# The Delta Garden Study: A quasi-experimental, cross-sectional, nested, pair-matched design with zero inflated endpoints

Page C. Moore[1] Ph.D., Amy Schrader[1] M. S., Judith L. Weber[2] Ph. D., RD.
[1]Department of Biostatistics, University of Arkansas for Medical Sciences, Little Rock, AR
[2]Department of Pediatrics, Arkansas Children's Hospital Research Institute, Little Rock, AR

## Abstract

The Delta Garden Study is a science-based school garden intervention designed to increase fruit and vegetable (FV) intakes and minutes of moderate to vigorous physical activity (PA), as well as to improve academic achievement and reduce social risk behaviors, in middle school students in Arkansas. The study used a quasi-experimental, cross-sectional, nested, pair-matched design utilizing 6 intervention (school garden) schools and 6 control (no school garden) schools, including over 2,000 6$^{th}$, 7$^{th}$ and 8$^{th}$ grade students nested within 4-6 science teachers per school. Observed as continuous outcomes, the two primary endpoints were zero-inflated and were collected across multiple time-points. We present the design and analysis of this quasi-experimental, cross-sectional, nested, pair-matched design and the handling of longitudinally collected zero-inflated outcomes.

**Key Words:** nested, zero inflated, hierarchical, repeated measures

## 1. Introduction

Nationally, 19% of children and adolescents ages 6-19 years are obese, and a combined 35% are overweight or obese with the highest prevalence found in Hispanic boys and African-American girls[1]. In Arkansas, 21% of children and adolescents ages 6-19 years are obese, and a combined 38% are overweight or obese with the highest prevalence found in Hispanic boys, African-American girls and Hispanic girls[2]. Data collected from the 2009 YRBS found that less than 23% of kids consume fruit and vegetables 5 or more times per day and less than 14% consume vegetables 3 times per day[3]. However, 29% of kids consume soda at least once per day. The same study reported less than 19% of kids were physically active at least 60 minutes per day and 33% of students watched TV 3 or more hours per day on an average school day.

Existing school garden literature had studies with small sample sizes (n<400), usually no control schools, less than 3 schools overall, mostly limited to elementary schools, brief exposure of 12 weeks or less and focused primarily on knowledge and willingness to

taste new fruits and vegetables. In an effort to combat childhood obesity and increase fruit & vegetable intake as well as physical activity, the Delta Garden Study (DGS) was designed as a randomized control trial to test the effectiveness of school gardens in increasing mean fruit and vegetable intakes and minutes of physical activity.

## 1.1 Objective

To adapt established US Dietary Guidelines for Americans eating and physical activity patterns for children in the Delta to prevent childhood obesity through:

- Development and testing of a school gardening program;
- Designed to increase students' school bonding as the mediating mechanism;
- For increasing their physical activity and fruit/vegetable intake, improving academic achievement and decreasing social risk behaviors.

## 1.2 Study Design

Six garden schools (intervention) and 6 control schools were pair-matched on the percent of African-American, free/reduced lunch, and obese students with a total of over 2,000 $6^{th}$, $7^{th}$ and $8^{th}$ grade students nested within 4-6 science teachers per school. Table 1 shows the number of students within each pair-matched school as well as how many students within each timeframe. All schools completed the fruit/vegetable intake and physical activity questionnaires at baseline (BL), interim (INT) and final (FNL) time-points spread over the course of a school year. All schools also completed height, weight, and body fat assessments as well as school bonding and knowledge testing at only the baseline and final time-points. Therefore, some endpoints had 3 time-points where other endpoints only had 2 time-points.

