Longitudinal School-Based BMI Surveillance: Informing Obesity Prevention Strategies

BY

STEPHANIE SHAPIRO BERKSON
B.A., Brandeis University, Waltham, MA, 1993
M.P.H., University of Illinois at Chicago, Chicago, 2001

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Public Health Sciences
in the Graduate College of the
University of Illinois at Chicago, 2012

Chicago, Illinois

Defense Committee:

Noel Chavez, Chair and Advisor
Jesus Ramirez
Timothy Johnson, Public Administration
Virgina Chomitz, Tufts School of Medicine
Kathryn Keim, Rush University
DEDICATION

To my beloved family members, Grandma “B,” Mom Berkson, and Uncle Joel, each an eternal inspiration, who have passed away over the course of my doctoral work and to my supportive and loving family, Eric, Jadon, Brennan, Femur, Mom, Dad, and Dana.
ACKNOWLEDGEMENTS

I would like to thank my dissertation committee advisor, Dr. Noel Chavez, for her wisdom and unwavering support and understanding throughout this process. I would also like to thank my dissertation committee for their support, assistance, and understanding over many years. I especially would like to thank Dr. Virginia Chomitz and the Cambridge Public Schools for generously supplying me with the data, as well as teaching me about childhood and adolescent obesity through participatory evaluation.

Thank you to my husband, Eric Berkson, for his tireless support and helpful suggestions for improving this research and to Jeff Berkson for his editing adroitness.

I would finally like to thank my children and dog, Jadon, Brennan, and Femur, and their caretakers without whom I would have never completed this research, my parents and sister, Lynne, Marty, and Dana Shapiro, and Altagracia Peña.

SBSB
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td></td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>A. Problem Statement</td>
<td>1</td>
</tr>
<tr>
<td>B. Significance of the Problem</td>
<td>3</td>
</tr>
<tr>
<td>C. Purpose of the Proposed Research</td>
<td>4</td>
</tr>
<tr>
<td>1. Significance of the proposed research</td>
<td>4</td>
</tr>
<tr>
<td>2. Specific aims</td>
<td>8</td>
</tr>
<tr>
<td>II. CONCEPTUAL FRAMEWORK</td>
<td>9</td>
</tr>
<tr>
<td>A. A Social-Ecological Approach: Overview of the Cambridge Body Mass Index Surveillance System Framework</td>
<td>9</td>
</tr>
<tr>
<td>B. Environmental Rationale for the Dramatic Rise in the Prevalence of Obesity</td>
<td>12</td>
</tr>
<tr>
<td>C. Philosophical Underpinnings of the Ecological Model</td>
<td>13</td>
</tr>
<tr>
<td>D. Specific Uses of Body Mass Index as Part of an Ecological Model</td>
<td>15</td>
</tr>
<tr>
<td>1. Individual level information</td>
<td>16</td>
</tr>
<tr>
<td>2. Population level information</td>
<td>17</td>
</tr>
<tr>
<td>E. A Best Practice Methodological Framework for Body Mass Index Surveillance Analysis</td>
<td>18</td>
</tr>
<tr>
<td>III. REVIEW OF RELATED LITERATURE</td>
<td>19</td>
</tr>
<tr>
<td>A. The Role of School-Based Obesity Prevention and Intervention</td>
<td>19</td>
</tr>
<tr>
<td>B. Longitudinal Childhood and Adolescent Obesity Surveillance and the Role of Schools</td>
<td>22</td>
</tr>
<tr>
<td>C. Validity of Body Mass Index as a Measure of Body Fat</td>
<td>24</td>
</tr>
<tr>
<td>IV. METHODOLOGY AND RESULTS</td>
<td>27</td>
</tr>
<tr>
<td>A. Permission for Use of the Data and Database Cleaning</td>
<td>27</td>
</tr>
<tr>
<td>B. Manuscript 1—Reliability of Height and Weight Measurements Collected by Physical Education Teachers for a School-Based Body Mass Index Surveillance and Screening System</td>
<td>29</td>
</tr>
<tr>
<td>1. Abstract</td>
<td>29</td>
</tr>
<tr>
<td>2. Introduction</td>
<td>30</td>
</tr>
<tr>
<td>3. Methods</td>
<td>32</td>
</tr>
<tr>
<td>a. Participants</td>
<td>32</td>
</tr>
<tr>
<td>b. Instruments</td>
<td>33</td>
</tr>
<tr>
<td>c. Procedures</td>
<td>33</td>
</tr>
<tr>
<td>d. Data analysis</td>
<td>34</td>
</tr>
<tr>
<td>4. Results</td>
<td>35</td>
</tr>
<tr>
<td>a. Controlled setting assessment</td>
<td>35</td>
</tr>
<tr>
<td>i. Student sample characteristics</td>
<td>35</td>
</tr>
<tr>
<td>ii. Intra-rater reliability: visual assessment</td>
<td>36</td>
</tr>
<tr>
<td>iii. Intra-rater reliability: mean absolute differences</td>
<td>38</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>iv. Intra-rater and inter-rater reliability: intraclass correlation coefficients</td>
<td>39</td>
</tr>
<tr>
<td>v. Effect of measurement discrepancy on body mass index</td>
<td>39</td>
</tr>
<tr>
<td>b. Natural classroom setting assessment</td>
<td>40</td>
</tr>
<tr>
<td>i. Student sample characteristics</td>
<td>40</td>
</tr>
<tr>
<td>ii. Inter-rater reliability: visual assessment</td>
<td>41</td>
</tr>
<tr>
<td>iii. Inter-rater reliability: mean absolute differences</td>
<td>43</td>
</tr>
<tr>
<td>iv. Inter-rater reliability: intraclass correlation coefficients</td>
<td>44</td>
</tr>
<tr>
<td>v. Effect of measurement discrepancy on body mass index</td>
<td>44</td>
</tr>
<tr>
<td>c. Physical education teacher survey and classroom observations</td>
<td>44</td>
</tr>
<tr>
<td>5. Discussion</td>
<td>45</td>
</tr>
<tr>
<td>6. Limitations</td>
<td>47</td>
</tr>
<tr>
<td>7. Conclusions</td>
<td>47</td>
</tr>
<tr>
<td>8. Implications for school health</td>
<td>48</td>
</tr>
<tr>
<td>9. Human participants approval statement</td>
<td>49</td>
</tr>
</tbody>
</table>

C. Manuscript 2—School-Based Obesity Surveillance: Cross-Sectional Versus Longitudinal Analysis | 50 |
1. Abstract | 50 |
2. Introduction | 51 |
3. Methods | 53 |
   a. Study setting and study population | 53 |
   b. Study design | 54 |
   c. Statistical analysis | 55 |
4. Results | 55 |
5. Discussion | 66 |
6. Strengths and limitations | 69 |

D. Manuscript 3—Trajectory of Body Mass Index Percentile Growth among Sociodemographically Diverse Elementary School Students Exposed to Eight Years of System-Wide Ecological Nutrition and Physical Activity Intervention | 71 |
1. Abstract | 71 |
2. Introduction | 72 |
3. Methods | 75 |
   a. Data source | 75 |
   b. Study population | 76 |
TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>c. Intervention</td>
<td>77</td>
</tr>
<tr>
<td>d. Statistical analysis</td>
<td>77</td>
</tr>
<tr>
<td>4. Results</td>
<td>78</td>
</tr>
<tr>
<td>5. Discussion</td>
<td>89</td>
</tr>
<tr>
<td>6. Strengths and limitations</td>
<td>92</td>
</tr>
<tr>
<td>7. Conclusions</td>
<td>94</td>
</tr>
<tr>
<td>V. OVERALL CONCLUSION</td>
<td>96</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>100</td>
</tr>
<tr>
<td>APPENDICIES</td>
<td>109</td>
</tr>
<tr>
<td>Appendix A</td>
<td>110</td>
</tr>
<tr>
<td>Appendix B</td>
<td>178</td>
</tr>
<tr>
<td>Appendix C</td>
<td>179</td>
</tr>
<tr>
<td>VITA</td>
<td>183</td>
</tr>
</tbody>
</table>
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>40</td>
</tr>
<tr>
<td>II.</td>
<td>57</td>
</tr>
<tr>
<td>III.</td>
<td>58</td>
</tr>
<tr>
<td>IV.</td>
<td>59</td>
</tr>
<tr>
<td>V.</td>
<td>60</td>
</tr>
<tr>
<td>VI.</td>
<td>79</td>
</tr>
<tr>
<td>VII.</td>
<td>81</td>
</tr>
<tr>
<td>VIII.</td>
<td>84</td>
</tr>
<tr>
<td>IX.</td>
<td>85</td>
</tr>
<tr>
<td>X.</td>
<td>87</td>
</tr>
</tbody>
</table>

I. CONTROLLED SETTING ASSESSMENT: EFFECT OF HEIGHT AND WEIGHT MEASUREMENT OUTLIERS ON CDC WEIGHT STATUS CATEGORY

II. CPS STUDENT CHARACTERISTICS AT EACH STUDY TIME POINT

III. WEIGHT STATUS OF CPS STUDENTS AT EACH STUDY TIME POINT

IV. CROSS-SECTIONAL ODDS OF OBESITY AT KEY INTERVENTION TIME POINTS

V. LONGITUDINAL ODDS OF OBESITY AT KEY INTERVENTION TIME POINTS

VI. PERCENTAGE DISTRIBUTIONS FOR DEMOGRAPHIC CHARACTERISTICS OF FULL BMI SURVEILLANCE DATABASE POPULATION AND STUDY POPULATION

VII. WEIGHT STATUS OF STUDY POPULATION COMPARED TO NATIONAL PREVALENCE AT THE FIRST AND LAST STUDY TIME POINTS

VIII. MODEL I. UNCONDITIONAL GROWTH MODEL ESTIMATING BMI PCT GROWTH TRAJECTORIES OF CHILDREN AND ADOLESCENTS FROM KINDERGARTEN TO GRADE 8

IX. MODEL II. GROWTH MODEL ESTIMATING BMI PCT GROWTH TRAJECTORIES OF CHILDREN AND ADOLESCENTS FROM KINDERGARTEN TO GRADE 8 AFTER ADDING CUMULATIVE EXPOSURE TO INTERVENTION AND THE INTERACTION OF CUMULATIVE INTERVENTION EXPOSURE WITH TIME

X. MODEL IV. GROWTH MODEL ESTIMATING BMI PCT GROWTH TRAJECTORIES OF CHILDREN AND ADOLESCENTS FROM KINDERGARTEN TO GRADE 8 AFTER ADDING CUMULATIVE EXPOSURE TO INTERVENTION AND THE INTERACTION OF CUMULATIVE INTERVENTION EXPOSURE WITH TIME AND ALL SIGNIFICANT DEMOGRAPHIC PREDICTORS FOR EACH PARAMETER
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Social-ecological framework of the PEP intervention</td>
</tr>
<tr>
<td>2.</td>
<td>BMI surveillance system logic model</td>
</tr>
<tr>
<td>3.</td>
<td>Controlled setting assessment: Bland-Altman plot of PE teacher-rater height measurements</td>
</tr>
<tr>
<td>4.</td>
<td>Controlled setting assessment: Bland-Altman plot of PE teacher-rater weight measurements</td>
</tr>
<tr>
<td>5.</td>
<td>Natural setting assessment: Bland-Altman plot of PE teacher-rater height measurements</td>
</tr>
<tr>
<td>6.</td>
<td>Natural setting assessment: Bland-Altman plot of PE teacher-rater weight measurements</td>
</tr>
<tr>
<td>7.</td>
<td>Cross-sectional and longitudinal relative odds of obesity over time for Black and African American students</td>
</tr>
<tr>
<td>8.</td>
<td>Cross-sectional and longitudinal relative odds of obesity over time by for Latino students</td>
</tr>
<tr>
<td>9.</td>
<td>Cross-sectional and longitudinal relative odds of obesity over time for Asian students</td>
</tr>
<tr>
<td>10.</td>
<td>Cross-sectional and longitudinal relative odds of obesity over time for socioeconomic status</td>
</tr>
<tr>
<td>11.</td>
<td>Cross-sectional and longitudinal relative odds of obesity over time for gender</td>
</tr>
</tbody>
</table>
### LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAP</td>
<td>The American Academy of Pediatrics</td>
</tr>
<tr>
<td>B-A plot</td>
<td>Bland-Altman Plot</td>
</tr>
<tr>
<td>BMI</td>
<td>Body Mass Index</td>
</tr>
<tr>
<td>BMI PCT</td>
<td>Body Mass Index Percentile</td>
</tr>
<tr>
<td>CBPR</td>
<td>Community-based Participatory Research</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CI</td>
<td>Confidence Interval</td>
</tr>
<tr>
<td>CPS</td>
<td>Cambridge Public Schools</td>
</tr>
<tr>
<td>EST</td>
<td>Ecological Systems Theory</td>
</tr>
<tr>
<td>HLM</td>
<td>Hierarchical Linear Modeling</td>
</tr>
<tr>
<td>ICC</td>
<td>Intraclass Correlation Coefficients</td>
</tr>
<tr>
<td>ICH</td>
<td>Institute of Community Health</td>
</tr>
<tr>
<td>IOM</td>
<td>Institute of Medicine</td>
</tr>
<tr>
<td>IRB</td>
<td>Institutional Review Board</td>
</tr>
<tr>
<td>NCHS</td>
<td>National Center for Health Statistics</td>
</tr>
<tr>
<td>NHANES</td>
<td>National Health and Nutrition Examination Survey</td>
</tr>
<tr>
<td>PE</td>
<td>Physical Education</td>
</tr>
<tr>
<td>SD</td>
<td>Standard Deviation</td>
</tr>
<tr>
<td>SES</td>
<td>Socioeconomic Status</td>
</tr>
</tbody>
</table>
SUMMARY

In an effort to address childhood and adolescent obesity prevention strategies, school-based body mass index (BMI) surveillance has been legislated in 30 states. This research project analyzed data from the Cambridge Public Schools (CPS) BMI surveillance system, the longest and most complete school-based BMI surveillance system in the country, in order to further the knowledge base, from a methodological perspective, of a best practice approach for the use of school-based BMI surveillance in the prevention of childhood and adolescent obesity. This project was divided into phases. Each phase is summarized by a manuscript: assessing measurement reliability, analysis, and correlation to simultaneous prevention efforts.

The first manuscript assesses the reliability of the BMI surveillance and screening system data. As part of the Physical Education (PE) curriculum, PE teachers have been responsible for annually collecting student height and weight measurements. Bland-Altman plots (B-A plots), mean absolute differences, and intraclass correlation coefficients (ICC) were used to estimate intra- and inter-rater reliability among PE teachers in a controlled setting and PE teacher versus expert inter-rater reliability in a natural classroom setting. Barriers to reliability were also qualitatively assessed. Results showed that according to ICC criteria, reliability of PE teachers’ measurements was “excellent,” but the criteria for mean absolute differences were not consistently met. Results therefore supported PE teacher height and weight measurements for reliable and practical school-based BMI surveillance and potentially for BMI screening if training, adequate staffing, and safeguards are built into the system to check for outliers and equipment problems.
SUMMARY (continued)

The second manuscript presented as part of this thesis explores best practice analyses of the data from the CPS BMI surveillance system. Differences were compared in the calculations of relative odds of childhood and adolescent obesity as measured by cross-sectional versus longitudinal analysis. Using all data collected over eight years from the sociodemographically diverse kindergarten to eighth grade student body, cross-sectional and longitudinal relative odds ratios of obesity (dichotomized: BMI<95th versus ≥95th percentile) at baseline and at two different intervention time points were compared. Cross-sectional relative odds were determined by logistic regression and longitudinal relative odds by hierarchical linear modeling (HLM). Differences were assessed by overlap of confidence intervals. Results showed that analyzing school-based obesity surveillance data cross-sectionally to report annual obesity may be reasonable; however, to track trends or to measure program intervention effects over time, cross-sectional analysis may not be a reliable method since, over time, it may misrepresent obesity prevalence.

The third manuscript follows the BMI percentile (PCT) trajectories of the diverse CPS student body of kindergarten to eighth grade students during an eight-year time period of ecological school-based participatory physical activity and nutrition intervention. Multilevel growth curve modeling for panel data was used to estimate students’ BMI PCT growth trajectories as a function of cumulative exposure to intervention while controlling for gender, race/ethnicity, and socioeconomic status (SES). The correlation between cumulative exposure to school-based intervention and decreased BMI PCT growth reinforces the need for longitudinal intervention and analysis and supports the role of schools in efforts to address childhood and adolescent obesity.
I. INTRODUCTION

A. **Problem Statement**

The first U.S. Department of Health and Human Services *Healthy People* disseminated as a report by the Surgeon General (Healthy People: The Surgeon General's Report on Health Promotion and Disease Prevention) in 1979 was primarily devoted to health and disease prevention as they related to smoking; there was no mention of childhood and adolescent obesity. (U.S. Department of Health, Education, and Welfare, 1979) However, from 1970 to 2000, the prevalence of childhood and adolescent obesity more than tripled. (Hedley et al., 2004; Joliffè, 2004; Ogden et al., 2012) By 2010, the rapid increase in obesity prevalence at the end of the twentieth century was seen to slow and possibly level. Nonetheless, the prevalence of obesity, as measured by BMI, remains unequivocally high despite vast and multifaceted research and prevention and interventional endeavors over the past three decades. (Ogden et al., 2010; Ogden et al., 2012)

The prevalence of high BMI among children and adolescents is currently estimated at 18.7% at or above the 95th percentile for age (benchmark to determine obese youth). (Ogden et al., 2012) Including those who are overweight (BMI ≥ 85th percentile for age) increases this number to include 33.2% of six to nineteen year-olds. (Ogden et al., 2012)

Most notable are the persistent disparities in childhood and adolescent obesity: minority children and adolescents, males overall, and those from lower SES families are more likely to be obese. (Delva et al., 2007; Dwyer et al., 1998; Freedman et al., 2006) Data from the National Health and Nutrition Examination Survey (NHANES) from 1999 to 2010 demonstrated significantly higher prevalence of obesity among Latino and Black and African American youth (Black refers to those who self-identify as Black such as those who were not born in the United
States, and African American refers to those who self-identify as African American), 22.9% and 25.7%, respectively, compared with non-Latino White youth, 15.2%; disparate effects are most dramatic among Latino boys, 25.3%, and non-Latino Black and African American girls, 26.1%, compared to 17.2% among non-Latino White boys and 13.0% among non-Latino White girls. (Crespo et al., 2001; Hanson & Chen, 2007; Madsen et al., 2010; Ogden et al., 2012) Thus, at the same time that there is evidence of the overall childhood and adolescent obesity epidemic leveling, trends in the data suggest that disparities in the prevalence of obesity are expected to continue to worsen over time, widening the race/ethnicity gap. (Madsen et al., 2010)

Regarding SES disparities, lower household income and lower parental education are both associated with higher rates of childhood and adolescent obesity. These two risk factors were found to explain as much as one third of obesity prevalence in the sample of adolescents participating in the National Longitudinal Study of Adolescent Health (Add Health). (Goodman, 1999; Goodman et al., 2003) Across school districts, the proportion of students eligible for free and reduced-price school meals, one index of SES, is a reliable predictor of childhood obesity rates. (Caprio et al., 2008) The most recent update of Healthy People, Healthy People 2020 published in 2010, includes 22 objectives dedicated to nutrition and weight status and an additional 15 objectives dedicated to physical activity with an overarching goal to eliminate disparities. (U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion, 2010)

Another facet of the persistently high rates of childhood and adolescent obesity is the notion that society has yet to view obesity as a societal problem to solve, unlike most other threats to children and adolescents. Public education and legal measures, such as toy and other children’s product recalls, bicycle helmet laws, and car seat use laws, safeguard the risks of
injury or death by helping to reduce rates of unintentional injuries. (Schwartz & Puhl, 2003)

When a child chokes on a toy from a fast food “kid’s meal,” society blames the maker of the toy, not the child or parent for allowing the child to eat the toy. However, when the child becomes obese from the same meal, society does blame the child or parent for allowing the child to eat the meal. Historically, prevention for obesity has primarily focused solely on behavior modification and thus blaming the victim (personal responsibility) mentality. (Minkler, 1999; Schwartz & Puhl, 2003) Despite the overall societal perception of the mechanisms of obesity, more recent public health prevention efforts have been more ecologically minded. The field and practitioners of public health, though, are challenged to equitably reconcile personal and collective responsibility approaches in ways that best serve the public good. (Brownell et al., 2010)

B. Significance of the Problem

The detriment of the impact of obese youth is two-fold. Firstly, the rapid increase in the prevalence of obesity at the end of the twentieth century resulting in unprecedented prevalence at the beginning of the twenty-first century has been coupled with an increasing prevalence of obesity-related disease surfacing in youth. What were once thought to be adult-onset diseases such as type 2 diabetes, hypertension, and hyperlipidemia in addition to obstructive sleep apnea, orthopedic disorders, and social and psychological effects, such as low self-esteem and discrimination, have now become prevalent in children and adolescents. (Dietz & Gortmaker, 2001; Schonfeld-Warden & Warden, 1997; Strauss, 2002) Even a high prevalence of the metabolic syndrome (a combination of medical disorders that, when occurring together, increase the risk of developing cardiovascular disease and diabetes) is now reported in children and adolescents. (Weiss et al., 2004) Approximately 60% of obese children have developed one cardiovascular risk factor, and about 20% have two or more. (Schwartz & Puhl, 2003)
Furthermore, up to 80% of obese youth become obese adults, especially if they are obese by age six. (Council on Sports, Medicine, and Fitness, & Council on School Health, 2006) The odds ratio for adult obesity is more than 20 times greater for obese children than for non-obese children. (Saha et al., 2005) Obese adults are, likewise, at increased risk for type 2 diabetes, hypertension, stroke as a result of hypertension, and cardiovascular disease, as well as for certain cancers, orthopedic problems, and respiratory problems. (Must et al., 1999; Ogden et al., 2010; Strauss, 2002; Styne, 2001) The rapid increase in childhood and adolescent obesity and associated cardiac risk factors will likely lead to a reversal of the past half century of decreasing trends in rates of cardiovascular disease. (Olshansky et al., 2005; Perrin et al., 2007) For the first time in history, researchers such as Olashanski et al. (2005) have calculated that children will have shorter life spans than their parents.

