly selected and columns represent average estimates over 100 samples for each condition (e.g., 10% tracts with n per tract).

[b] ICC with individual-level variance calculated using the formula of Snijders based on an underlying continuous variable with $V_{individual} = \Pi^2/3$ (Snijders and Bosker, 1999). Because of limitations of the ICC for non-linear outcomes, the Median Odds Ratio (MOR) (Merlo et al., 2004) was also calculated.

Table 5

Impact of Group Size = 2 on Model _with Aggregated Group_-Level Covariate – DMV Study Data – Linear Random Effects Model

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	**75**	**50**	**25**	**10**
Intercept	**29.57**	28.52 (25.79–32.05)	29.52 (27.01–32.02)	29.44 (27.80–31.09)	29.57 (28.66–30.47)	29.62 (29.08–30.15)
Standard error	**0.76**	1.81 (1.51–2.10)	1.37 (1.19–1.56)	1.03 (0.94–1.12)	0.87 (0.81–0.92)	0.80 (0.77–0.82)
Mean Age	**−0.096**	−0.08 (−0.15–0.03)	−0.095 (−0.15–−0.034)	−0.093 (−0.13–−0.053)	−0.096 (−0.12–−0.074)	−0.097 (−0.11–−0.084)
Standard error	**0.018**	0.043 (0.036–0.051)	0.033 (0.029–0.038)	0.025 (0.023–0.026)	0.021 (0.019–0.022)	0.019 (0.018–0.019)
Random Effects						
Variance between tract intercepts	**1.065**	1.10 (0.63–1.58)	1.08 (0.83–1.32)	1.07 (0.93–1.21)	1.06 (0.98–1.14)	1.06 (1.02–1.10)
Variance within tracts	**24.52**	24.48 (22.71–26.24)	24.49 (23.59–25.38)	24.55 (24.06–25.04)	24.51 (24.14–24.89)	24.52 (24.33–24.70)
ICC%	**4.16**	4.32 (2.45–6.19)	4.22 (3.23–5.19)	4.18 (3.64–4.72)	4.14 (3.83–4.44)	4.16 (3.99–4.32)

Impact of Group Size = 2 on Model with _Aggregate Group-Level Covariate – DMV Study Data – Logistic Random Effects Model_

	n=4 to 1549, average=447 [a]	% of Tracts with 2 per Tract				
		90	**75**	**50**	**25**	**10**
Intercept	0.75	0.72 (0.44–0.99)	0.75 (0.54–0.96)	0.74 (0.61–0.88)	0.75 (0.67–0.82)	0.76 (0.71–0.79)
Standard error	0.062	0.16 (0.13–0.19)	0.12 (0.10–0.13)	0.08 (0.08–0.09)	0.07 (0.066–0.075)	0.06 (0.062–0.067)
Mean Age	−0.0064	−0.0056 (−0.012–0.001)	−0.0064 (−0.011–−0.013)	−0.0062 (−0.009–−0.003)	−0.0064 (−0.008–−0.005)	−0.0064 (−0.0073–−0.0054)
Standard error	0.0015	0.0038 (0.0031–0.0045)	0.0028 (0.0024–0.0032)	0.0020 (0.0018–0.0022)	0.0017 (0.0016–0.0018)	0.0016 (0.0015–0.0016)
Random Effects						
Variance between tract intercepts	.0068	0.0068 (0.0038–0.0097)	0.0068 (0.0052–0.0084)	0.0068 (0.0058–0.0077)	0.0067 (0.0062–0.0072)	0.0068 (0.0065–0.0071)
Variance within tracts	0.24	0.24 (0.240–0.246)	0.24 (0.241–0.245)	0.24 (0.242–0.244)	0.24 (0.243–0.244)	0.24 (0.2428–0.2434)
ICC% [b]	0.20	0.21 (0.11–0.30)	0.21 (0.16–0.25)	0.21 (0.18–0.24)	0.20 (0.19–0.22)	0.20 (0.20–0.21)
Median Odds Ratio (MOR) [b]	1.08	1.08 (1.06–1.100)	1.08 (1.071–1.091)	1.08 (1.075–1.087)	1.08 (1.078–1.084)	1.08 (1.079–1.083)

[a] Based on full sample of individuals included in study, range and average per tract (group size) randomly selected and columns represent average estimates over 100 samples for each condition (e.g., 10% tracts with n per tract).

[b] ICC with individual-level variance calculated using the formula of Snijders based on an underlying continuous variable with $V_{individual} = \Pi^2 / 3$ (Snijders and Bosker, 1999). Because of limitations of the ICC for non-linear outcomes, the Median Odds Ratio (MOR) (Merlo et al., 2004) was also calculated.

Table 6

Impact of Decreased Number of Groups, *with 90% of Tracts with 2 per Tract* – Empty Model for Body Mass Index, Linear Random Effects Model

Empty Model	All Tracts N=459 [a]	Number of Groups / Census Tracts		
		100	50	30
Intercept (Standard error)	25.61187 (0.12257)	25.66173 (0.25406)	25.58857 (0.36100)	25.62700 (0.47108)
Between tract variance	1.23	1.26	1.43	2.19
Within tract variance	24.54	24.63	24.56	23.89
Mean ICC%	4.80	4.85	5.42	7.07
Median ICC%	4.63	4.27	4.20	4.43
Individual Covariate				
Intercept (Standard error)	23.53140 (0.15324)	23.55377 (0.32990)	23.63344 (0.47711)	23.58695 (0.58209)
Age (Standard error)	0.04989 (0.00220)	0.04890 (0.00472)	0.04996 (0.00680)	0.04832 (0.00857)
Between tract variance	1.27	1.30	1.95	2.00
Within tract variance	23.98	23.85	23.88	24.07
Mean ICC%	5.06	5.18	6.86	7.24
Median ICC%	4.94	4.79	4.84	4.27
Group-level Covariate				
Intercept (Standard error)	25.60311 (0.08179)	25.64850 (0.16242)	25.63875 (0.23657)	25.62497 (0.29773)
Socioeconomic status (Standard error)	−0.33398 (0.03122)	−0.34754 (0.06457)	−0.35392 (0.10230)	−0.32015 (0.14065)
Between tract variance	0.33	0.32	0.44	0.72
Within tract variance	24.38	24.35	24.88	24.21
Mean ICC%	1.34	1.29	1.62	2.66
Median ICC%	1.25	1.00	0.78	0.67
Aggregate Group-level Covariate				
Intercept (Standard error)	29.21490 (1.831984)	29.50759 (3.90602)	28.42699 (5.69244)	28.88522 (7.76234)
Mean Age (Standard error)	−0.08738 (0.04412)	−0.08988 (0.09415)	−0.08161 (0.13689)	−0.08567 (0.18862)
Between tract variance	1.13	1.27	1.18	1.87
Within tract variance	24.62	24.75	24.32	24.09
Mean ICC%	4.39	4.90	4.48	6.44
Median ICC%	4.25	4.37	3.62	3.28

[a] Based on full sample of tracts (N=459) with 90% having 2 individuals per tract. Other group sizes randomly selected, and then 90% of these tracts included random sample of 2 individuals per tract. Columns represent average estimates over 100 samples for each condition (e.g., N=100 with 90% with 2 per tract).