**Table 1.** Pairings of intervention and control schools as well as number of students participating at each school by time-point.

| Intervention Schools Pair-match | Time-point | | | Control Schools Pair-match | Time-point | | |
|---|---|---|---|---|---|---|---|
| | Baseline (n) | Midpoint (n) | Final (n) | | Baseline (n) | Midpoint (n) | Final (n) |
| **1** | 235 | 233 | 219 | **1** | 256 | 236 | 237 |
| **2** | 207 | 188 | 188 | **2** | 193 | 191 | 192 |
| **3** | 209 | 193 | 207 | **3** | 168 | 152 | 155 |
| **4** | 208 | 199 | 201 | **4** | 241 | 219 | 229 |
| **5** | 172 | 166 | 156 | **5** | 185 | 153 | 183 |
| **6** | 217 | 146 | 149 | **6** | 244 | 211 | 208 |
| **Total** | 1248 | 1125 | 1120 | **Total** | 1287 | 1162 | 1204 |

Intervention schools were provided with one full time garden program specialist for 1 year to design and develop the garden, build a greenhouse, develop planting/harvesting calendar and co-teach all DGS lessons in the garden. Garden lesson curriculum aligned with science, math, language arts and health/wellness state frameworks. Science teachers at the intervention schools taught curriculum related to science topics on garden days. Students spent a minimum of 2 days/week working and learning in the garden including physical activity, nutrition education, tasting/prep of fruits & vegetables. In exchange for

data collection, control schools were provided a $2,000 stipend towards an academic program and assistance with state-mandated BMI measurements.

## 2. Methods

This study had a complex design consisting of nested levels (students within teachers within schools), pair-matched schools (matches based on % African-American, free/reduced lunch and obesity), 3 time-points, and adjusting for student-level covariates (gender, race, BMI and meal status). The primary endpoints were continuous variables where total intake was averaged over 2 days:

- Total fruits
- Total vegetables
- Total fruit & vegetables
- Total moderate physical activity
- Total vigorous physical activity
- Total physical activity minutes.

As shown in Figure 1, the primary endpoints contain excess zero-count data.