The second, and nonetheless important, detrimental impact of obesity is the economic burden placed on the health care system. In light of the Patient Protection and Affordable Care Act and steps toward greater universal health care coverage, health care cost containment is vitally essential. Considering the estimated $147 billion in medical costs for obesity per year, real cost savings and thus greater ability of near universal coverage are more likely to be achieved through reducing obesity and related risk factors. (Finkelstein et al., 2009)

C. **Purpose of the Proposed Research**

1. **Significance of the proposed research**

Interventions that promote healthy eating and physical activity during childhood and adolescence may, therefore, not only prevent some of the leading causes of death but also may decrease direct health care costs and improve quality of life. (Centers for Disease Control and Prevention, 1996) However, reversing what has been defined as a childhood and adolescent
obesity epidemic has in its own right become a seemingly insurmountable problem. Over the past three decades, effective prevention and intervention initiatives have been widely explored. Schools have been at the forefront of such childhood and adolescent obesity prevention and intervention research for a variety of powerful and plausible reasons, including the ready access to children and adolescents, the near universal enrollment, the amount of time children spend, the proportion of total daily energy consumed (35% to 40%), and the powerful influence of the physical, social, and normative environments. (French et al., 2003; Gortmaker et al., 1999; Gortmaker et al., 1999; Lytle et al., 2004) In addition, the near universal enrollment of children and adolescents in school secures a way to reach the often hard-to-reach underserved, low-income populations that suffer the greatest obesity disparities.

As an important obesity prevention strategy, the Institute of Medicine (IOM) and the American Academy of Pediatrics (AAP) recommend the routine tracking of BMI in all children and adolescents. (Stoddard et al., 2008) For the reasons mentioned above, schools are an ideal setting to track childhood and adolescent obesity. More precisely monitoring obesity prevalence, incidence, and trend data through longitudinal observation beginning in childhood and adolescence can be used to identify and track disparities, design and evaluate targeted interventions, and advocate for policy changes and funding. (Longjohn et al., 2010) Since 2003, approximately 30 states have either enacted or proposed school-based BMI surveillance laws and legislation. (Longjohn et al., 2010)

Though the Massachusetts Public Health Council did not approve legislation requiring all public schools to measure BMI and send individual student BMI results to families until 2009, the CPS in Massachusetts have monitored all kindergarten to eighth grade students’ BMI status since 2000 and have sent results home to families since 2003. The PE staff at CPS has been
responsible for conducting height and weight assessments on all kindergarten to eighth graders each spring as part of the PE curriculum. To our knowledge, this makes the CPS system the longest, most complete computerized school-based BMI database in the country to date.

Because height and weight measurement is part of the CPS PE curriculum, data from the CPS surveillance system does not have the limitations inherent in the few other studies that had at least eight years of longitudinal data: selection bias. Such a bias could result in an underestimation of obesity prevalence since heavier children and adolescents may be less inclined to volunteer for a study in which repeated measurements of body size would take place. (Saha et al., 2005)

In addition, the CPS Controlled Choice Plan includes in its school assignment policy SES, determined by free and reduced-price meals eligibility status, and race/ethnicity in order to have the multi-faceted diversity in each school that will provide all students with equitable educational opportunities and with improved achievement. (Cambridge Public Schools, 2001)

Because a variety of diversity factors in addition to neighborhood proximity are used for school assignment, CPS has established not only ethnic and racial diversity in each school, but also diversity in linguistic and socioeconomic status of the CPS student population. Because exposure to school-based obesity prevention in CPS is inherently inclusive of a diverse population, this environment may be ideal to permit an understanding of the racial/ethnic and socioeconomic disparities in obesity.

This BMI surveillance research project is the first to include nearly a full elementary and middle school system population. In addition to including a larger and more complete sample size, this study analyzes these data longitudinally—not cross-sectionally as has been performed by the few studies available from other states where school-based BMI surveillance and
screening are legislated. (Madsen et al., 2010; Centers for Disease Control and Prevention, 2006) Cross-sectional analysis of longitudinal data may misrepresent increases or decreases in obesity prevalence over time. As an example, CPS has a dynamic population. From the 1999–2000 school year to the 2007–2008 school year, the pre-kindergarten to twelfth grade population decreased from 7,294 to 5,682, witnessing an annual average 4% decline in population. (Massachusetts Department of Elementary and Secondary Education, 2011) Analyzing BMI surveillance system data cross-sectionally to determine obesity prevalence among CPS students and to evaluate the physical activity and nutrition interventions during this time period may not accurately represent change in obesity prevalence since cross-sectional analysis does not account for the dynamic school population. Overall obesity prevalence rates and, in turn, intervention impact may therefore be misrepresented, especially since sociodemographic differences between the population that has left the district and the population that has remained cannot be tracked cross-sectionally. Furthermore, cross-sectional data are not useful to explain effects of age or of a cohort shift in risk.

In 2000, CPS BMI surveillance was added as an obesity prevention strategy (screening with notification was included in 2003). Over the following eight years, nine BMI time points, supported by various and substantial grants, CPS longitudinally implemented system-wide ecologically minded obesity intervention. Therefore, longitudinally studying growth trajectories among a diverse child and adolescent population from this unique CPS, now mandated, surveillance system could have a significant impact on informing obesity prevention programming and policy.
2. **Specific aims**

**Specific aim 1:** To assess the reliability of CPS BMI surveillance data by evaluating the reliability of height and weight measurements collected by CPS PE teachers.

**Specific aim 2:** To evaluate the differences between relative odds of obesity determined cross-sectionally and longitudinally to inform best analysis practices for school-based BMI surveillance.

**Specific aim 3:** To longitudinally analyze data from the CPS BMI surveillance database in order to assess CPS students’ BMI percentile (PCT) trajectories over an eight-year period of ecological school-based intervention, as well as to assess sociodemographic differences in these trajectories.
II. CONCEPTUAL FRAMEWORK

A. A Social-Ecological Approach: Overview of the Cambridge Body Mass Index

**Surveillance System Framework**

As an obesity prevention strategy, BMI surveillance in CPS was initiated in 2000 (and screening with notification in 2003). By the 2002–2003 school year and continuing through the 2007–2008 school year, BMI surveillance and screening was funded, improved, and then sustained through a combination of components of seven substantial physical activity and nutrition grants from various funders. The funds from these grants primarily helped CPS overhaul their PE program, the extracurricular physical activity opportunities offered, the school nutrition environment, the nutrition curriculum, related school policy, and related family activities and involvement (this overall effort was referred to as, “Healthy Living Cambridge Kids”).

With these grants, the school district has taken an ecological approach to reducing childhood and adolescent overweight and obesity among students enrolled in CPS. Figure 1 applies the Social-Ecological model first proposed by Bronfenbrenner and later by McLeroy et al. (1988) to the Healthy Living Cambridge Kids intervention. As seen in Figure 1, Healthy Living Cambridge Kids was a multi-level initiative aiming to alter the macrosystem via school policy, the physical environment through healthy food and physical activity availability and the elimination of non-nutritious foods, and the social environment by providing faculty and student learning experiences. Incorporating BMI surveillance and screening into the PE curriculum fits into this model at the school policy level and, as of 2009, at the state policy level. Since 2000, as part of the PE curriculum, all students in kindergarten to eighth grade have been required to have their heights and weights measured in PE class unless parents sign a passive consent form stating
that they do not want their children to participate. In 2009, the Massachusetts Public Health Council approved regulation requiring all public schools to measure BMI and send individual student BMI results to families of students in first, fourth, seventh, and tenth grades.
Figure 1. Social-ecological framework of the PEP intervention.
B. **Environmental Rationale for the Dramatic Rise in the Prevalence of Obesity**

An ecological approach to reducing the prevalence of obesity is warranted by the environmental etiology of the dramatic rise in obesity at the end of the twentieth century and subsequent sustained high prevalence. Though an exact reason for why the prevalence of obesity increased so suddenly and markedly is unknown, researchers agree that it is not likely that the dramatic rise was caused by a sudden shift in genetic or biologic factors: genetic shifts take much longer to occur. (Perrin et al., 2007) Instead, the vast body of obesity research shows that the marked increase of obesity involves a complex interaction between multiple factors at each level of the Ecological model. (Davison & Birch, 2001) Ecological Systems Theory (EST) states that individual change is not distinct from the context, or “ecological niche” in which one is embedded. (Davison & Birch, 2001) For a child, an ecological niche would include both his/her own personal attributes, such as gender and age, which are influenced by the family and school in which he/she is embedded, which in turn are influenced by the community and society in which the family and school are embedded. Thus, to promote health, an individual’s ecological niche must offer economic and social conditions that are conducive to health and healthy lifestyles since the environment controls or sets limits on behavior that occurs in it. (Green et al., 1996)

The interaction of this web of contributing factors has negatively altered the two primary non-genetic determinants of obesity: food (energy intake) and physical activity (energy expenditure). Though consuming a nutritious and portion controlled diet, being physically active, and avoiding sedentary activities are individual health behaviors, EST shows that the greater structural environment has had an overpowering influence on the foods that people choose to consume and the amount of time that they spend being physically active and inactive.
(Hill & Peters, 1998) Ultimately, the present economic and social conditions have created an environment that has promoted behaviors which have caused energy intakes greater than energy needs coupled with low energy expenditure. (Hill & Peters, 1998) Less space and opportunity for physical activity due to contemporary amorphous designs of neighborhoods, more unsafe neighborhoods, advances in technology and transportation, and cutbacks in mandatory PE classes are some examples of the structural changes that have altered the amount of time people spend being physically active. (Hill & Peters, 1998; Hill et al., 2004; Perrin et al., 2007) Greater availability and affordability of non-nutritious foods has led to greater portion sizes. (Hill & Peters, 1998; Hill et al., 2004; Perrin et al., 2007) There has also been a multi-million dollar increase in media advertisement of non-nutritious foods at the same time as the affordability of nutritious foods like fruits and vegetables has decreased. (Hill & Peters, 1998; Hill et al., 2004; Perrin et al., 2007) Such changes have negatively influenced energy intake. Structural changes like the ones mentioned have hindered the reduction of obesity prevalence rates. (Hill & Peters, 1998; Hill et al., 2004; Perrin et al., 2007) These contextual changes are compounded by more fundamental changes such as greater inequality of household incomes and greater racial and ethnic diversity coupled with lingering structural racism. (Perrin et al., 2007) Figure 1 thus shows how the Healthy Living Cambridge Kids intervention intervened at each level of a CPS student’s ecological niche and how factors from outer rings moderate the influence of factors from inner rings.

C. **Philosophical Underpinnings of the Ecological Model**

The socioeconomic and racial/ethnic disparity in the distribution of obesity supports the notion that changes in the structural environment causing changes in access to nutritious foods and physical activity have played a large role in the continuous rise of the prevalence of obesity.
Those at increased vulnerability, reflected by lower levels of power and resources, are automatically at an increased risk of obesity. (Sorensen et al., 2003) It is no coincidence that low-income neighborhoods are more unsafe, have fewer grocery stores, and more fast food restaurants than higher income neighborhoods and that residents of such low-income neighborhoods are more likely to be obese and to have obesity-related chronic disease. Also, minorities disproportionately reside in these neighborhoods. (Mari Gallagher Research & Consulting Group, 2006)

Another way to understand how the environment influences individual behavior is through Max Weber’s theory of the power of life chances over individual choices. Weber defines “chance” to be the probability of the occurrence of events which bring about satisfaction. The probability of acquiring satisfaction is based on structural conditions such as income, property, and the opportunity for profit, as well as on rights, norms, and social relationships. Weber asserted that “choice” is then based on what is structurally possible. Ultimately, based on Weber’s theory, people are not entirely free to determine their lifestyle but have the freedom to choose within their social constraints. Theorists after Weber included gender, age, and race/ethnicity as part of the structural conditions that influence life chances and choices. (Cockerham et al., 1997)

For the vast majority of people in this society, the life circumstances leading to poor health outcomes are not a matter of choice but instead are a matter of social and economic circumstances into which they were born. Low-income living conditions are a limiting social context, which also, in turn, limits the impact of individual actions. (Brown & Margo, 1978) People at progressively lower SES have correspondingly less opportunity to and feel that they are less able to control the circumstances and events that affect their lives. (Minkler, 1999)
Thus, a behavior-focused framework implicitly, and often, explicitly, reinforces a victim-blaming etiology by ignoring the social context in which individual decision-making and health-related actions take place. (Brown & Margo, 1978; Minkler, 1999)

Altering the environment and introducing more affordable and nutritious foods and safe places to exercise provides individuals with the opportunity, or chance, to make healthier behavioral choices. With greater access, more individuals are more likely to choose these healthier food and physical activity options. Because food intake and physical activity are directly correlated to weight, having access to and consuming more nutritious foods and being more physically active will ultimately contribute to decreasing obesity. Thus, macro-level or environmental interventions grounded in the notion of social in addition to personal responsibility for health can exert a powerful effect in changing behavior on a broad scale. (Minkler, 1999)

D. **Specific Uses of Body Mass Index as Part of an Ecological Model**

Weight status data from a BMI surveillance system is valuable at numerous levels under the ecological umbrella and is therefore becoming a universal standard in the effort to prevent the continuous rise of obese children and adolescents. The CPS BMI Surveillance System is likely the longest and most complete school-based BMI database in the country. The system was started in 2000, years before any BMI surveillance legislation. Figure 2 adapted from the Childhood Obesity Working Group of the U.S. Preventive Services Task Force is an analytical framework showing how BMI surveillance works as an obesity prevention and intervention strategy. (Moyer et al., 2005)
1. **Individual level information**

The CPS BMI surveillance system is primarily used to gather individual level information (screening). Using a surveillance system at the individual level makes it possible to
identify overweight and obese students at onset as well as to target individual students with prevention and intervention efforts before obesity becomes established. (Kim et al., 2005a)

When documentation of high BMI occurs, screening, counseling, and referral rates for overweight and obese children and adolescents increase. (Perrin et al., 2004) Individualized health and fitness “progress reports” are generated from the surveillance and screening system and sent home to inform parents/guardians of their child’s weight status. The information is intended to motivate parents/guardians to work with their child’s pediatrician for weight control and to take positive lifestyle steps to promote healthy weight. (Chomitz et al., 2003)

2. **Population level information**

The CPS BMI Surveillance and Screening System has also been used to calculate yearly cross-sectional overweight and obesity prevalence rates. The breadth and longevity of the surveillance and screening system makes it a highly valuable tool for informing public health programming and policy. At the population level, ongoing school-based BMI surveillance can potentially close persistent gaps in knowledge of childhood and adolescent obesity. Tracking BMI trajectories as well as prevalence, incidence, and remission rates over time will further our understanding of the natural history of childhood and adolescent obesity as well as provide the necessary data for rigorous program evaluation.

The information collected in an ongoing, mandated surveillance system could be used to compare obesity status with information from surveillance systems in other locales, such as other schools, other communities, and even other states. This level and consistency of surveillance could also lead to the development of school-wide or community-wide efforts or even national efforts at improving or reducing factors contributing to rising rates of obesity. (Ikeda et al., 2006)
Furthermore, BMI surveillance can serve to assess the nutritional status of populations and to evaluate policy. (Chomitz et al., 2003)

E. **A Best Practice Methodological Framework for Body Mass Index Surveillance**

**Analysis**

At present, there is not yet a recommended best practice approach to analyzing BMI surveillance system data to assess the prevalence and incidence of overweight and obesity and monitor their trends over time in children and adolescents. (Caroli et al., 2007) Along with legislation of school-based BMI surveillance, developing a methodological best practice approach to surveillance system analysis will be essential in order to most precisely achieve BMI surveillance objectives. Methodological discrepancies can be found in BMI measurement reliability as well as in the analysis of prevalence, incidence, and trends over time. Regarding BMI reliability, poor precision in measurement of an anthropometric variable will lead to not only inaccurate BMI calculations and inaccurate estimations of obesity incidence, prevalence, and trends, but also to an underestimation of correlations to other variables. (Nordhamn et al., 2000) Regarding estimating obesity incidence, prevalence, and their trends, implications of analyzing BMI surveillance data cross-sectionally versus longitudinally must be determined. In so doing, a methodological best practice framework for BMI surveillance analysis can ultimately be established. This thesis seeks to use standard evaluation and statistical methods as the framework to determine best analytical practices for BMI surveillance. These will include anthropometry reliability testing, the effects of cross-sectional versus longitudinal analysis on BMI outcomes, and trends over time.
III. REVIEW OF RELATED LITERATURE

A. The Role of School-Based Obesity Prevention and Intervention

School-based interventions may be an ideal target for reducing childhood and adolescent obesity. The two primary determinants of obesity, dietary intake and physical activity patterns, are developed during school-age years. Schools have near universal enrollment (greater than 95% of all children and adolescents) and offer ready access to an intervention. (Carter, 2002; French et al., 2003; Gortmaker et al., 1999; Gortmaker et al., 1999; Lytle et al., 2004; Story, 1999) In addition, the amount of time children spend and the proportion of total daily energy consumed (35% to 40%) within the school environment offer meaningful targets for intervention. (Carter, 2002; French et al., 2003; Gortmaker et al., 1999; Gortmaker et al., 1999; Lytle et al., 2004; Story, 1999) Also, the powerful influence of the physical, social, and normative school environments can be directed towards ultimately reducing obesity prevalence. (Carter, 2002; French et al., 2003; Gortmaker et al., 1999; Gortmaker et al., 1999; Lytle et al., 2004; Story, 1999)

Most schools nationwide participate in the National School Lunch and Breakfast Programs. These programs annually serve roughly five billion lunches and 1.6 billion breakfasts. (Peterson & Fox, 2007) The National School Lunch Program also provides more than 164 million after-school snacks to children and adolescents. (Peterson & Fox, 2007) Schools also provide ready access to physical activity resources—such as gyms, equipment and outdoor playing fields—and physical education programs, as well as to school nurses who can provide screening and counseling. (Story, 1999) However, despite the potential to have a large impact on the determinants of obesity, results of school-based obesity prevention and intervention evaluation are variable at best. Since 1990 and the beginning of school-based prevention and
intervention evaluation, no conclusive evidence regarding the benefit of school-based efforts has been established. (Kropski et al., 2008) Evaluation of individual school-based interventions has shown widely varying results leading to conclusions of insufficient evidence to recommend school-based obesity prevention and intervention. (Gonzalez-Suarez et al., 2009; Kropski et al., 2008)

The Guide to Community Preventive Services (Community Guide) presents systematic reviews of the effectiveness of selected population-based interventions aimed at promoting healthy growth and development in children and adolescents (Centers for Disease Control and Prevention, 2011). School-based interventions were one of their four foci. Upon completion of the review, the investigators likewise concluded that despite the large body of research regarding school-based approaches to reducing childhood and adolescent obesity, obesity status has not meaningfully improved and that more evidence is needed to determine the effectiveness of school-based programs to control obesity. (Guide to Community Preventive Services, 2011)

There have been numerous reviews, meta-analyses, and syntheses of school-based obesity prevention and intervention efforts, and none of these studies have found conclusive results. In a recent synthesis published in 2012 of five systematic reviews and three meta-analyses of school-based behavioral interventions for controlling and preventing obesity, Khabalia, Dickinson et al. concluded that limited evidence was found on which to base recommendations. (Khabalia et al., 2012)

On the surface, the minimal effects of school-based initiatives point to the lack of effective interventions; however, the lack of informative congruent results may have more to do with methodology than with statistical significance. The inconsistency of school-based obesity intervention outcome evaluation is in part attributed to the lack of a commonly accepted
convention for reporting results, including a standardized set of outcome measures. (Baranowski et al., 2002; Gonzalez-Suarez et al., 2009; Kropski et al., 2008; McGraw, 2000) Studies are also heterogeneous in design, participants, and intervention. As no single intervention will fit all schools and populations, further high-quality research needs to focus on identifying specific program characteristics predictive of success. (Khabalia et al., 2012) Another systematic review similarly found that there is lack of insight into what interventions work for whom and that future studies should apply stronger methodology to test moderating effects of important potential target group segmentations. (Yildirim et al., 2011) In a systematic review by Kropski et al. of all research published on this issue since 1990, the authors only identified fourteen studies that were of sufficient quality to draw any conclusions about best practices for school-based obesity prevention and intervention. (Kropski et al., 2008)

Another critical methodological issue to consider in regard to the lack of conclusive school-based obesity prevention and intervention recommendations is that the more recent prevention efforts have been ecologically minded. Ecological approaches to public health problems cannot be evaluated as thoroughly as clinical approaches since the units of analysis do not lend themselves to random assignment into experimental and control groups. (Green et al., 1996) By definition, ecological interventions do not yield independent variables given the interdependence of individuals and their environments. Also, each ecosystem is a subsystem of a larger system. Evaluating the entire system is impossible. Pinpointing where in the hierarchy of subsystems to take a slice for analysis is difficult, especially since ecosystems have a dynamic nature thus making any set of observations time-dependent. (Green et al., 1996) Considering the methodological difficulties of school-based obesity prevention and intervention evaluation and despite the rather poor body of evidence, many investigators continue to believe that schools will
play an important role in stemming current trends in overweight and obesity in children and adolescents. While the variability in school-based obesity prevention and intervention evaluation results also underscores the many influences outside of schools, at the same time, schools certainly provide an important opportunity for prevention and intervention efforts. (Carter, 2002)