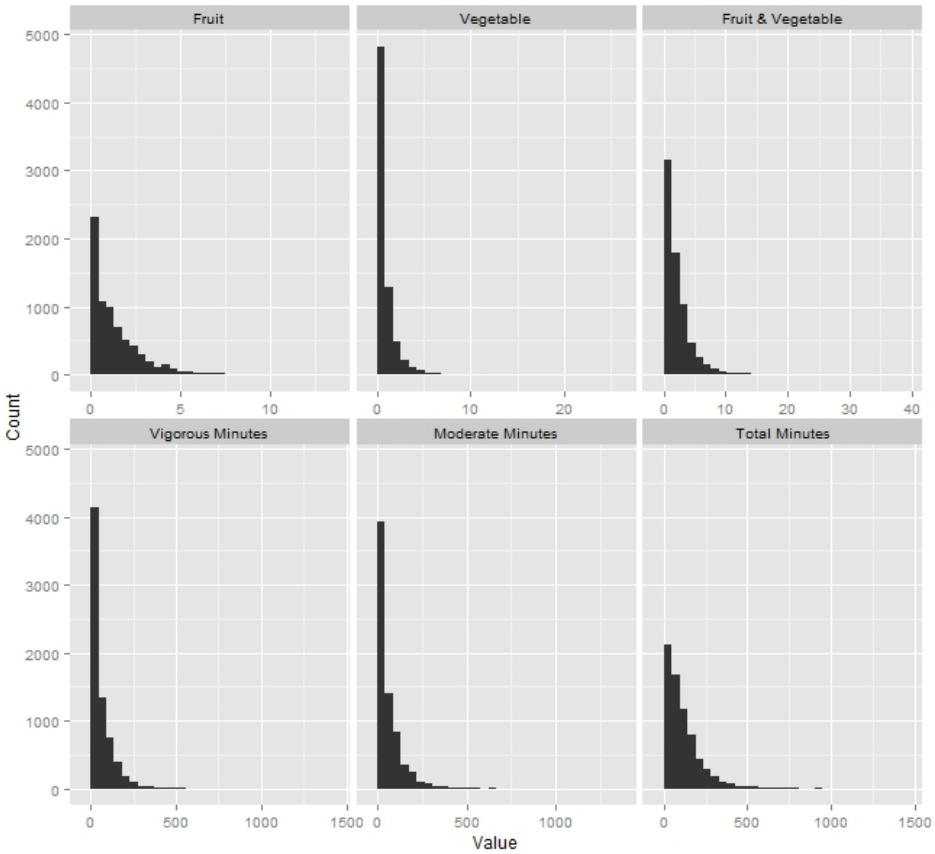


**Figure 1.** Histograms showing the zero-inflated distribution of each primary endpoint.

When attempting to model the continuous endpoints using PROC GLIMMIX, there was insufficient memory to run the full model (student level covariates, 3 time-points, 3

nested levels). We then decided to create new dichotomous outcome variables based upon a clinically meaningful increase in fruit & vegetable intake (at least ¼ cup) and physical activity minutes (at least 30 minutes).

Table 2 shows the SAS code for the final unadjusted and adjusted GLIMMIX models. Final unadjusted models accounted for nesting at the school level and teachers within schools where the outcome was the change in either fruit/vegetable intake or physical activity and the independent variable of interest was treatment group (i.e. garden school vs. control school). Final adjusted models also accounted for student level characteristics: gender, race, BMI, and meal status. In the original models, there was a lack of variation within the student within teachers within school nested level[4]; therefore, it was removed from the model.

**Table 2.** SAS code for unadjusted and adjusted generalized linear mixed models.

| **Unadjusted Model** | **Adjusted Model** |
|---|---|
| PROC GLIMMIX DATA=in<br>NOCLPRINT;<br>CLASS school teacher treatment;<br>MODEL y = treatment /<br>DIST=BINARY;<br>RANDOM INT / TYPE=VC<br>SUBJECT=school;<br>RANDOM INT / TYPE=VC<br>SUBJECT=teacher(school);<br>RUN; | PROC GLIMMIX DATA=in<br>NOCLPRINT;<br>CLASS school teacher treatment<br>gender meals race bmi;<br>MODEL y = treatment gender meals<br>race bmi / DIST=BINARY;<br>RANDOM INT / TYPE=VC<br>SUBJECT=school;<br>RANDOM INT / TYPE=VC<br>SUBJECT=teacher(school);<br>RUN; |

## 3. Results

The primary outcomes of the Delta Garden Study revolve around increasing fruit/vegetable intake as well as physical activity, whether moderate physical activity minutes, vigorous physical activity minutes and total physical activity minutes. Each of these measures was collected at baseline, midpoint and final. Table 3 shows the fruit and vegetable measures at each time-point for the intervention and control schools along with the unadjusted and adjusted p-values from the GLIMMIX models. Table 4 shows the physical activity measures at each time-point for the intervention and control schools along with the unadjusted and adjusted p-values from the GLIMMIX models.

From baseline to the midpoint of the year, the proportion of students within the intervention schools (garden schools) who increased their fruit & vegetable intake by at least ¼ cup was significantly higher than students in the control schools (unadjusted p-value = 0.0322; adjusted p-value = 0.0116). Although the differences are not statistically significant, students within the intervention schools (garden schools) during the baseline to midpoint timeframe had a higher proportion of increasing their fruit alone intake and vegetables alone intake by ¼ cup than students in the control schools. Increases in fruit and vegetable intake from the midpoint to final timeframe were not significantly different

between the intervention (garden schools) and control schools. After adjusting for student-level covariates, the proportion of students within the intervention schools (garden schools) that increased their fruit alone intake, vegetable alone intake as well as their fruit & vegetable intake was significantly higher than students in the control schools.