Within the inconsistent milieu of school-based obesity prevention and intervention evaluation results, investigators are in agreement about one recommendation: long-term follow-up is sparse but must be included in the evaluation. (Shaya et al., 2008; Story, 1999) Results of systematic reviews and meta-analyses of school-based obesity intervention reveal long-term multiple-level interventions seem to have the potential to be the most conclusively effective. (Gonzalez-Suarez et al., 2009; Kropski et al., 2008; Strauss, 2002) In a review conducted by Story, she found modest positive treatment effects for school-based high risk interventions, but only two of twelve interventions had follow-up data of at least six months. (Story, 1999) She also found that interventions with younger students appeared to be more successful than those with adolescents. (Story, 1999) Still another meta-analysis showed that longer-running programs were more effective than shorter programs. (Gonzalez-Suarez et al., 2009)

B. **Longitudinal Childhood and Adolescent Obesity Surveillance and the Role of Schools**

Over the course of the childhood and adolescent obesity epidemic, obesity status has primarily been determined from cross-sectional data derived from a handful of national studies. The NHANES survey, a complex, multistage probability sample of the U.S. civilian, non-institutionalized population, conducted by the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC) is the most frequently quoted data source for childhood and adolescent obesity. (Ogden et al., 2010; Yanovski & Yanovski, 2011)
Similarly, the majority of research studies, especially school-based obesity prevention and intervention, have either cross-sectionally derived their outcome measures of obesity such as BMI or have only longitudinally analyzed less than five years of data. (Gonzalez-Suarez et al., 2009; Kambalia et al., 2012; Kropski et al., 2008) This methodology has made measuring program effectiveness difficult. Changes in the prevalence of obesity can take many years to detect. (Yanovski & Yanovski, 2011) Therefore, longitudinal analysis of program effects might not yield changes over a short time period. Also, cross-sectional analysis cannot account for directionality such as obese students who are less obese or for the transient nature of schools where new students who were not exposed to school-based interventions and students whose weight improved but left the system could mask program effects and make the data appear relatively stable. (Weiner & Long, 2004) Furthermore, cross-sectional data do not allow for distinguishing age differences in obesity from secular trends. It took over 30 years to reach the current epidemic levels of obesity, yet most programs’ evaluations have not extended over three years. (Gill, 1997; Gonzalez-Suarez et al., 2009; Kambalia et al., 2012; Kropski et al., 2008; Yanovski & Yanovski, 2011) Most longitudinal evaluations to date have not followed participants long enough to detect meaningful change. (Gill, 1997; Kim et al., 2005a)

Hence, a school-based surveillance system could potentially fill the gap in longitudinal school-based program evaluation through monitoring intervention effects by tracking prevalence and incidence rates over a longer duration of time. Results from multiple years of surveillance can be used to determine how long it takes to detect changes in BMI and thus inform follow-up requirements of evaluation. In addition to informing intervention efforts, BMI surveillance results can also be used to inform policy.
School-based BMI surveillance has been propelled to the forefront of childhood and adolescent obesity research as a prevention strategy. In a 2004 report released by the IOM, they recommended that schools conduct annual assessments of each student’s height and weight and BMI PCT and make this information available to parents as a childhood obesity prevention initiative. (Kubik et al., 2007) The IOM report spurred great interest in school-based BMI surveillance. As a result, to date, approximately 30 states have either enacted or proposed school-based BMI surveillance laws and legislation. (Longjohn et al., 2010) School-based BMI surveillance serves as the necessary resource for longitudinal analysis because its comprehensiveness can provide the insight to fully understand the scope of childhood and adolescent obesity and to track progress made in combating this epidemic. (Longjohn et al., 2010; Thompson & Card-Higginson, 2009) Moreover, a BMI measurement is relatively easy, inexpensive, noninvasive, and time efficient unlike direct measures of body fat such as skinfold measures and underwater weighing. (Nihiser et al., 2007)

C. **Validity of Body Mass Index as a Measure of Body Fat**

The ratio of an individual's weight to height squared (kg/m²), BMI, is used to estimate risk of weight-related health problems. (Nihiser et al., 2007) Valid measures of total body fat mass, at best, are difficult to obtain. The methods considered the “gold standard” for body composition measurements, such as underwater weighing, magnetic resonance imaging, computed axial tomography, deuterium dilution, and dual-energy radiograph absorptiometry, are too cumbersome and costly for most public health investigations, especially community driven research. (Council on Sports, Medicine, and Fitness, & Council on School Health, 2006) However, BMI is an inexpensive body composition measurement and provides a valid indicator of trends over time. Therefore, for surveillance or screening purposes, the advantages of BMI outweigh its limitations. As a surveillance and screening tool, BMI is widely recommended.
However, BMI is not recommended as a tool for clinical decision diagnostics. The limitations of using BMI as a measure of obesity for surveillance and screening purposes must be understood.

The accuracy of BMI varies according to the degree of body fatness. Among children and adolescents, BMI is a good indicator of excess adiposity, but differences in the BMIs of children and adolescents with lower body fat percentage can be largely due to fat-free mass. (Freedman & Sherry, 2009) Although the accuracy of BMI in identifying children with excess body fatness depends on the chosen cut points, Freedman and Sherry found that a high BMI-for-age percentile has a moderately high (70%–80%) sensitivity and positive predictive value, along with a high specificity (95%). (Freedman & Sherry, 2009)

Children and adolescent BMI results need to be interpreted with caution because height, weight, bone mass, and percent body fat change at different times and rates during the growth spurts that characterize child and adolescent development, especially puberty. (Nihiser et al., 2007) Also, the association between percent body fat and BMI differs by sex (for an equivalent BMI, girls have greater amounts of body fat than boys), age (there is an increase in percent body fat with increasing age), and race (for equivalent BMI, Caucasians tend to have higher body fat than African Americans). (Daniels et al., 1997; Widhalm et al., 2001) Maturation stage is also correlated to BMI (BMI increases with increasing stage of maturation). (Daniels et al., 1997) In addition, children and adolescents who are of the same sex, age, race/ethnicity, and maturation stage can have a wide range of percent body fat values, as BMI cannot distinguish obesity due to excess fat mass from obesity due to excess lean mass. (Daniels et al., 1997; Widhalm et al., 2001) Overall, BMI has been found to be a highly specific but low to moderately sensitive test to measure body fat. (Malina & Katzmarzyk, 1999; Must & Anderson, 2006; Widhalm et al., 2001) When BMI is used as a measure of body fatness to provide a diagnosis of obesity,
considering maturation stage, race, and gender is important. (Daniels et al., 1997) However, using BMI for population surveillance is valid, cost effective, and efficient.
IV. METHODOLOGY AND RESULTS

A. Permission for Use of the Data and Database Cleaning

Secondary data from the Cambridge, Massachusetts Public Schools BMI Surveillance System were analyzed for this study. Through agreements with the Cambridge Public School Department, the Institute for Community Health (ICH) presently has access to the BMI Surveillance System datasets to help facilitate the health and fitness “report card” process and to report on the annual cross-sectional rates of underweight, normal weight, overweight, and obesity, both overall and sociodemographically stratified. Data for the present study were jointly obtained through CPS in the form of nine separate Microsoft Excel (Microsoft Corporation, Redmond, Washington) databases populated every year for nine years and from ICH in the form of nine SAS (SAS Institute Inc., Cary, North Carolina) databases imported from Excel. The SAS databases were cleaned and then merged for final analysis. Appendix A contains the code from all data cleaning as well as from all data analysis.

To access the data for the specific aims of the present study, a “Research Application” was completed and submitted to the deputy superintendent of CPS during an in person meeting. After discussing the request with the director of physical education, as well as with other administrators who were involved in accessing the sociodemographic data, CPS granted permission to use the data for the purposes of the present study. Appendix B contains the letter granting permission from the CPS deputy superintendent.

Approval was then obtained from both the Cambridge Health Alliance, the overseeing body of ICH, Institutional Review Board (IRB) and the University of Illinois at Chicago IRB (Appendix C). Because this study was restricted to secondary data analysis at the population level, risks to participants were minimal. Individual identifiers (student identification number) will remain
confidential and will never be reported or revealed for purposes of this study. Names were not included in any of the information received from ICH or CPS. Because of the long-standing relationship between CPS and ICH, the school administrators felt confident that the data would be used in the best interest of the students.

Given ICH’s commitment to participatory research, the specific aims of the present research have also been presented at various times to the nutrition and physical activity (5-2-1 Action Group) subcommittee of the Healthy Children Task Force, a coalition that unites Cambridge academia, grade schools, community organizations, city agencies, and the health care community to identify children’s health issues and to mobilize community partners to help improve these issues through programmatic research and policy action. These community stakeholders seemed to agree that the purpose of the present research study being proposed will ultimately benefit Cambridge school children, as well as the community at large.

The present research that follows explored the specific aims stated above in three different studies. The first two studies were conducted in order to determine and verify a best practice methodological approach for the third and principal study. The first study, therefore, evaluated the reliability of the kindergarten to eighth grade CPS students’ height and weight measurements collected by PE teachers. The second study explored differences in the calculations of relative odds of youth obesity as measured by cross-sectional compared to longitudinal analysis. The third and principal study then analyzed the kindergarten to eighth grade CPS students’ BMI PCT growth trajectories over an eight-year period of ecological school-based participatory physical activity and nutrition intervention. Methods, results, and conclusions are separately detailed in the three chapters that follow, respectively.
B. **Manuscript 1—Reliability of Height and Weight Measurements Collected by Physical Education Teachers for a School-Based Body Mass Index Surveillance and Screening System**

1. **Abstract**

   **Background.** As a strategy to address the persistent childhood and adolescent obesity epidemic, school-based BMI screening and surveillance is currently legislated in 30 states. In Cambridge, Massachusetts, PE teachers are responsible for collecting measurements. The purpose of this current research was to evaluate the reliability of height and weight measurements collected by CPS PE teachers.

   **Methods.** Using B-A plots, mean absolute differences, and ICCs, we estimated intra- and inter-rater reliability among PE teachers in a controlled setting and PE teacher- versus expert inter-rater reliability in a natural classroom setting. We also qualitatively assessed barriers to reliability.

   **Results.** Controlled setting: Out of 150 measurements, three height (2.0%) and two weight (1.33%) measurement outliers were detected; intra-rater mean absolute differences for height/weight were 0.52 inches (SD 1.61) and 0.8 lbs (SD 3.2); intra- and inter-rater height/weight ICCs were ≥ 0.96. Natural setting: Out of 105 measurements, one weight measurement outlier (0.9%) was detected; PE teacher- versus expert-rater mean absolute differences for height/weight were 0.22 inches (SD 0.21) and 0.7 lbs (SD 0.8), and ICCs were both 0.99. Equipment deficiencies, data recording issues, and lack of students’ preparation were identified as challenges to collecting reliable measurements.
**Conclusion.** According to ICC criteria, reliability of PE teachers’ measurements was “excellent.” However, the criteria for mean absolute differences were not consistently met. Results highlight the importance of staff training and data cleaning.

**Keywords.** Growth and development; obesity; school health services; public health; child and adolescent health

2. **Introduction**

The prevalence of childhood and adolescent obesity increased at alarming rates during the 1980’s and 1990’s and continued to stay high during the first decade of the 21st century. (Ogden et al., 2012; Olshansky et al., 2005) These trends threaten to diminish the health and life expectancy of current and future generations. (Olshansky et al., 2005) The IOM and the AAP recommend the routine tracking of BMI in all children and adolescents as an obesity prevention strategy. (Stoddard et al., 2008) School-based BMI surveillance and screening are approaches used to fulfill this recommendation. In public schools, BMI surveillance is intended to assess childhood obesity in a population, while BMI screening is intended to detect those at risk for weight-related health problems (for further clinical assessment) and to provide families with personalized health information about their child. (Nihiser et al., 2007) Starting with Arkansas in 2003, states have promoted and legislated school-based BMI surveillance and screening. (Longjohn et al., 2010; Nihiser et al., 2007) To date, approximately 30 states have either enacted or proposed school-based BMI surveillance laws and legislation. (Longjohn et al., 2010)

In states where BMI surveillance and screening legislation has been passed, data have been collected most typically by nurses and less often by PE teachers. (Nihiser et al., 2007) Few studies have reported data reliability. Height and weight measurement by school staff can serve...
as a sustainable method to monitor childhood and adolescent obesity. Since data could potentially be used broadly, understanding data reliability is critical.

In 2009, the Massachusetts Public Health Council approved regulation requiring all public schools to measure BMI and send individual student BMI results to families. As a result of this policy change, schools are challenged to implement this plan with existing staff and no new dollars for infrastructure. In Massachusetts, CPS have monitored kindergarten to eighth grade students’ BMIs and sent results home to families since 2000. The CPS PE staff has been responsible for conducting height and weight measurements from all kindergarten to eighth graders each spring as part of the PE curriculum. Prior to the data collection period each year, all CPS PE teachers attend a three-hour anthropometry, data entry, and sensitivity training session conducted by local public health professionals and PE administrators. Data are entered into the database, which is pre-programmed to indicate when implausible height and weight values have been entered. The database is also pre-programmed to automatically calculate BMIs and BMI-for-age percentiles based on the 2000 CDC growth charts. (Centers for Disease Control and Prevention, 2000) As an additional check, school nurses review data on all students categorized by CDC criteria as “underweight” or “obese” and, based on familiarity with students, may request a re-evaluation. Health Reports including individualized student BMI percentile information are then mailed home to students’ families.

Until the present study, there had been no rigorous assessment of the reliability of height and weight measurements collected by CPS PE teachers. The purpose of this study, therefore, was to evaluate the reliability of height and weight measurements collected by CPS PE teachers.
3. **Methods**

a. **Participants**

This study consisted of three components: a controlled setting assessment to estimate within (intra-rater) and between (inter-rater) rater reliability among CPS PE teachers; a natural classroom setting assessment to estimate inter-rater reliability between CPS PE teachers and an expert rater in the PE classroom measurement environment; and qualitative assessments of the measurement process to provide information for recommendations.

For the controlled setting assessment, we selected a purposive sample of 15 public school children and a randomly selected sample of five CPS PE teachers. Members of a local child health coalition were invited to enroll their kindergarten to eighth grade children in the study. A heterogeneous mix (in regard to weight, grade, age, race/ethnicity) of 15 kindergarten-eighth grade public school students was selected among 17 students with written parental consent.

During a mandatory meeting, five of 18 CPS PE teachers were randomly selected as study raters by drawing names from a hat. A sixth teacher who was randomly selected declined to participate in the study. One expert rater, a public health professional trained in anthropometry, participated as a comparison rater.

For the natural classroom setting assessment, PE teachers at the five largest CPS elementary schools were invited to participate. The PE teachers at one school declined, leaving eight teachers at four schools to participate. Students from grades one to four and seven were invited to participate through recruitment letters and informed consent sent twice to parents/legal guardians. Recruitment materials were translated into Haitian Creole, Spanish, and Brazilian Portuguese. Students who returned consents and were present on measurement day were
included in the study. The same trained expert rater from the controlled setting also participated in this assessment.

b. **Instruments**

The same school-issued wall-mounted stadiometers (*Seca 216 Accuhite, Snoqualmie, Washington*) and self-calibrating digital scales (*Seca 840 Bellissima-digital, Snoqualmie, Washington*) used annually by CPS PE teachers were also used to collect height and weight measurements, respectively, for this study.

c. **Procedures**

The controlled assessment consisted of a single measurement session conducted in winter, 2007 at one CPS school. Six measurement stations, with one rater assigned to each station, were set-up in the gymnasium. Two rounds of measurements allowed all raters to measure each participant’s height and weight in duplicate. Raters were blinded to his/her first measurements during the second round.

In the natural classroom setting assessment, first the PE teacher rater and then the expert rater measured each student’s weight and then height. The expert measured all students twice and was blinded to her first round of measurements.

A survey to assess PE teachers’ attitudes, self-efficacy, and perceived barriers to taking student height and weight measurements in the typical PE setting was administered to all 16 PE teachers attending a professional development meeting. Additional qualitative assessments included observation of CPS’ annual measurement sessions at each measurement site. Research staff collected information on equipment function, data recording, and students’ preparation for measurements.
d. **Data analysis**

All statistical analyses were performed using SAS 9.1 (SAS Institute Inc, Cary, North Carolina). Each student’s BMI was estimated using the average of his/her height and weight measurements collected by the PE and expert raters using the formula: \( \text{BMI} = \frac{\text{weight in lbs}}{(\text{height in inches})^2} \times 703 \). The English system of measurement was used to calculate BMI since CPS PE teachers record annual measurement data using the English system.

Measurement reliability was analyzed using three different statistical methodologies suggested in the literature. The Bland-Altman limits of agreement method (B-A plots) was used to visually assess rater measurement agreement, outliers, and trends in the data. The B-A plots were constructed by plotting the non-absolute difference of duplicate measurements by the average of the two measurements. (Bland & Altman, 1986) Mean absolute differences with standard deviations (SD) were calculated to quantitatively assess intra- and inter-rater measurement agreement. Anthropometry manuals suggest that a reliable difference between duplicate height and weight measurements taken on the same participant should not exceed 0.25 inches or 0.25 lbs, respectively. (Wisconsin Partnership for Activity and Nutrition, 2007) The mean absolute measurement differences from both assessments were compared to these established tolerances, but we adjusted our weight benchmark from \( \leq 0.25 \) lbs to \( \leq 0.2 \) lbs since our scales measured weight to the tenth rather than to the hundredth of a pound. The intraclass correlation coefficient (ICC) was used to estimate overall reliability of the measurements. (Fleiss, 1986) For each ICC point estimate, 95% confidence intervals (CIs) were calculated. To interpret ICC values, the following Shrout and Fleiss benchmarks were used: excellent (\( > 0.75 \)), fair to good (0.40 to 0.75), and poor reliability (\( < 0.40 \)). (Shrout & Fleiss, 1979)
The effect that outlier data (detected visually by the B-A plots) may have had on classifying students into appropriate weight status (CDC BMI) categories was estimated. (Centers for Disease Control and Prevention, 2000) Using the controlled setting data, “reliable BMI estimates” were calculated using the mean of the students’ reliable measurements (that is, excluding outlier values identified by the B-A plots) collected by the PE and expert raters, and “unreliable BMI estimates” were calculated using the outlier values. The differences between the “reliable” and “unreliable” BMI estimates were calculated for each student. The BMI-for-age percentiles for reliable and unreliable BMIs were calculated, and variations in percentile categories were examined. (Centers for Disease Control and Prevention, 2000) Using the natural classroom setting data, PE- and expert-measured BMI were calculated separately, and BMI-for-age percentiles were determined for each value. (Centers for Disease Control and Prevention, 2000) The differences between the PE- and expert-rater derived percentile categories were calculated by subtracting the PE teacher value from the expert value. The percentile category differences were then examined to determine if measurement variation affected students’ percentile categories.

The PE teacher survey results were analyzed using frequencies of item responses. Recorded observations from measurement sessions were categorized into salient themes related to challenges to reliable measurement collection.

4. **Results**

   a. **Controlled setting assessment**

      i. **Student sample characteristics**

         The student sample included nine boys and six girls who ranged in age from seven- to thirteen-years old and first to seventh grades. Their BMI scores ranged from
14.1 to 22.0, with a median of 17.6. The five PE teacher raters collected 75 height-weight measurement pairs in total, and the expert rater collected 15 height-weight measurement pairs.

ii. **Intra-rater reliability: visual assessment**

The B-A plots revealed that the differences between the PE teacher raters’ duplicate measurements were near zero, but three height and two weight measurement outliers were visually detected (Figure 3 and Figure 4). Of the height measurements, 83% of differences were ≤ 0.25 inches while 63% of the weight differences were ≤ 0.2 lbs. Of the five total outlier measurements observed, all were associated with a different student participant, but three were associated with one PE teacher rater. The PE teacher raters’ measurement differences were not linearly associated with the students’ average height or average weight (p < .05 for both).
Figure 3. Controlled setting assessment: Bland-Altman plot of PE teacher-rater height measurements.
iii. **Intra-rater reliability: mean absolute differences**

Compared with the referenced benchmarks, the mean absolute differences of the PE teacher raters’ duplicate height and weight measurements were not clinically tolerable: 0.52 inches (SD 1.61) and 0.8 lbs (SD 3.2), respectively. Excluding the outlier measurements, the mean absolute difference for heights reached clinically tolerable
levels, but weight measurement differences remained intolerable, with an absolute mean of 0.3 lbs. The expert rater’s mean absolute differences were 0.25 inches (SD 0.25) for height and 0.2 lbs (SD 0.3) for weight; thus, achieving clinically tolerable levels.

iv. **Intra-rater and inter-rater reliability: intraclass correlation coefficients**

Including all measurement values collected by PE teacher raters, intra- and inter-rater ICC values were all ≥ 0.96 for height and weight, demonstrating “excellent” reliability. Excluding outliers increased all ICC values to 1.00.

v. **Effect of measurement discrepancy on body mass index**

In total, the outlier rate was 2% (3/150) for height measurements and 1% (2/150) for weight measurements. Students’ BMI values calculated from the five outlier measurements caused BMIs to be under- and over-estimated (range of ∆BMI = -7.0 to 7.4), shifting BMI-for-age percentile categories for all five students with outlier measurements when compared to values derived from the expert rater’s measurements (Table I).
TABLE I
CONTROLLED SETTING ASSESSMENT:
EFFECT OF HEIGHT AND WEIGHT MEASUREMENT OUTLIERS ON CDC WEIGHT STATUS CATEGORY

<table>
<thead>
<tr>
<th>Δ BMI</th>
<th>BMI-for-age Percentile (Expert)</th>
<th>*CDC Weight Status Category (Expert)</th>
<th>BMI-for-age Percentile (PE)</th>
<th>CDC Weight Status Category (PE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-7.0</td>
<td>57.3</td>
<td>Healthy weight</td>
<td>95.0</td>
<td>Obese</td>
</tr>
<tr>
<td>7.4</td>
<td>79.2</td>
<td>Healthy weight</td>
<td>0</td>
<td>Underweight</td>
</tr>
<tr>
<td>-5.0</td>
<td>94.0</td>
<td>Overweight</td>
<td>98.6</td>
<td>Obese</td>
</tr>
<tr>
<td>2.0</td>
<td>8.4</td>
<td>Healthy weight</td>
<td>0.03</td>
<td>Underweight</td>
</tr>
<tr>
<td>5.0</td>
<td>71.7</td>
<td>Healthy weight</td>
<td>0.03</td>
<td>Underweight</td>
</tr>
</tbody>
</table>


b. **Natural classroom setting assessment**

i. **Student sample characteristics**

Of the 504 informed consent forms sent to parents/guardians, 149 forms were returned. Among these consented students, one refused to participate, and 44 were absent during data collection. Thus, the final sample was 105 students.

Participation across the four schools was fairly equal: 19%–29% of targeted students from their respective schools took part. The sample was skewed to younger students: 46% of study participants were in first and second grades, 45% were in third and fourth grades, and 9.5% were in seventh grade. The study sample was 59% female, and the median age was 9.3 years (range: seven to thirteen years). Median BMI was 17.2 (range: 16.5–19.0).
ii. **Inter-rater reliability: visual assessment**

The B-A plot of PE teacher rater and expert rater height measurements showed that PE teacher raters tended to measure students slightly taller than the expert rater (Figure 5). One weight outlier was identified visually. In addition, one PE-teacher rater systematically over-weighed students compared to the expert rater (Figure 6).
Figure 5. Natural setting assessment: Bland-Altman Plot of PE teacher-rater height measurements.
iii. **Inter-rater reliability: mean absolute differences**

Mean absolute differences of all inter-rater height and weight measurements were 0.22 (SD 0.21) inches and 0.7 lbs (SD 0.8), respectively; 62.5% of inter-rater height differences and 12.5% of inter-rater weight differences were equal to or below the clinical tolerance.
iv. **Inter-rater reliability: intraclass correlation coefficients**

The ICC's for inter-rater data demonstrated “excellent” reliability (ICC = 0.99 for height and weight).

v. **Effect of measurement discrepancy on body mass index**

The outlier rate for weight measurements was 0.9% (1/105), and no outliers occurred for height measurements. We also examined five weight discrepancies due to a rater who systematically over-weighed students. The weight discrepancies resulted in a difference of 0.6 to 3.0 percentile points between raters but did not change students’ BMI percentile category. In the case of the outlier (6.2 lbs), though the student’s BMI percentile category did not change, the child’s BMI percentile changed from 53.7 to 16.1.

c. **Physical education teacher survey and classroom observations**

In total, 16 of 18 Cambridge PE teachers completed the survey designed to assess PE teachers’ perceptions regarding taking BMI measurements in the classroom setting. All teachers were at least “moderately confident” in their ability to take accurate measurements. Most PE teachers believed that BMI surveillance was at least moderately important for children’s health (69%), was at least a moderately important job function (69%), that measurement training was helpful (75%), and that measurements taken in the PE setting were not difficult to perform (75%).

Observation revealed issues that could pose challenges to collecting reliable measurements: equipment problems such as missing stadiometer head pieces and scale inaccuracies, data recording issues such as inconsistent use of data collection forms and transcription errors, and lack of students’ preparation for measurements which included interference from students’ hairstyles and bulky clothing.
5. **Discussion**

The findings of this study suggest that PE teachers could provide reliable measures of height and weight. Our review of the literature regarding reliability standards did not uncover a “gold standard” for assessing the reliability of school-based surveillance and/or screening. This is one of the few studies to systematically examine the reliability of school staff in measuring height and weight, and the first, to our knowledge, that focuses on PE teacher data.

While the PE teachers in both studies met “excellent” reliability standards according to one reliability criterion (ICC values), they did not meet all clinical tolerance criteria set for the mean absolute differences. The controlled assessment mean absolute height difference of 0.52 inches did not meet the tolerance criteria of 0.25 inches, but the natural setting mean absolute height difference of 0.22 inches did meet the reliability standard set in this study. The mean absolute difference for weight measured by the PE teachers in the controlled and natural setting assessments were 0.8 lb and 0.7 lb, respectively, neither meeting the established tolerance goal of 0.2 lb.

In the only other study to assess the reliability of a school-based surveillance system, Stoddard et al. reported “good” absolute mean difference reliability for measurements taken by school nurses compared to those taken by trained research staff in two study groups. (Stoddard et al., 2008) Although the mean absolute difference “tolerable” weight reliability measurement criterion was different ($\leq 1.1$ lb) than the clinical criteria used in this study, Stoddard et al. calculated a mean absolute difference between nurse and expert height measurements of 0.15 inches in one group and 0.12 inches in the other and a weight measurement difference of 0.15 lbs in each study group. (Stoddard et al., 2008) This highlights the lack of reliability measurement standardization. The mean absolute difference for weight measured by the PE teachers in the
controlled and natural setting assessments, 0.8 lb and 0.7 lb, respectively, would both meet the 1.1 lb tolerance identified in the Stoddard et al. study. Based on observation, rater measurement discrepancies were caused by either systematic mis-measurement due to equipment error or by human error, including measurement recording and transcription errors.

Another obvious discrepancy between the present study and that of Stoddard was the accuracy of the digital scales. Stoddard et al. concluded that the use of digital scales enhanced reliable collection of weight measurement when measurement protocols were followed. (Stoddard et al., 2008) However, in our study, inaccurate digital scales seemed to contribute to a reduction in reliable weight collection. In the controlled setting, a three-pound weight difference was detected between the five different scales used to measure the children even though they were from the same manufacturer. This added error to our reliability estimates not caused by rater differences. In the Stoddard study, it is unclear if different raters used the same equipment. Although using different scales may contribute to inaccurate and unreliable measurements, it should be noted that most school districts would likely supply each school with its own scale(s) (as is the case in CPS).

That the measurements consistently met the ICC reliability criteria but not the clinical tolerance criteria suggests that it is important to consider the choice of criteria as well as the purpose, surveillance or screening, to which the criteria are to be applied. A few discrepant measures are less likely to impair the interpretation of BMI data used for surveillance or identification of at-risk populations. (Nihiser et al., 2007) However, for screening purposes, one discrepant measure reported back to families can provide false weight category positives or negatives information, and thus, may require a higher standard of accuracy and reliability.
In the natural classroom setting assessment study, the 0.9% outlier rate did not change any student’s weight status categorization based on BMI-for-age percentiles. (Centers for Disease Control and Prevention, 2000) In this case, neither reports of surveillance nor screening results of BMI categories would have been compromised. However, 1.7% of the individual-level BMI category reports based on measures collected from the controlled setting assessment would have been compromised. As more states legislate the use of BMI for surveillance and/or screening, it will be important to consider the level of “tolerance” or “rigor” expected for determining reliably of the data collected.

6. **Limitations**

There are several limitations to our study. The small number of teacher raters and students who participated may limit the generalizability of results. Since the sample of students for both assessments and teacher raters for the natural setting assessment were not random, selection bias may have affected the reliability estimates of this assessment. Among the students, overweight and obese children, who may be more difficult to measure, were underrepresented in the study sample, potentially inflating the reliability estimates. As mentioned above, raters did not use the same equipment, which introduced equipment variability into the assessment. Furthermore, the scales were not calibrated until after the study. However, this study represents a “real world” situation where different schools likely use different equipment of different levels of maintenance.

7. **Conclusion**

This study supports using PE teachers to measure height and weight for reliable and practical school-based BMI surveillance and potentially for screening if training, adequate staffing, and safeguards are built into the system to check for outliers and equipment problems.
Higher reliability standards are important when using individual level data to send personal information home to students and families.

8. Implications for school health

Schools are a valuable resource for childhood and adolescent obesity prevention; hence, states are increasingly legislating school-based BMI surveillance and screening. (Longjohn et al., 2010) Legislation has made school-based surveillance systems an important pathway for tracking the childhood and adolescent obesity epidemic. Using BMI surveillance is an important tool to assess progress in school as well as community interventions. (Dietz et al., 2009) Therefore, assessing measurement reliability is essential. This study supported the assertion that PE teachers can collect reliable measurements, particularly for surveillance purposes, but also reinforced the importance of taking further steps to continually improve measurement reliability. The following recommendations to maximize reliability of a school-based surveillance and screening system were compiled based on study results and field observations.

Comprehensive training: In addition to measurement protocols, training should include review of potential sources of error: measurement, equipment, recording, and transcribing. Training should also highlight accountability for data collection and results, including training teachers to scan data and recognize outliers and irregularities and the consequences of inaccurate or unreliable data collection, particularly when results are sent home to families.

Duplicate measures: Most height and weight measurement protocols recommend duplicate measures of a participant followed by a third measure if the discrepancy between the measures exceeds the set tolerance. If staffing is adequate, duplicate or triplicate measures would improve reliability.
Built in checks and balances: The data entry database should indicate when implausible height and weight values are entered. Other checks could include “an outlier check” by school nurses or other personnel to check for implausible results and to verify measurements for all students classified as underweight or obese.

Periodic reliability assessments: A reliability assessment such as the one presented here should be regularly conducted to monitor the reliability of the surveillance and screening program over time. This is particularly necessary if changes are made in the measurement protocols or if equipment is changed. Results of the assessment should be provided to the raters to demonstrate the challenges and importance of reliable and accurate data collection.

Regular equipment maintenance and calibration: Regular equipment maintenance, calibration, and upgrades would help to limit equipment related measurement errors.

Parent/guardian notification: Notification should be sent to parents regarding the date of measurement collection, including a reminder to send their child to school with hairstyles or clothing that will not impede measurement taking.

Teacher helpers: Since PE teachers must often measure students and teach class at the same time, help from school nurses, older students, parent volunteers, or interns from area colleges would be beneficial so the teacher can concentrate on the task of measuring students. The accuracy and reliability of measurement in the classroom setting are enhanced by using three people: one to measure, one to record, and one to manage the class.

9. Human participants approval statement

The research protocol and materials were approved by the Cambridge Health Alliance IRB in 2007 and by the University of Illinois at Chicago IRB in 2010.
C. Manuscript 2—School-Based Obesity Surveillance: Cross-Sectional Versus Longitudinal Analysis

1. Abstract

Objectives. Differences in the calculations of relative odds of youth obesity as measured by cross-sectional compared to longitudinal analysis were assessed for the purpose of recommending best analysis practices for school-based BMI surveillance.

Methods. Using data collected over eight years from a sociodemographically diverse kindergarten to eighth grade student body, we compared cross-sectional and longitudinal relative odds ratios of obesity (dichotomized: BMI < 95th versus ≥ 95th percentile) at baseline and at two different intervention time points. We determined cross-sectional relative odds by logistic regression and longitudinal relative odds by HLM. Differences were assessed by overlap of CI’s.

Results. Increasingly over time, cross-sectional analysis overestimated relative odds of obesity compared to longitudinal analysis. The relative odds of obesity for low-income students in 2008 was the only significant cross-sectional and longitudinal difference where the 95% CI’s did not overlap. The directionality of the cross-sectional and longitudinal relative odds of obesity over time was similar.

Conclusions. Analyzing school-based obesity surveillance data cross-sectionally to report annual obesity rates may be reasonable; however, to track trends in the obesity epidemic over time or to measure program intervention effects over time, cross-sectional analysis may not be a reliable method since it may overestimate the longitudinal rate. Longitudinal analysis naturally provides the most accurate results of longitudinal obesity surveillance.
2. **Introduction**

Since 1980, obesity measured by BMI for age at or above the 95th percentile has tripled among school-age children and adolescents, and prevalence remains high at 18.7%. (Ogden et al., 2010) Though the persistent increasing prevalence rate of childhood and adolescent obesity seems to have somewhat stabilized, there is no evidence of a declining trend. (Madsen et al., 2010) As a response to youth obesity, approximately 30 states since 2003 have either enacted or proposed school-based BMI surveillance laws and legislation that typically have mandated annual BMI measurements aggregated for state-level tracking and disaggregated as obesity screening results for families. (Longjohn et al., 2010) School-based BMI surveillance can be used to describe and track trends in the obesity epidemic over time among youth; to identify demographic or geographic subgroups at greatest risk of obesity to target intervention programs; to create local awareness of the extent of youth obesity thereby providing an impetus to improve policies, practices, and prevention and treatment services in communities; to evaluate the effects of school-based physical activity and nutrition programs and policies; and to monitor progress toward achieving related health objectives. (Longjohn et al., 2010; Nihiser et al., 2009) These uses of mandated statewide surveillance have the potential to positively impact youth obesity rates.

Cambridge, Massachusetts Public Schools have been systematically tracking BMI for all students in kindergarten to eighth grade since 2000. During this time, CPS has simultaneously implemented large-scale physical activity and nutrition interventions. (Chomitz et al., 2010) In conjunction with a research institution, CPS has had the capacity to annually analyze the BMI data cross-sectionally. (Kim et al., 2005b) The BMI data from Arkansas, the first state to legislate school-based BMI surveillance, have also been analyzed cross-sectionally. (Centers for
Disease Control and Prevention, 2006) A recent study reporting obesity prevalence data from the California school-based BMI surveillance system, likewise, analyzed obesity prevalence cross-sectionally. (Madsen et al., 2010) However, cross-sectional analysis of longitudinal data may misrepresent increases or decreases in obesity prevalence over time. As an example, CPS has a dynamic population. From the 1999–2000 school year to the 2007–2008 school year, the pre-kindergarten to twelfth grade population decreased from 7,294 to 5,682, witnessing an annual average decline of 4%. (Massachusetts Department of Elementary and Secondary Education, 2011) Analyzing CPS BMI surveillance system data cross-sectionally to determine obesity prevalence among CPS students and to evaluate the physical activity and nutrition interventions during this time period may not accurately represent change in obesity prevalence since cross-sectional analysis does not account for the dynamic school population. Overall obesity prevalence rates and in turn intervention impact may therefore be misrepresented, especially since sociodemographic differences between the population who has left the district and the population who has remained cannot be tracked, cross-sectionally. Furthermore, cross-sectional data are not useful to explain effects of age or of a cohort shift in risk.

The purpose of this study was to evaluate the differences between relative odds of obesity determined cross-sectionally and longitudinally in order to inform best analysis practices for school-based BMI surveillance. We hypothesize that the results of a longitudinal analysis are statistically different than those from a cross-sectional analysis. Because CPS instituted physical activity and nutrition interventions between 2001 and 2004 and then increased efforts between 2005 and 2008, these time periods were assessed using 2000 data as a baseline.
3. **Methods**
   
a. **Study setting and study population**

   Student-level data were obtained from CPS, a district with a diverse student body, via separate Microsoft Excel (Microsoft Corporation, Redmond, Washington) databases populated every year for nine years. Student identification numbers were used to link data since student names were not included in any database. The University of Illinois at Chicago and the Cambridge Health Alliance IRBs approved this study as an exempt protocol.

   As part of the elementary physical education curriculum of CPS, each spring since 2000, PE teachers measure heights and weights on all students in kindergarten through eighth grade. Prior to the data collection period each year, all CPS PE teachers attend a three-hour anthropometry, data entry, and sensitivity training session conducted by local public health professionals and PE administrators. Height and weight measurements are collected with school-issued wall-mounted stadiometers (*Seca 216 Accuhite, Snoqualmie, Washington*) and calibrating digital scales (*Seca 840 Bellissima-digital, Snoqualmie, Washington*) by one PE teacher per school often with assistance from a school nurse or another PE teacher. Height and weight are assessed with one measure each. Height is reported to the nearest quarter inch and weight to the nearest two-tenths of a pound. Teachers and research assistants enter data into a standardized, password-protected Microsoft Excel (Microsoft Corporation, Redmond, Washington) spreadsheet, which is pre-programmed to indicate when biologically implausible values have been entered using the CDC SAS protocol for biologically implausible values. (Division of Nutrition, 2007) The Excel database is also pre-programmed to automatically calculate BMIs and BMI-for-age percentiles based on the 2000 CDC growth charts. (Centers for Disease Control and Prevention, 2000) In addition to heights, weights, BMIs, and BMI-for-age
percentiles, each Excel database also included, populated from the CPS student database, students’ gender, self-identified race/ethnicity, date of birth, and National School Lunch Program eligibility status. A reliability study conducted by Shapiro Berkson et al. (Chapter IV.B.) revealed that BMI measurements by CPS PE teachers used for surveillance purposes were reliable.

Between 2000 and 2008, CPS underwent three distinct district-wide, school-based physical activity and nutrition intervention periods. From the 1999–2000 school-year to the 2000–2001 school-year, there were minimal levels of intervention; from the 2001–2002 school-year to the 2003–2004 school-year, CPS implemented a low-level intervention; and, from the 2004–2005 school-year to the 2007–2008 school-year, CPS implemented a high level intervention. The specific intervention components are detailed in Chomitz et al. (Chomitz et al., 2010)

b. **Study design**

The BMI z-scores were calculated in SAS 9.1.3 (SAS Institute Inc, Cary, North Carolina) using the CDC’s program, which is based on the 2000 gender-specific BMI-for-age growth charts. (Centers for Disease Control and Prevention, 2000) Each student’s BMI percentile was calculated from the BMI z-score and, consistent with CDC recommendations, weight status was dichotomized into greater than or equal to the 95th BMI percentile (obese) and less than the 95th BMI percentile (not obese). Other covariates included gender, race/ethnicity, age, SES using National School Lunch Program eligibility as a proxy, and study year. We controlled for study year in both longitudinal and cross-sectional models to account for any potential difference due to year in study. Year was a categorical predictor using 2000 as the reference year. The SES variable was dichotomized into low income (receiving free and/or
reduced lunch) and not low income (no assistance). The race/ethnicity variable included Asian, Black and African American, Latino, and White students. Students who self-identified as “other” race/ethnicity were excluded from analysis. White students were used as the reference group in both analyses.

c. **Statistical analysis**

A cross-sectional analysis was performed to determine relative odds of being obese at baseline (2000) and at the end of each intervention period (2004, 2008) using logistic regression analysis. Cross-sectional statistical analyses were generated using SAS software, Version 9.1.3 of the SAS system for Windows (SAS Institute Inc, Cary, North Carolina. To determine longitudinal relative odds of obesity at the end of each intervention period, HLM for repeated measures was performed. For the 2004 time point, all 2000–2004 measurements per student were included. Similarly for the 2008 time point, all 2000–2008 measurements per student were included. Since all measurements were included, the longitudinal analysis had greater statistical power than the cross-sectional analysis. Using HLM 7 (Scientific Software International, Chicago, Illinois, longitudinal odds ratios were estimated using Maximum Likelihood estimation. All models were adjusted for the previously described covariates. To determine any difference between cross-sectional and longitudinal relative odds, 95% confidence intervals were calculated and overlapping ranges between the two different analyses were assessed.

4. **Results**

All CPS students in the BMI surveillance database, ages four to seventeen years, with reliable height and weight measurements from 2000–2008 were included in analyses. On average over the study period, 2.3% of students per year were not measured due to either refusal
or absence on measurement days. Due to substantial school population attrition over the study period, there was a decreasing range of 4,847 students measured in 2000 to 3,739 students measured in 2008, yielding a total of 11,871 student records. Each student had between one to nine BMI measurements, with an average of 4.19 measurements, yielding 39,064 total BMI measurements. We linked BMI measurements by student ID for longitudinal analysis. There were a total of 1,519 missing height and/or weight values in addition to 28 other records that could not be included due to missing values such as age. Despite the Biologically Implausible Value (BIV) pre-programmed check in each Excel database, 97 students, or 0.25%, were excluded due to a biologically implausible BMIz (BMIz<-4). Of the remaining records, 237 with “other” race/ethnicity were also excluded from these analyses. Students classified in the “other” race/ethnicity category were not included in the study because this category comprised less than 5% of the sample at each study time point. After all data exclusions, there were 11,503 students with 37,113 total valid BMI measurements who were aged four to seventeen years included in the final analysis.