**Table 3.** Results of fruit and vegetable intake where the outcome variable is the proportion of students who increased their fruit and vegetable intake by at least ¼ cup from the previous time-point.

| Timeframe | Measure | Intervention | Control | Unadjusted | Adjusted |
|---|---|---|---|---|---|
| | | N (%) | N (%) | p-value | p-value |
| Baseline to Midpoint | Fruit | 395 (36.2) | 387 (34.6) | 0.4584 | 0.4309 |
| | Vegetable | 341 (31.3) | 274 (24.5) | 0.1681 | 0.2846 |
| | Fruit & Vegetable | 435 (39.8) | 386 (34.5) | 0.0322* | 0.0116* |
| Midpoint to Final | Fruit | 337 (31.9) | 326 (29.9) | 0.5983 | 0.3844 |
| | Vegetable | 268 (25.5) | 279 (25.6) | 0.8779 | 0.6623 |
| | Fruit & Vegetable | 356 (33.7) | 340 (31.2) | 0.2468 | 0.0985 |
| Baseline to Final | Fruit | 358 (32.8%) | 323 (27.8%) | 0.0210* | 0.0096** |
| | Vegetable | 266 (24.4%) | 231 (19.9%) | 0.1303 | 0.0229* |
| | Fruit & Vegetable | 375 (34.3%) | 309 (26.5%) | 0.0022** | <0.0001** |

*p<0.05, **p<0.01

For all measures of physical activity and across all three timeframes, differences in the proportion of students who increased their activity by 30 minutes were not significantly different between the intervention (garden schools) and control schools. However, intervention schools (garden schools) had a higher proportion of students who increased their moderate physical activity minutes and total physical activity minutes by 30 minutes compared to control schools within all three timeframes.

**Table 4.** Results of physical activity measures where the outcome variable is the proportion of students who increased their physical activity by at least 30 minutes from the previous time-point.

| Timeframe | Measure | Intervention | Control | Unadjusted | Adjusted |
|---|---|---|---|---|---|
| | | N (%) | N (%) | p-value | p-value |
| Baseline to Midpoint | Moderate | 190 (17.5) | 174 (15.6) | 0.5480 | 0.5702 |
| | Vigorous | 190 (17.5) | 211 (18.9) | 0.4696 | 0.5579 |
| | Total Minutes | 262 (24.2) | 241 (21.6) | 0.5072 | 0.5260 |
| Midpoint to Final | Moderate | 336 (31.9) | 325 (29.7) | 0.9264 | 0.8183 |
| | Vigorous | 207 (19.6) | 204 (18.7) | 0.8296 | 0.4587 |
| | Total Minutes | 376 (35.7) | 377 (34.6) | 0.8781 | 0.9276 |
| Baseline to Final | Moderate | 304 (28.0%) | 261 (22.3%) | 0.4666 | 0.3992 |
| | Vigorous | 191 (17.6%) | 194 (16.6%) | 0.8818 | 0.8706 |
| | Total Minutes | 335 (31.0%) | 286 (24.5%) | 0.2423 | 0.2624 |

*p<0.05, **p<0.01

## 4. Conclusions

**Recommendation.** When faced with a complex study design that ends up with a large number of zero-count data, try creating a categorical variable from the outcome variable that will still address the research question. PROC GLIMMIX will then be able to still support the study design including the repeated measures and nested levels.

**Limitations.** Although there are other potential ways in which these data could have been modeled, due to the nature of this complex study design and time constrictions, they were not explored.

## Acknowledgements

This research is based upon work supported by the USDA-Agricultural Research Service, Delta Obesity Prevention Research Unit under Agreement No. 58-6251-8-039.

## References

1. Ogden CL., Carrol MD., Curtin LR., Lamb MM., Flegal KM., Prevalence of High Body Mass Index in US Children and Adolescents, 2007-2008, JAMA. 2010; 303(3):242-249.
2. Arkansas Center for Health Improvement, Year Seven Assessment of Childhood and Adolescent Obesity in Arkansas (Fall 2009-Spring 2010), Little Rock, AR: ACHI, December 2010.
3. Centers for Disease Control and Prevention. Youth Risk Behavior Surveillance-United States. Surveillance Summaries, 2009. MMWR 2010; 59.
4. Kiernan K, Tao J, Gibbs P. Tips and Strategies for Mixed Modeling with SAS/STAT® Procedures. SAS Global Forum 2012, Paper 332-2012.