Just over half of the study sample was male (51.3%) at baseline and slightly decreased to 50.5% by 2008. At baseline, 41.6% of students were White, 32.7% were Black and African American, 13.6% were Latino, and 11.5% were Asian. Study sample demographics slightly shifted by the end of the study in 2008: 37.6% of students were White, 32.2% were Black and African American, 13.5% were Latino, and 12.5% were Asian. The proportion of low-income students increased from 38.4% at baseline to 42.5% of the population at the final time point. The characteristics of CPS students at each study time point are shown in Table II.
TABLE II
CPS STUDENT CHARACTERISTICS AT EACH STUDY TIME POINT

<table>
<thead>
<tr>
<th></th>
<th>2000 % (n)</th>
<th>2004 % (n)</th>
<th>2008 % (n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51.2 (2231)</td>
<td>51.6 (2220)</td>
<td>50.5 (1880)</td>
</tr>
<tr>
<td>Female</td>
<td>48.8 (2130)</td>
<td>48.4 (2079)</td>
<td>49.5 (1843)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black and African American</td>
<td>32.7 (1410)</td>
<td>36.9 (1582)</td>
<td>32.2 (1190)</td>
</tr>
<tr>
<td>Latino</td>
<td>13.6 (589)</td>
<td>14.5 (620)</td>
<td>13.5 (500)</td>
</tr>
<tr>
<td>White</td>
<td>41.6 (1798)</td>
<td>37.1 (1589)</td>
<td>37.6 (1391)</td>
</tr>
<tr>
<td>Asian</td>
<td>11.5 (496)</td>
<td>10.9 (465)</td>
<td>12.5 (461)</td>
</tr>
<tr>
<td>Population</td>
<td>4511</td>
<td>4560</td>
<td>3739</td>
</tr>
<tr>
<td>Lunch status</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low income (Free or reduced-price lunch)</td>
<td>38.4 (1631)</td>
<td>43.4 (1864)</td>
<td>42.5 (2123)</td>
</tr>
<tr>
<td>Higher Income (Paid lunch)</td>
<td>61.6 (2620)</td>
<td>56.6 (2438)</td>
<td>57.6 (1566)</td>
</tr>
</tbody>
</table>

At baseline in 2000, 19.55% of the student population was obese. The prevalence of obesity climbed to 20.96% in 2004 and decreased to 16.81% by the end of the study period in 2008. At baseline, 16.91% of White students were obese as were 24.04% of Black and African American students, 25.81% of Latino students, and 9.27% of Asian students. By the end of the study period, obesity among all racial/ethnic study populations decreased. Table III shows obesity prevalence at each study time point as well as racial/ethnic specific obesity prevalence at each study point.
<table>
<thead>
<tr>
<th></th>
<th>2000</th>
<th>2004</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% (n)</td>
<td>% (n)</td>
<td>% (n)</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.34 (102)</td>
<td>1.58 (68)</td>
<td>2.18 (81)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>60.76 (2648)</td>
<td>59.12 (2544)</td>
<td>64.28 (2393)</td>
</tr>
<tr>
<td>Overweight</td>
<td>17.35 (756)</td>
<td>18.34 (789)</td>
<td>16.73 (623)</td>
</tr>
<tr>
<td>Obese</td>
<td>19.55 (852)</td>
<td>20.96 (902)</td>
<td>16.81 (626)</td>
</tr>
<tr>
<td><strong>White</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.67 (48)</td>
<td>1.83 (29)</td>
<td>1.51 (39)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>63.79 (1147)</td>
<td>65.96 (1048)</td>
<td>72.97 (1015)</td>
</tr>
<tr>
<td>Overweight</td>
<td>16.63 (299)</td>
<td>16.61 (264)</td>
<td>13.87 (193)</td>
</tr>
<tr>
<td>Obese</td>
<td>16.91 (304)</td>
<td>15.61 (248)</td>
<td>10.35 (144)</td>
</tr>
<tr>
<td><strong>Black and African American</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>1.77 (25)</td>
<td>1.33 (21)</td>
<td>1.51 (18)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>55.74 (786)</td>
<td>52.72 (834)</td>
<td>55.13 (656)</td>
</tr>
<tr>
<td>Overweight</td>
<td>18.44 (260)</td>
<td>19.34 (306)</td>
<td>19.83 (236)</td>
</tr>
<tr>
<td>Obese</td>
<td>24.04 (339)</td>
<td>26.61 (421)</td>
<td>23.53 (280)</td>
</tr>
<tr>
<td><strong>Latino</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>1.36 (8)</td>
<td>0.48 (3)</td>
<td>0.80 (4)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>55.18 (325)</td>
<td>51.61 (320)</td>
<td>54.40 (272)</td>
</tr>
<tr>
<td>Overweight</td>
<td>17.66 (104)</td>
<td>20.00 (124)</td>
<td>20.80 (104)</td>
</tr>
<tr>
<td>Obese</td>
<td>25.81 (152)</td>
<td>27.90 (173)</td>
<td>24.00 (120)</td>
</tr>
<tr>
<td><strong>Asian</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>4.03 (20)</td>
<td>2.58 (12)</td>
<td>3.47 (16)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>70.77 (351)</td>
<td>67.96 (316)</td>
<td>72.23 (333)</td>
</tr>
<tr>
<td>Overweight</td>
<td>15.93 (79)</td>
<td>17.85 (83)</td>
<td>12.15 (56)</td>
</tr>
<tr>
<td>Obese</td>
<td>9.27 (46)</td>
<td>11.61 (54)</td>
<td>12.15 (56)</td>
</tr>
</tbody>
</table>
Table IV presents the results of cross-sectional analysis of the relative odds of being obese at the end of each intervention period.

### TABLE IV
CROSS-SECTIONAL ODDS OF OBESITY AT KEY INTERVENTION TIME POINTS

<table>
<thead>
<tr>
<th>YEAR</th>
<th>2000 Baseline (Confidence Interval)</th>
<th>2004 End of Phase I (Confidence Interval)</th>
<th>2008 End of Phase II (Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>1.350 (1.157, 1.577)*</td>
<td>1.038 (0.892, 1.207)</td>
<td>1.166 (0.973, 1.396)</td>
</tr>
<tr>
<td>Low income**</td>
<td>1.426 (1.203, 1.691)*</td>
<td>1.497 (1.273, 1.761)*</td>
<td>1.827 (1.502, 2.223)*</td>
</tr>
<tr>
<td>Age</td>
<td>1.014 (0.974, 1.055)</td>
<td>1.023 (0.981, 1.066)</td>
<td>1.031 (0.982, 1.082)</td>
</tr>
<tr>
<td>Black/African American</td>
<td>1.376 (1.140, 1.660)*</td>
<td>1.703 (1.408, 2.059)*</td>
<td>2.037 (1.609, 2.579)*</td>
</tr>
<tr>
<td>Latino</td>
<td>1.513 (1.198, 1.911)*</td>
<td>1.796 (1.421, 2.271)*</td>
<td>2.135 (1.609, 2.832)*</td>
</tr>
<tr>
<td>Asian</td>
<td>0.514 (0.369, 0.717)*</td>
<td>0.719 (0.522, 0.990)*</td>
<td>1.080 (0.773, 1.508)</td>
</tr>
</tbody>
</table>

*p<0.05

**Low income was determined by using National School Lunch Program eligibility as a proxy and includes those receiving free and reduced lunch combined.

Male students, Black and African American students and Latino students compared to White students, and low-income students had the highest relative odds of being obese at every time point. At baseline, Asian students had the lowest relative odds of being obese. However, relative odds of obesity among Asian students steadily increased through 2008. Relative odds of being obese increased at the end of each intervention period for each of these populations except for male students whose relative odds of being obese decreased in 2004 and slightly increased in 2008 but were still lower than baseline. Table V presents the longitudinal results of the relative odds of being obese at the end of each intervention period.
Similar to the cross-sectional analysis, the longitudinal relative odds of being obese were highest at every time point for male students compared to female students, Black and African American students and Latino students compared to White students, and low-income students compared to higher-income students. However, the rate of increase in the longitudinal relative odds of obesity at the end of each intervention point was more moderate for Black and African American, Latino, and Asian students. In addition, relative odds of obesity for male students steadily decreased. Contrary to the cross-sectional findings, relative odds of being obese for low-income students decreased and then stabilized.

For each racial/ethnic variable studied, the directionality of the cross-sectional and longitudinal relative odds of obesity over time was identical: compared to White students, relative odds of obesity for Black and African American (Figure 7), Latino (Figure 8), and Asian students (Figure 9) increased. However, cross-sectional and longitudinal relative odds ratios differed. In all instances, cross-sectional relative odds of obesity were higher than longitudinal
relative odds of obesity. Figures 7–9 clearly show the steeper cross-sectional rate of increase. Though the relative odds of obesity for low-income students (Figure 10) in 2008 was the only significant cross-sectional and longitudinal difference where the 95% CI’s did not overlap, the CI’s of the cross-sectional relative odds ratios were much wider than the longitudinal relative odds ratios due to power differences. Though the 95% CI’s overlapped, the difference in the Asian students’ relative odds of obesity was most notable. Longitudinally, Asian students’ relative odds of obesity slightly increased but were still less than the relative odds of obesity for White students. However, cross-sectionally, the relative odds of obesity among Asian students increased and slightly surpassed the relative odds of obesity among White students.

The directionality of the relative odds of obesity over time differed between the SES and gender (Figure 11) cross-sectional and longitudinal analyses. For low-income students, cross-sectional relative odds of obesity increased while longitudinal relative odds slightly decreased and then stabilized. The cross-sectional relative odds of obesity for males decreased and then increased while the longitudinal relative odds steadily decreased.
Figure 7. Cross-sectional and longitudinal relative odds of obesity over time for Black and African American students.
Figure 8. Cross-sectional and longitudinal relative odds of obesity over time for Latino students.
Figure 9. Cross-sectional and longitudinal relative odds of obesity over time for Asian students
Figure 10. Cross-sectional and longitudinal relative odds of obesity over time for socioeconomic status.
5. **Discussion**

This study compared cross-sectional versus longitudinal analysis of data from one of the first school-based BMI surveillance systems in the nation. While the directionality of the cross-sectional and longitudinal relative odds of obesity over time was similar, cross-sectional analysis overestimated relative odds of obesity compared to longitudinal analysis. Childhood and adolescent obesity is an important public health priority. School-based interventions are abundant, as schools offer access to children, the facilities necessary for classroom or physical education and physical activity interventions, and qualified personnel to carry out these efforts.
(Kropski et al., 2008) Given the limited resources of school systems dedicated to data analysis, studying the implications of cross-sectional analysis of longitudinal data is important to not only recommending best practice approaches to analysis of school-based BMI surveillance but also to childhood and adolescent obesity prevention. Thus, a precise understanding of the information provided by school-based BMI surveillance systems is likewise critical to best understand the problem as well as any intervention effects. To our knowledge, this is the first study to compare cross-sectional and longitudinal analysis of school-based BMI surveillance.

Results from this study suggest that BMI surveillance data are more accurately analyzed over time using longitudinal analysis and that cross-sectional analysis may overestimate changes in obesity rates relative to longitudinal analysis that tracks the same individuals over time. Comparing the cross-sectional and longitudinal relative odds of obesity at each study time point showed that relative odds ratio values became more disparate with time. The cross-sectional and longitudinal relative odds of obesity for Black and African American, Latino, Asian, and low-income students were apparently and increasingly different at each time point. By 2008, the cross-sectional and longitudinal comparison of the relative odds of obesity for low-income students demonstrated a significant difference shown by non-overlapping CI’s. The increasing difference between cross-sectionally and longitudinally derived relative odds of obesity with time was true of each race/ethnicity variable. The overlapping CI’s for each of these comparisons, though, may be better explained by power differences than lack of significance. The cross-sectional relative odds ratios had wider confidence intervals than the longitudinal calculations. For example, the cross-sectional and longitudinal comparison of relative odds of obesity for Asian students in 2004 overlapped by a value of 0.163 but by 2008, only overlapped by a value of 0.023. In 2004, cross-sectionally, Asian students were 0.719 (0.522, 0.990) times
as likely as White students to be obese and longitudinally were only 0.585 (0.499, 0.685) times as likely. This effect increased over time. By 2008, relative odds of obesity had increased to 1.08 (0.773, 1.508) cross-sectionally but had only increased to 0.697 (0.610, 0.796), longitudinally. Cross-sectionally, compared to White students, Asian students’ relative odds of being obese shifted from Asian origin being protective to exceeding the relative odds of being obese for White students, while the relative odds of obesity increased but remained protective longitudinally.

While to our knowledge there are no other published studies comparing obesity rates derived cross-sectionally versus longitudinally, other studies comparing cross-sectional and longitudinal analyses have also found significant differences between analyses. In a study comparing perceived health changes among elderly participants, Orfila et al. (2000) found that cross-sectional analysis undervalued the level to which deterioration in perceived health was associated with aging by a factor of 8.7. In two studies examining the effect on the relationship between maternal age and the risk of delivering a child with spinal bifida and stillbirth ratio, respectively, each also found differing cross-sectional and longitudinal results. (Janerich, 1973; Resseguiem, 1976) In both of these cases, unlike this study, cross-sectional analysis substantially underestimated results. (Janerich, 1973; Resseguiem, 1976) Cross-sectional analysis is conceptually incongruous with assessment of risk that changes with both time and age, like BMI. (Resseguiem, 1976; Weiner & Long, 2004) In addition, cross-sectional analysis also masks dynamic effects. For example, a study examining performance assessments in the management of diabetes found that while cross-sectional results indicated a modest improvement in diabetes control, longitudinal results illustrated a more dynamic picture that was missed cross-sectionally. (Weiner & Long, 2004) It is important to note that BMI changes with both time and
age, and weight status determined by BMI-for-age percentiles can be dynamic, especially for those who fall on cut-off points. Losing a dynamic picture in cross-sectional analysis may be one explanation of the inflated obesity risk determined cross-sectionally.

The sociodemographic disparities of obesity have been well established and well researched. (Babey, 2010; Madsen et al., 2010) Though examining such disparities was not the objective of this study, the longitudinal analysis demonstrated an increased relative odds of obesity for Black and African American, Latino, and Asian students compared to White students despite a decrease in overall obesity prevalence. Obesity rates among White students are declining at a greater speed. This is a similar result to Madsen et al. (2010) who used cross-sectional methodology of longitudinal data to look at disparities in peaks, plateaus, and declines in prevalence of high BMI among adolescents and found that prevalence of high BMI is declining for some groups but has not declined for American Indian and Black girls, indicating greater disparities over time. Relative odds of obesity among low-income students versus higher income students likewise increased despite an overall decrease in prevalence. In another study looking at obesity rates among California adolescents, Babey et al. (2001) found that between 2001 and 2007, obesity prevalence significantly increased among lower-income adolescents but showed no statistically significant differences among higher-income adolescents after adjustment for age, gender, and race/ethnicity.

6. **Strengths and limitations**

In 2003, Arkansas was the first state to legislate BMI surveillance and screening, but CPS started their BMI surveillance and screening system in 2000. (Longjohn et al., 2010) The large size of this data set and the inclusion of all kindergarten to eighth graders are unique strengths. By using HLM in conjunction with maximum likelihood estimation, we were able to
overcome one common disadvantage of longitudinal analysis of missing follow-up data due to, for example, a student switching school systems or absenteeism during one of the time points. An advantage of HLM is that all students who have been observed at least once can be incorporated into the analysis. (Raudenbush, 2002) HLM was used to analyze the data longitudinally and therefore enabled inclusion of students with missing data. Since state legislation has helped to provide essential weight status surveillance information, it is important to use the information in the most valuable way. On the contrary, Madson et al. (2010) reported that a total of 2,477,450 youth were excluded from analysis because of missing or invalid data.

The differing cross-sectional and longitudinal results may be due to differing characteristics of the population who leaves the school-system or who does not get measured from year to year. This study was limited in that there was no way to assess the characteristics of the population who came into and out of the school system and differentiate between who was in the school system but not measured during any particular year.

This study emphasizes the importance of longitudinal statistical methodology for school-based BMI surveillance system analysis for the purpose of tracking trends in the obesity epidemic over time and for measuring program intervention effects over time; cross-sectional analysis may not be a reliable method for these purposes since relative odds of obesity over time may be overestimated, and any shift in risk over time especially with a dynamic population may be missed. One of the objectives of Arkansas’s legislation to mandate BMI surveillance was to link annual BMI results and show an individual student’s trend. (Justus, 2007) Based on the results of this study, showing an individual student’s growth trajectory over time would be most accurately analyzed using longitudinal statistical methodology. However, analyzing BMI surveillance data cross-sectionally, for the purpose of reporting annual obesity may be
appropriate, especially since, in general, the directionality of change in relative odds of obesity over time was the same despite the differing values of the relative odds ratios. For school districts lacking analytical resources and the capacity to conduct complex analyses such as HLM, cross-sectional analysis of a BMI surveillance system may provide a general idea of a school district’s annual obesity rates.

D. **Manuscript 3—Trajectory of Body Mass Index Percentile Growth among Sociodemographically Diverse Elementary School Students Exposed to Eight Years of System-Wide Ecological Nutrition and Physical Activity Intervention**

1. **Abstract**

   **Objectives.** Kindergarten to eighth grade CPS students’ BMI PCT growth trajectories were assessed over an eight-year period of an ecological school-based participatory physical activity and nutrition intervention. Sociodemographic growth trajectory differences were also assessed.

   **Methods.** Multilevel growth curve modeling for panel data was used to estimate students’ BMI PCT growth trajectories as a function of cumulative exposure to the intervention, while controlling for gender, race/ethnicity, and SES. Percentage distributions of demographic characteristics and obesity status were determined at the study entry point in 2000 and final time point in 2008.

   **Results.** Controlling for gender, race/ethnicity, and SES, the mean BMI PCT trajectory was estimated to start at 57.64 BMI PCT at the first study time point, to slope upward at a rate of 0.81 BMI PCT units for each 12-month unit of time, but to decrease this growth rate by -0.19 BMI PCT units with each additional year of exposure to the nutrition and physical activity intervention. Cumulative exposure to the intervention was significantly associated with
decreased BMI PCT growth across most sociodemographic groups although at varying rates.

**Conclusions.** The correlation of cumulative exposure to school-based intervention and decreased BMI PCT growth reinforces the importance of longitudinal obesity prevention methodology and analysis and supports the role of schools in efforts to address childhood and adolescent obesity.

2. **Introduction**

The dramatic increase in childhood and adolescent obesity in the United States has been well established. From 1970 to 2000, the prevalence of childhood and adolescent obesity more than tripled. (Hedley et al., 2004; Joliffe, 2004; Ogden et al., 2006) By the end of the first decade of this century, the rapid increase in obesity prevalence at the end of the twentieth century was seen to slow and possibly level. (Ogden et al., 2012) The prevalence of high BMI among children and adolescents remains high and is currently estimated at 18.7% at or above the 95th percentile (benchmark for obese youth). (Ogden et al., 2012)

Most notable are the persistent disparities. Minority children and adolescents and those from lower SES families are more likely to be obese. (Delva et al., 2007; Dwyer et al., 1998) Males are sometimes reported to have higher BMIs, but not always, and NHANES data from 1999 to 2008 have demonstrated a significantly higher prevalence of obesity among Latino and Black and African American youth, currently 23.2% and 22.3%, respectively, compared with non-Latino White youth, 17.0%. (Crespo et al., 2001; Hanson & Chen, 2007; Ogden et al., 2012) Despite the evidence of a possible stabilization of youth obesity, among males with the highest BMIs, a significant increase in obesity has been observed. (Madsen et al., 2010) Moreover, substantial racial/ethnic disparities persist with Latino boys, 26.7%, and non-Latino Black girls, 25.9%, disproportionately continuing to be affected by obesity. (Madsen et al., 2010; Ogden et
Thus, at the same time that there is evidence of the childhood and adolescent obesity epidemic leveling, trends in the data suggest that disparities in the prevalence of obesity are expected to continue to worsen over time, widening the race/ethnicity gap. (Madsen et al., 2010)

The rapid increase in the prevalence of obesity in children and adolescents has led to an increase in associated pediatric cardiovascular disease risk factors and psychosocial disorders. (Dietz & Gortmaker, 2001; Schonfeld-Warden & Warden, 1997) Furthermore, obese children are likely to become obese adults, especially if obese by age six. (Schonfeld-Warden & Warden, 1997) Obese adults are, likewise, at increased risk for type 2 diabetes, hypertension, stroke as a result of hypertension, and cardiovascular disease, as well as for certain cancers, orthopedic problems, and respiratory problems. (Dietz & Gortmaker, 2001; Must, 1999; Ogden et al., 2010; Schonfeld-Warden & Warden, 1997; Styne, 2001) Obesity in the United States has not only led to an estimated $147 billion in medical costs but also to the possibility that, for the first time in history, children and adolescents would have shorter life spans than their parents. (Finkelstein et al., 2009; Olshansky et al., 2005)

Interventions that promote healthy eating and physical activity during childhood and adolescence may, therefore, not only prevent some of the leading causes of death but may also decrease direct health care costs and improve quality of life. (Centers for Disease Control and Prevention, 1996) Recent reviews of the literature regarding childhood and adolescent obesity have called for more prospective studies with larger samples and more sophisticated and longer-term longitudinal analyses that control for potentially confounding factors such as SES, race/ethnicity and gender. (Danner, 2008) School-based approaches to obesity prevention have been advocated because of near-universal enrollment of children and adolescents in school, the amount of contact time with children and adolescents, and the potential to affect behaviors of
children that track into adolescence and adulthood through already existing organizational, social, and communication structures. (Dietz & Gortmaker, 2001; Schmitz & Jeffery, 2000) In addition, the school atmosphere provides ample opportunity to offer a range of activities and environments to improve diet and increase physical activity. (Dietz & Gortmaker, 2001) Schools may also be able to ensure coordination of comprehensive nutrition education programs, a healthy school environment, and policies to uphold a healthy environment. (Dietz & Gortmaker, 2001)

Beyond being a recommended setting for intervention, schools are an ideal setting to track childhood and adolescent overweight and obesity. As an important obesity prevention strategy, the IOM and the AAP recommend the routine tracking of BMI in all children and adolescents. (Stoddard et al., 2008) Obesity prevalence and trend data can be used to identify and track disparities, design and evaluate targeted interventions, and advocate for policy changes and funding. (Longjohn et al., 2010) Since 2003, approximately 30 states have either enacted or proposed school-based BMI surveillance laws and legislation. (Longjohn et al., 2010) Though the Massachusetts Public Health Council didn’t approve legislation requiring all public schools to measure BMI and send individual student BMI results to families until 2009, the CPS in Massachusetts have monitored all kindergarten to eighth grade students’ BMI status and sent results home to families since 2000.

The purpose of the present study was to longitudinally analyze data from the CPS BMI surveillance database to assess CPS students’ BMI PCT trajectories over an eight-year period of an ecological school-based intervention, as well as to assess sociodemographic differences in these trajectories.
3. **Methods**

a. **Data Source**

   The CPS physical education PE department developed a district-wide BMI surveillance and screening program in 2000, three years before any BMI surveillance and screening state legislation and nine years before legislation in Massachusetts. (Chomitz et al., 2003; Longjohn et al., 2010) Student BMI data for this study were obtained from the CPS surveillance and screening system via separate Microsoft Excel (Microsoft Corporation, Redmond, Washington) databases populated every year for nine years. Student identification numbers were used to link data since student names were not included in any database. The University of Illinois at Chicago and the Cambridge Health Alliance IRBs approved this study as an exempt protocol.

   As part of the elementary physical education curriculum of CPS, each spring since 2000, PE teachers measure heights and weights on all students enrolled in kindergarten through eighth grade. Prior to the data collection period each year, all CPS PE teachers attend a three-hour anthropometry, data entry, and sensitivity training session conducted by local public health professionals and PE administrators. Height and weight measurements are collected with school-issued wall-mounted stadiometers (*Seca 216 Accuhite, Snoqualmie, Washington*) and calibrating digital scales (*Seca 840 Bellissima-digital, Snoqualmie, Washington*) by one PE teacher per school often with assistance from a school nurse or another PE teacher. Height and weight are assessed with one measure each. Height is reported to the nearest quarter inch and weight to the nearest two-tenths of a pound. Teachers and research assistants enter data into a standardized, password-protected Microsoft Excel (Microsoft Corporation, Redmond, Washington) spreadsheet, which is pre-programmed to indicate when biologically implausible
height and weight values have been entered using the CDC SAS protocol for biologically implausible values. (Wisconsin Partnership for Activity and Nutrition, 2007) The Excel database is also pre-programmed to automatically calculate BMIs and BMI-for-age percentiles calculated using the SAS program (SAS Institute, Inc., Cary, North Carolina) developed by the CDC based on the 2000 gender-specific BMI-for-age growth charts. (Centers for Disease Control and Prevention, 2000) In addition to heights, weights, BMIs, and BMI-for-age percentiles, each Excel database also included, populated from the CPS student database, students’ gender, self-identified race/ethnicity, date of birth, and National School Lunch Program eligibility status. A reliability study conducted by Shapiro Berkson et al. (Chapter IV.B.) demonstrated that BMI measurements by CPS PE teachers used for surveillance purposes were reliable.

b. **Study population**

The CPS student population is ethnically, racially, linguistically, and socioeconomically diverse. The CPS Controlled Choice Plan includes in its school assignment policy SES, determined by free and reduced-price meals eligibility status, and race/ethnicity in order to have, in each school, the multi-faceted diversity that will provide all students with equitable educational opportunities and with improved achievement. (Cambridge Public Schools, 2001) At the start of the study in 2000, all CPS students in kindergarten to eighth grade, aged four to fifteen years, with reliable height and weight measurements were included, and students with reliable height and weight measurements from each subsequent incoming kindergarten class and students transferring into the school system through the final time point in 2008 were also included. Students aged out of the study population upon graduating eighth grade or transferring out of the study population by leaving the school system.
c. **Intervention**

Between 2000 and 2008, guided by the social-ecological model and community-based participatory research (CBPR) principles and methodology, CPS underwent an incrementally increasing physical activity and nutrition intervention. (Chomitz et al., 2010) During 2000 and 2001, the first two study time points, the dose of physical activity and nutrition intervention was minimal. By the 2004 time point, two separate pilot nutrition and physical activity grants supported a system-wide intervention, though some activities varied by school, focusing on PE enhancements, food service reforms, farm-to-school-to-home and school garden programs, and family outreach in addition to BMI and Fitness surveillance and reports. (Chomitz et al., 2010) By the final time point in 2008, three primary grants supported the full implementation of the PE enhancements, food service reforms, farm-to-school-to-home and school garden programs, and family outreach in addition to BMI and fitness surveillance and reports, which facilitated community awareness campaigns and ultimately the enactment of supportive school system and city policies. Chomitz et al. (2010) further detail this final implementation phase.

d. **Statistical analysis**

Data analyses focused on the trajectory of BMI PCT growth of children and adolescents beginning anywhere from kindergarten to eighth grade, depending on when they entered the study, and the correlation of their cumulative exposure to system-wide physical activity and nutrition intervention, while controlling for gender, race/ethnicity, age, SES using National School Lunch Program eligibility as a proxy, and time. The race/ethnicity variable was entered as a dummy code with White students serving as the reference group. The gender and SES variables were dichotomized, with SES divided into low income (receiving free and/or
reduced-price lunch) and not low income (no assistance). The time variable was defined as a student’s grade, from kindergarten to eighth grade with kindergarten being set at zero. Cumulative exposure was the number of years that the student was exposed to the intervention and was calculated for each student at each time point. The 2000 and 2001 time points were excluded in this calculation since intervention activities were minimal. Cumulative exposure was modeled as a continuous variable. The maximum cumulative exposure value was seven years. Student BMI PCT rather than BMI was used as the dependent variable because BMI PCT measures from multiple time points are directly comparable: an increase in the BMI PCT can be interpreted as excess weight while an increase in BMI may only reflect the natural growth curve of the child or adolescent. (Fernandes & Sturm, 2011) These analyses were addressed with growth curve modeling using HLM 7 software (Scientific Software International, Chicago, Illinois) and procedures described by Raudenbush and Bryk and Singer and Willett. (Danner, 2008; Raudenbush, 2002; Singer, 2003) Simple descriptive analyses were also completed of the percentage distributions for sociodemographic characteristics as well as for obesity status at the study entry point in 2000 and final time point in 2008.

4. Results

Linking BMI measurements from nine Excel databases yielded 11,692 student records and 37,318 BMI measurements. On average over the study period, 2.3% of students per year were not measured due to either refusal or absence on measurement days. Due to substantial school population attrition over the study period, there was a decreasing range of 4,847 students measured in 2000 to 3,739 students measured in 2008. Each student had anywhere from one to nine BMI measurements, with an average of 4.19 measurements. To calculate the time variable, 12,699 measurements were dropped due to a missing grade value or a
grade value outside of the kindergarten to eighth grade range. An additional eleven measurements were dropped due to a missing value for SES. There were 24,608 BMI measurements from 8,975 student records included in the final analysis. Table VI presents comparative data on the demographic characteristics of the full student BMI surveillance database population and the final study population. The distribution of demographic characteristics was within one percentage point for all characteristics other than White ethnicity, which was still within two percentage points.

<table>
<thead>
<tr>
<th>TABLE VI</th>
<th>PERCENTAGE DISTRIBUTIONS FOR DEMOGRAPHIC CHARACTERISTICS OF FULL BMI SURVEILLANCE DATABASE POPULATION AND STUDY POPULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Full population % (n)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>51.27 (5997)</td>
</tr>
<tr>
<td>Female</td>
<td>48.73 (5700)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
</tr>
<tr>
<td>Black and African American</td>
<td>32.50 (3802)</td>
</tr>
<tr>
<td>Latino</td>
<td>14.32 (1675)</td>
</tr>
<tr>
<td>White</td>
<td>36.37 (4554)</td>
</tr>
<tr>
<td>Asian</td>
<td>13.77 (1675)</td>
</tr>
<tr>
<td>National School Program Lunch Eligibility</td>
<td></td>
</tr>
<tr>
<td>Free/reduced</td>
<td>43.47</td>
</tr>
<tr>
<td>Paid</td>
<td>56.53</td>
</tr>
</tbody>
</table>
Table VII presents the results of obesity status at the study entry point in 2000 and at the final time point in 2008, as well as the results of obesity status among the population exposed to the full intervention. Within the full study population, the prevalence of obesity decreased for all sociodemographic groups other than Asian students. The prevalence of obesity for this group increased by 1.77%. Notably, the prevalence of obesity did not equally decrease across all groups. Obesity prevalence decreased by 5.63% among White students and 5.26% among higher income students compared to only 0.97% among Black and African American students, 1.0% among Latino students, and 0.27% among low-income students.

However, among the full study population, 2.18% (269) of White students, 1.83% (207) of Black and African American students, 2.03% (94) of Latino students, and 1.64% (60) of Asian students were exposed to the full intervention. The population exposed to the full intervention was racially/ethnically representative of the full study sample (32.86% Black and African American students, 14.92% Latino students, 9.53% Asian students, and 42.70% White students). Among this population exposed to the full intervention, obesity prevalence decreased for all groups other than Latino students who experienced a 4.74% increase. Obesity prevalence among White students decreased by 6.45%, among Black and African American students by 3.39%, among Asian students (counter to the overall rate) by 1.49%, and among low- and high-income students by 0.90% and 6.19%, respectively.
TABLE VII
WEIGHT STATUS OF STUDY POPULATION COMPARED TO NATIONAL PREVALENCE AT THE FIRST AND LAST STUDY TIME POINTS

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>2000 % (n)</td>
<td>2008 % (n)</td>
<td>2000 %</td>
<td>2008 %</td>
</tr>
<tr>
<td>Underweight</td>
<td>2.25 (83)</td>
<td>1.91 (269)</td>
<td>1.26 (8)</td>
<td>1.26 (8)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>60.57 (2234)</td>
<td>62.78 (8833)</td>
<td>64.94 (413)</td>
<td>64.94 (413)</td>
</tr>
<tr>
<td>Overweight</td>
<td>17.03 (628)</td>
<td>16.94 (2383)</td>
<td>16.67 (106)</td>
<td>16.67 (106)</td>
</tr>
<tr>
<td>Obese</td>
<td>20.15 (743)</td>
<td>18.37 (2584)</td>
<td>17.14 (109)</td>
<td>17.14 (109)</td>
</tr>
<tr>
<td>Gender: Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.25 (42)</td>
<td>2.02 (146)</td>
<td>0.96 (3)</td>
<td>0.96 (3)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>59.59 (1112)</td>
<td>62.08 (4485)</td>
<td>66.03 (206)</td>
<td>66.03 (206)</td>
</tr>
<tr>
<td>Overweight</td>
<td>16.08 (300)</td>
<td>16.76 (1211)</td>
<td>15.06 (47)</td>
<td>15.06 (47)</td>
</tr>
<tr>
<td>Obese</td>
<td>22.08 (412)</td>
<td>19.14 (1383)</td>
<td>17.95 (56)</td>
<td>17.95 (56)</td>
</tr>
<tr>
<td>Gender: Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.25 (41)</td>
<td>1.82 (126)</td>
<td>1.54 (5)</td>
<td>1.54 (5)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>61.55 (1122)</td>
<td>63.36 (4381)</td>
<td>63.89 (207)</td>
<td>63.89 (207)</td>
</tr>
<tr>
<td>Overweight</td>
<td>17.99 (328)</td>
<td>17.14 (1185)</td>
<td>18.21 (59)</td>
<td>18.21 (59)</td>
</tr>
<tr>
<td>Obese</td>
<td>18.21 (1823)</td>
<td>17.67 (1222)</td>
<td>16.36 (53)</td>
<td>16.36 (53)</td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.58 (40)</td>
<td>2.49 (129)</td>
<td>2.23 (6)</td>
<td>2.23 (6)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>63.48 (984)</td>
<td>71.57 (3715)</td>
<td>72.86 (196)</td>
<td>72.86 (196)</td>
</tr>
<tr>
<td>Overweight</td>
<td>16.71 (259)</td>
<td>14.35 (745)</td>
<td>14.13 (38)</td>
<td>14.13 (38)</td>
</tr>
<tr>
<td>Obese</td>
<td>17.23 (267)</td>
<td>11.60 (602)</td>
<td>10.78 (29)</td>
<td>10.78 (29)</td>
</tr>
<tr>
<td>Black and African American</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>1.53 (18)</td>
<td>1.62 (87)</td>
<td>0.97 (2)</td>
<td>0.97 (2)</td>
</tr>
<tr>
<td>Normal weight</td>
<td>54.77 (643)</td>
<td>55.44 (2985)</td>
<td>58.45 (121)</td>
<td>58.45 (121)</td>
</tr>
<tr>
<td>Overweight</td>
<td>18.57 (218)</td>
<td>18.78 (1011)</td>
<td>18.84 (39)</td>
<td>18.84 (39)</td>
</tr>
<tr>
<td>Obese</td>
<td>25.13 (295)</td>
<td>24.16 (1301)</td>
<td>21.74 (45)</td>
<td>21.74 (45)</td>
</tr>
<tr>
<td>Latino</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>1.42 (7)</td>
<td>0.76 (17)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal weight</td>
<td>55.67 (275)</td>
<td>54.03 (1207)</td>
<td>51.06 (48)</td>
<td>51.06 (48)</td>
</tr>
<tr>
<td>Overweight</td>
<td>16.80 (83)</td>
<td>20.10 (449)</td>
<td>18.09 (17)</td>
<td>18.09 (17)</td>
</tr>
<tr>
<td>Obese</td>
<td>26.11 (129)</td>
<td>25.11 (561)</td>
<td>30.85 (29)</td>
<td>30.85 (29)</td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>4.11 (18)</td>
<td>2.72 (48)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Normal weight</td>
<td>71.69 (314)</td>
<td>70.41 (1242)</td>
<td>73.33 (44)</td>
<td>73.33 (44)</td>
</tr>
<tr>
<td>Overweight</td>
<td>14.38 (63)</td>
<td>15.42 (272)</td>
<td>18.33 (11)</td>
<td>18.33 (11)</td>
</tr>
<tr>
<td>Obese</td>
<td>9.82 (43)</td>
<td>11.45 (202)</td>
<td>8.33 (5)</td>
<td>8.33 (5)</td>
</tr>
</tbody>
</table>
TABLE VII (continued)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>National School Program Lunch Eligibility</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Free/reduced</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.11 (40)</td>
<td>1.83 (125)</td>
<td>1.11 (3)</td>
<td></td>
</tr>
<tr>
<td>Normal weight</td>
<td>54.97 (1040)</td>
<td>54.75 (3732)</td>
<td>51.66 (140)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>18.02 (341)</td>
<td>18.79 (1281)</td>
<td>23.25 (63)</td>
<td></td>
</tr>
<tr>
<td>Obese</td>
<td>24.89 (371)</td>
<td>24.62 (1678)</td>
<td>23.99 (65)</td>
<td>**</td>
</tr>
<tr>
<td>Paid</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Underweight</td>
<td>2.37 (76)</td>
<td>2.06 (176)</td>
<td>1.37 (5)</td>
<td></td>
</tr>
<tr>
<td>Normal weight</td>
<td>62.39 (2004)</td>
<td>69.43 (5942)</td>
<td>74.73 (272)</td>
<td></td>
</tr>
<tr>
<td>Overweight</td>
<td>16.77 (545)</td>
<td>15.49 (1326)</td>
<td>11.81 (43)</td>
<td></td>
</tr>
<tr>
<td>Obese</td>
<td>18.28 (587)</td>
<td>13.02 (1114)</td>
<td>12.09 (44)</td>
<td>**</td>
</tr>
</tbody>
</table>

*Prevalence only includes Mexican-American youth; **Prevalence not available
To analyze BMI trajectories, first, an unconditional model, model I, was run to determine if there was significant variance in the intercept (BMI PCT at study entry point) and slope (rate of increase or decrease of BMI PCT over time). Below are the equations for this unconditional model:

**Level-1 Model**

\[ \text{BMIPCT}_{it} = \pi_{0i} + \pi_{1i}(\text{TIME}_{it}) + e_{it} \]

**Level-2 Model**

\[ \pi_{0i} = \beta_{00} + r_{0i} \]
\[ \pi_{1i} = \beta_{10} + r_{1i} \]

Results from Model I indicated that there was significant (p < 0.001) variance in the intercept and slope, so all three of these parameters were retained for further analyses. A random slope on time improved the fit of the growth curve models. The mean BMI PCT trajectory was estimated to start at 64.20 BMI PCT at the first study time point and to slope upward at a rate of 0.54 BMI PCT units for each 12-month unit of time. Table VIII presents the fixed and random effect estimates from this unconditional model.
Cumulative exposure to the intervention was added to Model II as a time-varying covariate and in an interaction term with time (BMI PCT slope). This cumulative exposure by time interaction term was added to determine if cumulative exposure to nutrition and physical activity intervention significantly interacted with BMI PCT slope. The results are presented in Table IX.
The significant negative cumulative exposure by time ($p < .001$) indicates that cumulative exposure to intervention was significantly associated with decreased BMI growth. In this model, the mean BMI PCT trajectory was estimated to start at 62.62 BMI PCT at the first study time point, to slope upward at a rate of 0.88 BMI PCT units for each 12-month unit of time, but to
decrease this growth rate by -0.20 BMI PCT units with each additional year of exposure to the nutrition and physical activity intervention.

Model III (results not shown) added gender, race/ethnicity, age, and SES in order to determine if the associations between cumulative exposure and BMI PCT growth were still present after accounting for each of these variables. A student’s age was not significant, but the cumulative exposure by time interaction remained significant (p < 0.001). In model IV, the final model, the age predictor was removed. Table X presents the fixed and random effect estimates.
TABLE X
MODEL IV. GROWTH MODEL ESTIMATING BMI PCT GROWTH
TRAJECTORIES OF CHILDREN AND ADOLESCENTS FROM KINDERGARTEN TO
GRADE 8 AFTER ADDING CUMULATIVE EXPOSURE TO INTERVENTION AND THE
INTERACTION OF CUMULATIVE INTERVENTION EXPOSURE WITH TIME AND ALL
SIGNIFICANT DEMOGRAPHIC PREDICTORS FOR EACH PARAMETER

<table>
<thead>
<tr>
<th>Fixed effects</th>
<th>Coefficient</th>
<th>SE</th>
<th>t-ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI upon entry to study (Mean initial status) $\beta_{00}$</td>
<td>57.64</td>
<td>0.66</td>
<td>87.23</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Gender $\beta_{01}$</td>
<td>2.01</td>
<td>0.56</td>
<td>3.59</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Black/African American $\beta_{02}$</td>
<td>8.27</td>
<td>0.67</td>
<td>12.34</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Asian $\beta_{03}$</td>
<td>-2.49</td>
<td>0.93</td>
<td>-2.67</td>
<td>0.008</td>
</tr>
<tr>
<td>Latino $\beta_{04}$</td>
<td>8.58</td>
<td>0.85</td>
<td>10.07</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SES $\beta_{10}$</td>
<td>1.29</td>
<td>0.31</td>
<td>4.11</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Cumulative exposure to intervention $\beta_{20}$</td>
<td>0.94</td>
<td>0.21</td>
<td>4.43</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time (Slope) (Mean growth rate) $\beta_{30}$</td>
<td>0.81</td>
<td>0.11</td>
<td>7.62</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time x Cumulative exposure to intervention $\beta_{40}$</td>
<td>-0.19</td>
<td>0.03</td>
<td>-5.86</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Random Effect</th>
<th>Standard Deviation</th>
<th>Variance Component</th>
<th>$\chi^2$</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI upon entry to study Initial status $r_{0i}$</td>
<td>26.78</td>
<td>717.26</td>
<td>23034.47</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Time (Slope) Growth rate $r_{3i}$</td>
<td>2.74</td>
<td>7.53</td>
<td>9248.18</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Level 1 error $e_{ni}$</td>
<td>11.24</td>
<td>126.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Controlling for sociodemographic variables, the mean BMI PCT trajectory was estimated to start at 57.64 BMI PCT at the first study time point, to slope upward at a rate of 0.81 BMI PCT units for each 12-month unit of time but to decrease this growth rate by -0.19 BMI PCT units with each additional year of exposure to the nutrition and physical activity intervention. As independent variables in model IV, Black and African American and Latino students compared to White students had a large effect on BMI PCT while Asian students’ effect compared to White students was protective. In addition, boys compared to girls and low-income students compared to higher income students each had a significant effect on BMI PCT.

Compared to White students, Black and African American and Latino students’ mean BMI PCT trajectories were estimated to start at higher BMI PCTs at the first study time point and to slope upward at a greater rate, but their BMI PCT growth rate was likewise decreased with each additional year of exposure to the nutrition and physical activity intervention. For Black and African American students, the mean BMI PCT trajectory was estimated to start at 67.78 BMI PCT at the first study time point, to slope upward at a rate of 0.91 BMI PCT units for each 12-month unit of time, and to decrease this growth rate by -0.23 BMI PCT units with each additional year of exposure to the intervention. For Latino students, the mean BMI PCT trajectory was estimated to start at 65.77 BMI PCT at the first study time point, to slope upward at a rate of 0.68 BMI PCT units for each 12-month unit of time, and to decrease this growth rate by -0.35 BMI PCT units with each additional year of exposure to the intervention. For Asian students, the mean BMI PCT trajectory was estimated to start at 50.25 at the first study time point, to slope upward at a rate of 0.89 BMI PCT units for each 12-month unit of time, and to decrease this growth rate by -0.26 BMI PCT units with each additional year of exposure to the intervention.
5. **Discussion**

To the best of the authors’ knowledge, the present study is the first to longitudinally analyze BMI PCT trajectories over eight years among nearly a full kindergarten to eighth grade school system population within an environment of longitudinal participatory ecological physical activity and nutrition intervention. The correlation of cumulative exposure to school-based intervention and decreased BMI PCT growth seen in the present study is important to consider despite the lack of a control group and randomization.

Through the first decade and now into the second decade of this century, recommendations for school-based obesity intervention are often touted as the great hope for reversal of the childhood and adolescent obesity epidemic. (Centers for Disease Control and Prevention, 1996) However, since 1990 and the beginning of school-based intervention evaluation, no conclusive evidence regarding the benefit of school-based intervention has been established. (Kropski et al., 2008) In a systematic review by Kropski et al. of all research published on this issue since 1990, the authors only identified fourteen studies that were of sufficient quality to draw any conclusions. (Kropski et al., 2008) The Guide to Community Preventive Services Task Force likewise concluded that there was insufficient evidence to determine the effectiveness of school-based programs to control obesity. (Guide to Community Preventive Services, 2011) Despite the rather poor body of evidence, Kropski et al. (2008), like many investigators, continue to believe that schools will play an important role in stemming current trends in overweight and obesity in children and adolescents given the near-universal enrollment, the amount of time children and adolescents spend in school, and the opportunity to influence nutrition and physical activity. (Centers for Disease Control and Prevention, 1996;
Dietz & Gortmaker, 2001; Schmitz & Jeffery, 2000; Story, 1999) The results of this study clearly support the role that schools can take in the effort to reduce youth obesity.

These data also reconﬁrmed the socioeconomic disparity of childhood and adolescent obesity. Black and African American and Latino youth as well as low-income youth are disproportionately affected by obesity, clearly indicated in this study. Looking at the effect of the ecological intervention across the full study population showed a universally positive correlation with BMI PCT trajectories, as well as a positive correlation with obesity prevalence across most groups but at unequal rates demonstrated by the socio-economic differential in BMI PCT growth.

The variable rates of growth across sociodemographic groups can most likely best be explained by a variety of factors outside of the school environment for which we unfortunately have no data. Using an ecological model within a CBPR framework, the CPS intervention focused on improving nutritional habits and increasing physical activity while decreasing sedentary behavior. Influences outside of school were not part of this study, but results from the Youth Risk Behavior Surveillance System (YRBS) showed that Black and African American and Latino students are less likely to participate in physical activity and more likely to watch three or more hours of television per day compared to White students. (Centers for Disease Control and Prevention, 2009) Black and African American and Latino youth are also more likely to live in nutritionally compromised environments. (Eisenhauer, 2001)

Previous studies have shown that access to school-based physical activity programs is limited among minority youth and that Black and African American youth receive less exercise in school than Caucasian youth, even at similar SES levels. (Lee, 2005; Lindquist et al., 1999) However, CPS is based on a lottery system that includes sociodemographic diversity. All youth
have the opportunity to be exposed to the same interventions, and furthermore, some of the CPS interventions involved changes at the policy level. The sociodemographic diversity of CPS within one school combined with the extensive long-term participatory intervention was a unique public health opportunity for which the full exposure to intervention effect may be explained.

Results of systematic reviews and meta-analyses of school-based obesity interventions reveal that long-term multiple-level interventions are the most conclusively effective. (Gonzalez-Suarez et al., 2009; Kropski et al., 2008; Story, 1999) Studying only the students who were exposed to the full intervention in the present study not only showed the most dramatic results, but also showed positive results across populations experiencing the greatest disparities. The decrease in obesity prevalence effect was more pronounced among the students exposed to the full intervention, especially among Black and African American and Asian students.

One clear example of the value of the long-term effect of the intervention is seen in the dichotomy among Asian students. Among the full study population of Asian students, results showed that obesity prevalence increased. However, among Asian students exposed to the full intervention, obesity decreased. The dichotomy seen in Asian students was also seen in Latino students but unfortunately in the opposite direction. Among the full study population of Latino students, obesity prevalence decreased; among students exposed to the full intervention, obesity prevalence increased by more than 5.5%. This dichotomy seen among the Latino population may be explained by acculturation, as the 1990’s marked an influx to Cambridge of an ethnically diverse Latino population. (Cambridge Public Schools, 2001) Future research is needed to better explain this paradox and a possible relationship of obesity and acculturation among Cambridge’s ethnically diverse community.
This study did not have a control group, but national obesity prevalence data from the NHANES survey suggest that between 2000 and 2008 among six to nineteen-year-olds, obesity increased overall by 2.7% (Table VI). (Hedley et al., 2004; Ogden et al., 2010) In contrast, in CPS, obesity decreased overall by 1.8% among the full study population and by 3.0% among the full intervention population. In 2000, CPS students had a higher prevalence of obesity than the national prevalence, but by 2008, not only was the gap narrowed, but obesity prevalence among the full intervention population of White students was 6.22% less than the national prevalence of White youth. (Hedley et al., 2004; Ogden et al., 2010)

The results of this study also strongly support the value of school-based BMI surveillance. Evaluating school-based intervention is often a reason cited for school-based surveillance. (Longjohn et al., 2010) CPS began a comprehensive surveillance system before any state legislation. This study using nine time points from this BMI surveillance system provides evidence of this intended use and value and reinforces the importance of long-term intervention and analysis and the longitudinal intervention possibilities within a school environment.

6. **Strengths and limitations**

The strengths of the present study lie within the inclusiveness of the study population and its longitudinal nature. A large majority of all students in CPS kindergarten to eighth grade were included in the study. The ability to include nearly all students is a unique strength, especially given state BMI surveillance and screening mandates. In Massachusetts, for example, legislation only requires BMI measurement of all students in first, fourth, seventh, and tenth grades. Along the same lines, the diversity of CPS and the universality of the intervention enabled a diverse study population that was also equally exposed to many components of the intervention. Using
multi-level modeling to analyze the data was likewise a strength since all students who have been observed at least once can be incorporated into the analysis. (Raudenbush, 2002) We were therefore able to retain students with missing data.

Among various limitations inherent in an eight-year study with nine time points, three are particularly important. The primary limitation of this study as well as seen throughout school-based obesity intervention research was the inability to maintain a control group and randomization, and as a result, the inability to determine causality. As mentioned above, many of the systematic reviews of school-based obesity intervention have determined that school-based intervention is difficult to evaluate due to lack of scientific rigor and standardization. However, the nature of school-based work and particularly the nature of CBPR and evaluation make traditional scientific methodology difficult to maintain. In agreement with Buchanan et al. it is also the belief of the authors of this study that the necessary evaluation standards be inclusive of alternative methodology and outcomes so that important results from school-based intervention are not discounted. (Buchanan et al., 2007) Buchanan et al. (2007) proposed alternative methodology such as “weight-of-evidence” standards, like Hill’s criteria, based on items such as biological plausibility, consistency in association, dose–response relationships, and triangulation of multiple methods to strengthen confidence in the research results.

Another limitation of this study was that there were no data regarding participation in intervention activities and therefore no data regarding specific intervention component contribution or dose. Also, we were not able to account for any school effects since they could not be incorporated into the analysis. However, all students were exposed to intervention components targeted at the school-system-wide and citywide policy level.
Finally, the calculation of the cumulative exposure variable excluded approximately 30% of the full database. However, Table VI demonstrated that data appeared to be missing at random thereby most likely not affecting overall results.

7. **Conclusions**

Few previous school-based obesity intervention studies have been longitudinal and still fewer studies have incorporated both longitudinal intervention and analysis. To our knowledge, the present study is one of the first studies to assess BMI PCT trajectories over eight years and nine time points among nearly an entire elementary school system within the context of an eight-year physical activity and nutrition intervention. Throughout the years in which obesity and interventions alike have increased, many systematic reviews and meta-analyses of school-based obesity intervention have been conducted. Evaluation of individual school-based interventions and thus results of reviews and analyses have shown widely varying results leading to conclusions of insufficient evidence to recommend school-based obesity intervention. (Baranowski et al., 2002; Gonzalez-Suarez et al., 2009; Kropski et al., 2008) The inconsistency of school-based obesity intervention outcome evaluation is in part attributed to the lack of a commonly accepted convention for reporting results, including a standardized set of outcome measures. (Gonzalez-Suarez et al., 2009; Kropski et al., 2008; McGraw, 2000) However, despite the lack of a control group and randomization, the findings of this study showing that after controlling for gender, race/ethnicity, and SES, the cumulative effect of longitudinal intervention was significantly associated with decreased BMI PCT growth among a large sample of sociodemographically diverse CPS students are important to consider because the longitudinal intervention and analytic methodology begin to fill in gaping holes.
Further investigation must continue to explore the unequal rates of the decrease in growth across sociodemographic groups, including sociodemographic variation in response to intervention. In so doing, given the nature of school-based research and evaluation, results should not be discounted simply because of the difficulty to conduct randomized controlled trials within a school environment. Instead, a standardized evaluation protocol that focuses on the assets of the school environment must be developed.

Finally, to implement obesity intervention at the level of CPS involves many resources and much time that may be difficult for many school districts to obtain. Ultimately, the positive results of this study may suggest the need for supportive legislative policy in order to appropriately execute this broad-based intervention. A broad population-based approach guided by legislation for childhood and adolescent obesity prevention was similarly recommended by the IOM in a 2010 report. (Committee on an Evidence Framework for Obesity Prevention Decision Making, Institute of Medicine, 2010) Legislative steps towards a broad-based obesity prevention approach have been established within Michelle Obama’s Let’s Move campaign. (White House Task Force on Childhood Obesity Report to the President, 2010) In light of the Patient Protection and Affordable Care Act and steps toward greater universal health care coverage, health care cost containment is vitally essential. Real cost savings and thus greater ability of near universal coverage is more likely to be achieved through reducing obesity and related risk factors. The importance of alternative methodology and outcomes standards to be inclusive of non-traditional scientific environments, therefore, is imperative. (Finkelstein et al., 2009)
V. OVERALL CONCLUSION

Preventing obesity in our children and adolescents, as well as evening out obesity-related racial/ethnic disparities, is one of the most important public health issues facing the nation today. (American Public Health Association, 2012; U.S. Department of Health and Human Services, 2010) In 2003, the AAP and later the IOM and the Expert Committee on Childhood Obesity recommended the routine tracking of BMI in all children and adolescents. School-based BMI surveillance has therefore attracted much attention across the nation from researchers, school officials, legislators, and the media as an approach to help monitor and address obesity among youth. (Kubik et al., 2007; Nihiser et al., 2007) Previous research has demonstrated that, particularly in its use for surveillance, BMI is not only a valid measure of body fat, but it is also an inexpensive and efficient way to track obesity in a population over time. (Nihiser et al., 2007) Many important uses of school-based surveillance system data have been cited throughout the literature such as the ability to describe and track trends in the obesity epidemic over time among youth and to evaluate the effects of school-based physical activity and nutrition programs and policies. (Longjohn et al., 2010) Though the potential of positively impacting obesity prevention via schools is great, literature reviews, syntheses and meta analyses have concluded that more evidence is needed to determine the effectiveness of school-based programs to control overweight and obesity. (Guide to Community Preventive Services, 2011)

Gathering data from a school-based surveillance system has the potential to fill a glaring gap in childhood and adolescent obesity prevention research: longitudinal evaluation. The majority of research studies, especially school-based obesity prevention and intervention, have either cross-sectionally derived their outcome measures of obesity such as BMI or have only longitudinally analyzed less than five years of data. (Gonzalez-Suarez et al., 2009; Kambalia et
Cross-sectional analysis cannot account for directionality such as obese students who are less obese or for the transient nature of schools where new students who were not exposed to school-based interventions and students whose weight improved but left the system could mask program effects and make the data appear relatively stable. (Weiner & Long, 2004) Also, changes in the prevalence of obesity can take many years to detect. (Yanovski & Yanovski, 2011) Therefore, longitudinal analysis of program effects might not yield changes over a short time period. The inconclusiveness of school-based obesity prevention and intervention evaluation may be as related to methodology as it is to ineffective interventions. (Baranowski et al., 2002) The purpose of this project, therefore, was to address some of these methodological deficiencies in the above-noted research and thereby provide more reliable and robust evidence to inform the use of school-based BMI surveillance. This was accomplished in three phases by (1) first, assessing the reliability and providing research-based guidance on improving BMI surveillance and (2) then, by evaluating the differences between relative odds of obesity determined cross-sectionally and longitudinally and thereby demonstrating the more nuanced and accurate results of a longitudinal analysis of substantial length (3) and, finally, by applying these findings to longitudinally analyzing the data to assess students’ BMI PCT trajectories and the impact of ecological school-based interventions over an eight-year period as well as to assess sociodemographic differences in these trajectories and impacts.

For surveillance purposes, reliability of PE teachers’ measurements was “excellent.” This is an important finding that supports the feasibility of school-based BMI surveillance. Resources in schools can often be limited. The excellent reliability of the CPS surveillance system data provides evidence that non-clinical staff can reliably participate in collecting student
height and weight measurements. However, results also showed that staff training and data cleaning are imperative to data reliability.

Results from this project also demonstrated that analyzing school-based obesity surveillance data cross-sectionally to report annual obesity may be reasonable, but to track trends in the obesity epidemic over time or to measure program intervention effects over time, cross-sectional analysis may not be a reliable method since it may overestimate or underestimate the problem. Longitudinal analysis naturally provides the most accurate results of longitudinal obesity surveillance. This finding from the second phase of this project reemphasizes the need to evaluate school-based obesity prevention and intervention initiatives longitudinally where the longitudinal follow-up extends beyond five years. Many school districts may not have the analytical capacity to analyze data longitudinally. However, if student anonymity can be preserved in school-based surveillance system databases by using unique identifiers such as student identification numbers, these databases can be shared with institutions with longitudinal analytical capacity.

Results from the third phase of this project have important implications for childhood and adolescent obesity prevention. Using nine time points of data from a school-based surveillance database to track BMI PCT trajectories of students within a school system having undergone eight years of ecological school-based participatory physical activity and nutrition intervention did indeed find decreased BMI PCT growth that was associated with cumulative exposure to intervention. The findings from this study were consistent with the data that have shown racial/ethnic and socioeconomic disparities in obesity prevalence, as non-minority students’ BMI PCT growth decreased at a greater rate than minority students’ BMI PCT growth. However, looking only at students who were exposed to the full eight years of intervention revealed a
steeper decrease in BMI PCT growth for non-minority and minority students alike as well as for non-low income and low-income students, though still at disparate rates.

This study supports school-based BMI surveillance legislation by demonstrating an intended use of longitudinal surveillance and by showing how it begins to address the gap in longitudinal obesity prevalence data. This study also provides critical evidence of and reinforces the importance of longitudinal intervention and analysis. The school-system studied was able to dedicate an incredible amount of resources to providing obesity prevention and analyzing surveillance data. Many school systems and communities do not have the resources to provide, analyze, and sustain a citywide initiative for eight years, and not all communities are as amenable as Cambridge to such change. Ultimately, the results of this study indicate and support the need for policy level intervention since policy may be the only feasible way to sustain school-based obesity prevention strategies long-term while at the same time inclusively addressing the many other influences in a child’s life outside of school. Thus, future research should focus on (1) proposing best practices obesity prevention research as policy opportunities; (2) evaluating already implemented policy such as soda tax legislation, banning trans fats legislation, and the Healthy, Hunger-Free Kids Act; and (3) creating an evaluation rubric that goes beyond scientific methodology and could therefore most effectively evaluate initiatives such as school-based initiatives and policy that do not lend themselves to scientific methodology.
CITED LITERATURE


103


APPENDIX A

Data Cleaning and Analysis SAS Code

********************************************;
* Choose one of the following library links *;
********************************************;
libname longbmi
'\campubhe\public\ICH\Epidemiology\Sandra\PEP\datasets for stephanie';
libname longbmi 'Z:\Documents\Ph.D. Research\LongBMIdata';
libname longbmi 'C:\Documents and Settings\Eric Berkson\Desktop\LongBMIdata';
libname longbmi 'C:sasstephfiles6210';
libname longbmi 'C:sasstephfiles72210';
dm log 'clear';
dm output 'clear';
options ls=100 ps=80 nocenter nodate nonumber;
title1;
********************************************;
* Proc contents each year ( 8 sets 00 - 08 ) *;
********************************************;
proc contents data=longbmi.set00 varnum;
title 'Set00';
run;
*data longbmi.set00; *set longbmi.set00;
*proc freq;
*tables bmi;
*title '--- BMI Numbers longbmi.set00 -------';
*run;
proc contents data=longbmi.set01 varnum;
title 'Set01';
run;
proc contents data=longbmi.set02 varnum;
title 'Set02';
run;
proc contents data=longbmi.set03 varnum;
title 'Set03';
run;
proc contents data=longbmi.set04 varnum;
title 'Set04';
run;
proc contents data=longbmi.set05 varnum;
title 'Set05';
run;
proc contents data=longbmi.set06 varnum;
title 'Set06';
run;
proc contents data=longbmi.set07 varnum;
title 'Set07';
proc contents data=longbmi.set08 varnum;
title 'Set08';
proc contents data=longbmi.gradeothree varnum;
title 'Set03 with Grade';
APPENDIX A (continued)

run;
proc contents data=longbmi.gradeofour varnum;
title 'Set04 with Grade';
run;
******************************************************************************;
* Assess missing values
1=male, 2=female;
******************************************************************************;
data sexmiss; set longbmi.set08;
proc freq;
tables grade ;title 'Grade 08----------';
run;
data sexmiss; set longbmi.set07;
proc freq;
tables grade ;title 'Grade 07----------';
run;
data sexmiss; set longbmi.set02;
proc freq;
tables grade ;title 'Grade 02----------';
run;
data sexmiss; set longbmi.set01;
proc freq;
tables grade ;title 'Grade 01----------';
run;
data sexmiss; set longbmi.set00;
proc freq;
tables grade ;title 'Grade 00----------';
run;

data sexmiss; set longbmi.gradeothree;
proc freq;
tables grade;title 'Grade 03 new----------';
run;
data sexmiss; set longbmi.gradeofour;
proc freq;
tables grade ;title 'Grade 04 new----------';
run;
data sexmiss; set longbmi.gradeofive;
proc freq;
tables grade ;title 'Grade 05 new----------';
run;
******************************************************************************;
*Assess implausible BMIz values 2000;
******************************************************************************;
data ipv00; set longbmi.set00;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz; title '--implausible bmi anal----------';
run;
data ipv00b; set longbmi.set00;
if bmiz gt 5;
APPENDIX A (continued)

```plaintext
proc sort;
by bmiz;
proc print;
by bmiz;
*******************************;
*Assess implausible BMIZ values 2001;
*******************************;
*******************************;
data ipv01; set longbmi.set01;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv01b; set longbmi.set01;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
*******************************;
*Assess implausible BMIZ values 2002;
*******************************;
*******************************;
data ipv02; set longbmi.set02;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv02b; set longbmi.set02;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
*******************************;
*Assess implausible BMIZ values 2003;
*******************************;
data ipv03; set longbmi.set03;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv03b; set longbmi.set03;
if bmiz gt 5;
proc sort;
by bmiz;
```
proc print;
by bmiz;
run;
**************************************;
*Assess implausible BMIz values 2004;
**************************************;
data ipv04; set longbmi.set04;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv04b; set longbmi.set04;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
**************************************;
*Assess implausible BMIz values 2005;
**************************************;
data ipv00; set longbmi.set05;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv05b; set longbmi.set05;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
**************************************;
*Assess implausible BMIz values 2006;
**************************************;
data ipv06; set longbmi.set06;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv06b; set longbmi.set06;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
*******************************;
*Assess implausible BMIz values 2007;
*******************************;
*************;
data ipv07; set longbmi.set07;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data ipv07b; set longbmi.set07;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
*******************************;
*Assess implausible BMIz values 2008;
*******************************;
data ipv08; set longbmi.set08;
if bmiz lt -4;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
data bipv; set longbmi.set08;
if bmiz gt 5;
proc sort;
by bmiz;
proc print;
by bmiz;
run;
*******************************;
* Start FINAL sets of cleaned data*;
*******************************;
*************;
* Each Variable*;
*************;
* SID      ;
* New Final Variable name :  sidf   *
*Making all IDs numeric format
*************;
data final00; set longbmi.set00;
  if trim(left(sid))="." then sid_num=.;
  else do;
    sid_num=sid+0;
    sidf=sid_num;

APPENDIX A (continued)

drop sid_num;
   end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title '--SET 00 MISSING---------';
run;
data final01; set longbmi.set01;
   if trim(left(sid))="." then sid_num=.;
   else do;
      sid_num=sid+0;
      sidf=sid_num;
      drop sid_num;
   end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title '--SET 01 MISSING---------';
run;
data final02; set longbmi.set02;
   if trim(left(sid))="." then sid_num=.;
   else do;
      sid_num=sid+0;
      sidf=sid_num;
      drop sid_num;
   end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title '--SET 02 MISSING---------';
run;
data final03; set longbmi.set03;
   if trim(left(sid))="." then sid_num=.;
   else do;
      sid_num=sid+0;
      sidf=sid_num;
      drop sid_num;
   end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title '--SET 03 MISSING---------';
run;
data final04; set longbmi.set04;
   if trim(left(sid))="." then sid_num=.;
   else do;
      sid_num=sid+0;
      sidf=sid_num;
      drop sid_num;
   end;
APPENDIX A (continued)

if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title 'SET 04 MISSING---------
run;
data final05; set longbmi.set05;
if trim(left(sid))="." then sid_num=.;
   else do;
sid_num=sid+0;
sidf=sid_num;
drop sid_num;
end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title 'SET 05 MISSING---------
run;
data final06; set longbmi.set06;
if trim(left(sid))="." then sid_num=.;
   else do;
sid_num=sid+0;
sidf=sid_num;
drop sid_num;
end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title 'SET 06 MISSING---------
run;
data final07; set longbmi.set07;
ethnic07 = ethnic_code+0;
run;
proc freq;
tables ethnic07;
run;
data final07; set final07;
if trim(left(sid))="." then sid_num=.;
   else do;
sid_num=sid+0;
sidf=sid_num;
drop sid_num;
end;
   if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title 'SET 07 MISSING---------
run;
data final08; set longbmi.set08;
ethnic08 = ethnic_code+0;
run;
APPENDIX A (continued)

proc freq;
tables ethnic08;
run;
data final08; set final08;
  if trim(left(sid))="." then sid_num=.;
  else do;
    sid_num=sid+0;
    sidf=sid_num;
    drop sid_num;
  end;
  if sidf=. then sidmissing=1; else sidmissing=0;
run;
proc freq;
tables sidmissing;
title 'SET 08 MISSING----------';
run;

*********************************************************************
* Race/ethnicity;
**Re-code--Changing ethnic to NEW DOE code*;
*********************************************************************
data final00; set final00;
  if race=6 then ethnic00=1;
  if race=5 then ethnic00=33;
  if race=1 then ethnic00=4;
  if race=2 then ethnic00=2;
  if race=3 then ethnic00=3;
run;
proc freq;
tables ethnic00;
run;
data final01; set final01;
  if ethnic=6 then ethnic01=1;
  if ethnic=5 then ethnic01=33;
  if ethnic=1 then ethnic01=4;
  if ethnic=2 then ethnic01=2;
  if ethnic=3 then ethnic01=3;
run;
proc freq;
tables ethnic01;
run;
data final02; set final02;
  if ethnic=6 then ethnic02=1;
  if ethnic=5 then ethnic02=33;
  if ethnic=1 then ethnic02=4;
  if ethnic=2 then ethnic02=2;
  if ethnic=3 then ethnic02=3;
run;
proc freq;
tables ethnic02;
run;
data final03; set final03;
  if ethnic=6 then ethnic03=1;
  if ethnic=5 then ethnic03=33;
APPENDIX A (continued)

if ethnic=1 then ethnic03=4;
if ethnic=2 then ethnic03=2;
if ethnic=3 then ethnic03=3;
run;
proc freq;
tables ethnic ethnic03;
run;
data final04; set final04;
if ethnic='White' then ethnic04=1;
if ethnic='Hispanic' then ethnic04=33;
if ethnic='Nat Amer' then ethnic04=4;
if ethnic='Black' then ethnic04=2;
if ethnic='Asian' then ethnic04=3;
run;
proc freq;
tables ethnic04;
run;
data final05; set final05;
if ethnic_code=6 then ethnic05=1;
if ethnic_code=5 then ethnic05=33;
if ethnic_code=1 then ethnic05=4;
if ethnic_code=2 then ethnic05=2;
if ethnic_code=3 then ethnic05=3;
run;
proc freq;
tables ethnic05;
run;
*06 & 07 were entered as the new DOE code already*;
data final06; set final06;
ethnic06 = ethniccode+0;
run;
proc freq;
tables ethnic06;
run;
********************************************;
** BDate    ;
** New Final Variable name :  bdateYR ;
********************************************;
data final00; set final00;
if trim(left(Bdate))="." then bdate00=.;
else do;
bdate00=datepart(bdate);
format bdate00 date.;
end;
run;
data final01; set final01;
bdate01=datepart(bdate);
format bdate01 date.;
run;
data final02; set final02;
bdate02=datepart(bdate);
format bdate02 date.;
run;
APPENDIX A (continued)

data final03; set final03;
bdate03=datepart(bdate);
format bdate03 date.;
run;
data final04; set final04;
bdate04=bdate1-21916; * Conversion from an Excel date by subtract 21916*;
format bdate04 date.;
run;
data final05; set final05;
bdate05=bdate;
format bdate05 date.;
run;
data final06; set final06;
bdate06=bdate;
format bdate06 date.;
run;
data final07; set final07;
bdate07=bdate;
format bdate07 date.;
run;
data final08; set final08;
bdate08=bdate;
format bdate08 date.;
run;
***************************************************************************;
** TDate    ;
** New Final Variable name :  tdateYR ;
***************************************************************************;
data final00; set final00;
tdate00=tdate-21916; * Conversion from an Excel date by subtract 21916*;
format tdate00 date.;
run;
data final01; set final01;
tdate01=tdate-21916; * Conversion from an Excel date by subtract 21916*;
format tdate01 date.;
run;
data final02; set final02;
tdate02=tdate-21916; * Conversion from an Excel date by subtract 21916*;
format tdate02 date.;
run;
data final03; set final03;
tdate03=tdate-21916; * Conversion from an Excel date by subtract 21916*;
format tdate03 date.;
run;
data final04; set final04;
tdate04=tdate1-21916; * Conversion from an Excel date by subtract 21916*;
format tdate04 date.;
run;
APPENDIX A (continued)

data final05; set final05;
tdate05=tdate;
format tdate05 date.;
run;
data final06; set final06;
tdate06=tdate;
format tdate06 date.;
run;
data final07; set final07;
tdate07=tdate;
format tdate07 date.;
run;
data final08; set final08;
tdate08=tdate;
format tdate08 date.;
run;
********************************************;
** age ht and wt    ;
** New Final Variable name :  ageYR htYR wtYR ;
********************************************;
data final00; set final00;
age00=age;
ht00=ht;
wt00=wt;
school00=cschool+0;
* Grade listed from 1 to 9????? and also K???:
grade00=0;
if grade ="UN" then grade00=.;
else do;
  if grade = "K" then grade00=0;
  else grade00 = grade+0;
end;
*if grade00~=. then grade00=grade00-1; * subtract 1 as 09 is really grade 8??????;
********************************************;
zip00=zipcode;
primlang00=primary_language;
sex00=sex;
homelang00=.;
format tdate00 date.;
run;
data final01; set final01;
age01=age;
ht01=ht;
wt01=wt;
school01=school+0;
* Grade listed from K to 08;
grade01=0;
if grade ="UN" then grade01=.;
else do;
  if grade = "K" then grade01=0;
  else grade01 = grade+0;
APPENDIX A (continued)

end;
***********************************************************************;
zip01=zip_code;
primlang01=primary_language;
sex01=sex;
lunch01=lunch+0;
homelang01=.;
run;
data final02; set final02;
age02=age;
ht02=ht;
wt02=wt;
school02=cschool+0;
***********************************************************************;
* Grade listed from K to 08 ;
grade02=0;
if grade ="UN" then grade02=.;
else do;
  if grade = "K" then grade02=0;
  else grade02 = grade+0;
end;
***********************************************************************;
zip02=zipcode;
primlang02=.;
sex02=sex;
lunch02=lunch+0;
homelang02=.;
run;
data final03; set final03;
age03=age;
ht03=ht;
wt03=wt;
school03=cschool+0;
zip03=zipcode;
primlang03=.;
sex03=sex;
lunch03=lunch+0;
homelang03=.;
run;
data final04; set final04;
age04=age;
ht04=ht;
wt04=wt;
school04=cschool+0;
zip04=zipcode;
primlang04=.;
sex04=sex;
lunch04=lunch;
homelang04=.;
run;
data final05; set final05;
age05=floor ((intck('month',bdate05,tdate05) + (day(tdate05) <
day(bdate05)))/12);
ht05=ht;
APPENDIX A (continued)

```
wt05=wt;
school05=school_code+0;
zip05=.;
primlang05=.;
sex05=sex;
lunch05=lunch_code+0;
homelang05=.;
run;
data final06; set final06;
age06=floor ((intck('month',bdate06,tdate06) - (day(tdate06) <
day(bdate06)))/12);
ht06=ht;
wt06=wt;
school06=school_code+0;
zip06=.;
primlang06=.;
sex06=sex;
lunch06=lunch_code+0;
homelang06=.;
run;
data final07; set final07;
age07=floor ((intck('month',bdate07,tdate07) - (day(tdate07) <
day(bdate07)))/12);
ht07=ht;
rename ht_ =ht;
rename wt_ =wt;
wt07=wt;
school07=school_code+0;
 ********************************************;
* Grade listed from blank to 8???? no K???
grade07=0;
if grade ="UN" then grade07=.
else do;
  if grade = "K" then grade07=0;
  else if grade = "JK" then grade08=99;
  else grade07 = grade+0;
end;
 ********************************************;
zip07=zipcode;
primlang07=Student_Primary_Language;
sex07=sex;
lunch07=Lunch_Code+0;
homelang07=home_language;
run;
data final08; set final08;
age08=age;
ht08=ht;
rename ht_ =ht;
rename wt_ =wt;
wt08=wt;
school08=school_code+0;
homelang08=home_language;
 ********************************************;
* Grade listed from 1 to 9???? and also K???
```
grade08=0;
if grade ="UN" then grade08=.;
else do;
    if grade = "K" then grade08=0;
    else if grade = "JK" then grade08=99;
    else if grade = "MK" then grade08=98;
    else grade08 = grade+0;
end;
*if grade08=-. then grade08=grade08-1; * subtract 1 as 09 is really grade 8??????;

zip08=zip_code;
primlang08=primary_language;
sex08=sex;
lunch08=lunch_code+0;
format tdate08 date.;
run;
*******************************************************************************;

** BMI
*******************************************************************************;
data final00; *set final00;
*proc freq;
*tables bmi;
*title 'Missing Final 00 PRE --------';
*run;
data final00; set final00;
bmi00=bmi;
bmiz00=bmiz;
bmipct00=bmipct;
*delete improbable weight values;
    if bmiz00 le -4 or bmiz00 ge 5 then delete;
run;
*proc freq;
*tables bmi00;
*title 'Missing Final 00 POST --------';
*run;
data final01; set final01;
bmi01=bmi;
bmiz01=bmiz;
bmipct01=bmipct;
*delete improbable weight values;
    if bmiz01 le -4 or bmiz01 ge 5 then delete;
run;

data final02; set final02;
bmi02=bmi;
bmiz02=bmiz;
bmipct02=bmipct;
*delete improbable weight values;
    if bmiz02 le -4 or bmiz02 ge 5 then delete;
run;

data final03; set final03;
bmi03=bmi;
APPENDIX A (continued)

bmiz03=bmiz;
bmipct03=bmipct;
*delete improbable weight values;
  if bmiz03 le -4 or bmiz03 ge 5 then delete;
run;
data final04; set final04;
bmi04=bmi;
bmiz04=bmiz;
bmipct04=bmipct;
*delete improbable weight values;
  if bmiz04 le -4 or bmiz04 ge 5 then delete;
run;
*/CHECKING WEIGHT STATUS TO FIND MISSING DATA;
*/data final04_A; */set final04;
*Weight status prevalence 2004 for full population;*
*categorize bmi for 2004;
  */if bmipct04=. then weightstatus=.;
  */else if bmipct04<5 then weightstatus=1;
  */else if 5<=bmipct04<85 then weightstatus=2;
  */else if 85<=bmipct04<95 then weightstatus=3;
  */else if bmipct04>=95 then weightstatus=4;
*delete people younger than 5 and older than 14;
  */if age04<5 then delete;
  */if age04>14 then delete;
*delete improbable weight values;
  */if bmiz04 le -4 or bmiz04 ge 5 then delete;
  */if waz le -6 or waz ge 6 then delete;
  */if haz le -6 or haz ge 6 then delete;*/
*create formats;
  */proc format;
    */value weightstatus 1="less than 5th percentile (underweight)"
       2="5th to less than 85th percentile (healthy weight)"
       3="85th to less than 95th percentile (overweight)"
       4="equal to or greater than 95th percentile (obese)";*
  */run;
  */apply formats;
  */data final04_A;
    */set final04_A;
    */format weightstatus weightstatus.;
  */run;
*/Data final04_1;
*/Set final04_1;
*/proc freq data = final04_A;
*/tables weightstatus;
*/run;
*/FINISHED CHECKING;
data final05; set final05;
APPENDIX A (continued)

bmi05=bmi;
bmiz05=bmiz;
bmpct05=bmpct;
*delete improbable weight values;
   if bmiz05 le -4 or bmiz05 ge 5 then delete;
run;
data final06; set final06;
bmi06=bmi;
bmiz06=bmiz;
bmpct06=bmpct;
*delete improbable weight values;
   if bmiz06 le -4 or bmiz06 ge 5 then delete;
run;
data final07; set final07;
bmi07=bmi;
bmiz07=bmiz;
bmpct07=bmpct;
*delete improbable weight values;
   if bmiz07 le -4 or bmiz07 ge 5 then delete;
run;
data final08; set final08;
bmi08=bmi;
bmiz08=bmiz;
bmpct08=bmpct;
*delete improbable weight values;
   if bmiz08 le -4 or bmiz08 ge 5 then delete;
run;
******************************************************************************;
*Add grade to years missing: 03, 04, 05, 06;
* Grade K = 0
* Grade UN = 96
* Grade JK = 99
* Grade MK = 98
* Grade PK = 97
******************************************************************************;
******************************************************************************;
*Add grade to years missing: 03;
******************************************************************************;
data addgrade03; set longbmi.grade03;
keep Student_Id Grade;
run;
data addgrade03; set addgrade03;
if trim(left(Student_Id))="." then sid_num=.;
   else do;
      sid_num=Student_Id+0;
      sidf=sid_num;
      drop sid_num;
   end;
if sidf=. then sidmissing=1; else sidmissing=0;
run;
data addgrade03; set addgrade03;
drop sidmissing Student_Id;
Grade03f=0;
if grade ="UN" then Grade03f=96;
else do;
   if grade = "K" then Grade03f=0;
   else if grade = "JK" then grade03f=99;
   else if grade = "MK" then grade03f=98;
   else if grade = "PK" then grade03f=97;
   else Grade03f = grade+0;
end;
drop Grade;
run;
proc sort data=addgrade03;
   by sidf;
run;
proc sort data=final03;
   by sidf;
run;
data grademerge03; set addgrade03;
   merge addgrade03 final03;
   by sidf;
run;
data grademerge03; set grademerge03;
   if sidf=.
   then delete;
run;
**********************************************************************;
*data grademerge03; *set grademerge03;
*Grade listed from K to 08  ;
*Grade03f=0;
*if Grade03f ="UN" then Grade03f=.;
*else do;
   *if Grade03f = "K" then Grade03f=0;
   *else Grade03f = Grade03f+0;
*end;
******
**********************************************************************;
*Add grade to years missing: 04;
**********************************************************************;
data addgrade04; set longbmi.grade04;
keep Student_Id Grade;
run;
data addgrade04; set addgrade04;
if trim(left(Student_Id))="." then sid_num=.;
else do;
   sid_num=Student_Id+0;
sidf=sid_num;
   drop sid_num;
end;
if sidf=. then sidmissing=1; else sidmissing=0;
run;
data addgrade04; set addgrade04;
drop sidmissing Student_Id;
Grade04f=0;
if grade ="UN" then Grade04f=96;
else do;
  if grade = "K" then Grade04f=0;
  else if grade = "JK" then grade04f=99;
  else if grade = "MK" then grade04f=98;
  else if grade = "PK" then grade04f=97;
  else Grade04f = grade+0;
end;
drop Grade;
run;
proc sort data=addgrade04;
  by sidf;
run;
proc sort data=final04;
  by sidf;
run;
data grademerge04; set addgrade04;
  merge addgrade04 final04;
  by sidf;
run;
data grademerge04; set grademerge04;
if sidf=. then delete;
run;
*******************************************************************************;
*Add grade to years missing: 05;
*******************************************************************************;
data addgrade05; set longbmi.gradeofive;
keep Student_Id Grade;
run;
data addgrade05; set addgrade05;
if trim(left(Student_Id))="." then sid_num=.;
else do;
sid_num=Student_Id+0;
sidf=sid_num;
drop sid_num;
end;
if sidf=. then sidmissing=1; else sidmissing=0;
run;
data addgrade05; set addgrade05;
drop sidmissing Student_Id;
Grade05f=0;
if grade ="UN" then Grade05f=96;
else do;
  if grade = "K" then Grade05f=0;
  else if grade = "JK" then grade05f=99;
  else if grade = "MK" then grade05f=98;
  else if grade = "PK" then grade05f=97;
  else Grade05f = grade+0;
end;
drop Grade;
run;
proc contents data=final05; run;
APPENDIX A (continued)

```sas
proc sort data=addgrade05;
  by sidf;
run;
proc sort data=final05;
  by sidf;
run;
data grademerge05; set addgrade05;
  merge addgrade05 final05;
  by sidf;
run;
data grademerge05; set grademerge05;
if sidf=. then delete;
run;
******************************************************************************;
*Add grade to years missing: 06;******************************************************************************;
data addgrade06; set longbmi.grade06;
  keep Student_Id Grade;
run;
data addgrade06; set addgrade06;
if trim(left(Student_Id))="." then sid_num=.;
else do;
  sid_num=Student_Id+0;
  sidf=sid_num;
  drop sid_num;
end;
if sidf=. then sidmissing=1; else sidmissing=0;
run;
data addgrade06; set addgrade06;
drop sidmissing Student_Id;
Grade06f=0;
if grade ="UN" then Grade06f=96;
else do;
  if grade = "K" then Grade06f=0;
  else if grade = "JK" then grade06f=99;
  else if grade = "MK" then grade06f=98;
  else if grade = "PK" then grade06f=97;
  else Grade06f = grade+0;
end;
drop Grade;
run;

proc contents data=final06; run;
proc sort data=addgrade06;
  by sidf;
run;
proc sort data=final06;
  by sidf;
run;
data grademerge06; set addgrade06;
  merge addgrade06 final06;
  by sidf;
run;
```
APPENDIX A (continued)

data grademerge06; set grademerge06;
if sidf=. then delete;
run;
******************************************************************************;
*Add lunch to years missing: 00;******************************************************************************;
data addlunch00; set longbmi.lunch00;
keep Student_Id Food_Service_Code;
run;
data addlunch00; set addlunch00;
lunch00=Food_Service_Code+0;
drop Food_Service_Code;
run;
data addlunch00; set addlunch00;
if trim(left(Student_Id))="." then sid_num=.;
else do;
sid_num=Student_Id+0;
sidf=sid_num;
   drop sid_num;
end;
if sidf=. then sidmissing=1; else sidmissing=0;
run;
data addlunch00; set addlunch00;
drop sidmissing Student_Id;
Lunch00f=lunch00+0;
drop lunch00;
run;
proc contents data=final00; run;
proc sort data=addlunch00 nodupkey;
   by sidf;
run;
proc sort data=final00;
   by sidf;
run;
data grademerge00; set addlunch00;
   merge addlunch00 final00;
   by sidf;
run;
data grademerge00; set grademerge00;
   if sidf=. then delete;
   if bmi00=. then delete;
run;
******************************************************************************;
*Checking missing Lunch*;
******************************************************************************;
data grademerge00; set grademerge00;
proc freq;
tables lunch00f;
title '--- Missing Lunch 00';
run;
data final01; set final01;
proc freq;
APPENDIX A (continued)

tables lunch01;
title '--- Missing Lunch 01';
run;
data final02; set final02;
proc freq;
tables lunch02; title '--- Missing Lunch 02';
run;
data final03; set final03;
proc freq;
tables lunch03; title '--- Missing Lunch 03';
run;
data final04; set final04;
proc freq;
tables lunch04; title '--- Missing Lunch 04';
run;
data final05; set final05;
proc freq;
tables lunch05; title '--- Missing Lunch 05';
run;
data final06; set final06;
proc freq;
tables lunch06; title '--- Missing Lunch 06';
run;
data final07; set final07;
proc freq;
tables lunch07; title '--- Missing Lunch 07';
run;
data final08; set final08;
proc freq;
tables lunch08; title '--- Missing Lunch 08';
run;
*****************************************************;
*Dropping unused variables*;
*****************************************************;
data grademerge00; set grademerge00;
drop bmi bmiz bmipct sid bdate tdate age ht wt Sex Grade ZipCode race
cschool Primary_Language sidmissing;
run;
data final01; set final01;
drop bmi bmiz bmipct sid bdate tdate age ht wt School_Code School Lunch
ethnic Ethnic_Code Lunch_code Sex Grade Zip_Code Home_Language
Student_Primary_Language race cschool Primary_Language sidmissing;
run;
data final02; set final02;
drop bmi bmiz bmipct sid bdate tdate age ht School_Code zipcodeschool Lunch
ethnic Ethnic_Code Lunch_code Sex Grade ZipCode Home_Language
Student_Primary_Language race cschool Primary_Language sidmissing;
run;
data final03; set final03;
drop bmi bmiz bmipct sid bdate tdate age ht School_Code School Lunch
ethnic Ethnic_Code Lunch_code Sex Grade grade03 ZipCode Home_Language
Student_Primary_Language race cschool Primary_Language sidmissing;
APPENDIX A (continued)

run;
data final04; set final04;
drop bmi bmiz bmipct sid bdate1 tdate1 age ht wt School_Code School
Lunch ethnic Ethnic_Code Lunch_code Sex Grade grade04 ZipCode Home_Language
Student_Primary_Language race cschool Primary_Language sidmissing;
run;
data final05; set final05;
drop bmi bmiz bmipct sid bdate tdate agemos ht wt School_Code
Ethnic_Code Lunch_code Sex sidmissing;
run;
data final06; set final06;
drop bmi bmiz bmipct sid bdate tdate ht wt SchoolCode agemos EthnicCode
Lunchcode Sex sidmissing;
run;
data final07; set final07;
drop bmi bmiz bmipct sid bdate tdate age ht wt School_Code Ethnic_Code
Lunch_code Sex Grade ZipCode Home_Language Student_Primary_Language
sidmissing;
run;
data final08; set final08;
drop bmi bmiz bmipct sid bdate tdate age ht wt School_Code Ethnic_Code
Lunch_code Grade Zip_Code Home_Language Primary_Language sidmissing gender;
run;

********************************************;
**Sorting Variables;
** Data needs sorting before merging
********************************************;

proc sort data=grademerge00;
   by sidf;
run;

proc sort data=final01;
   by sidf;
run;
proc sort data=final02;
   by sidf;
run;
proc sort data=final03;
   by sidf;
run;
proc sort data=final04;
   by sidf;
run;
proc sort data=final05;
   by sidf;
run;
proc sort data=final06;
   by sidf;
run;
proc sort data=final07;
   by sidf;
run;
**APPENDIX A (continued)**

```sas
proc sort data=final08;
  by sidf;
run;

*****************************************************************************;
**MERGE DATA**;
*****************************************************************************;

data finalMERGE;
  merge grademerge00 final01 final02 final03 final04 final05 final06 final07 final08;
    by sidf;
  run;

*****************************************************************************;
* Proc contents each Final Set ( 8 sets 00 - 07 ) *;
*****************************************************************************;

proc contents data=grademerge00 varnum;
  title 'Final00';
run;
proc contents data=final01 varnum;
  title 'Final01';
run;
proc contents data=final02 varnum;
  title 'Final02';
run;
proc contents data=final03 varnum;
  title 'Final03';
run;
proc contents data=final04 varnum;
  title 'Final04';
run;
proc contents data=final05 varnum;
  title 'Final05';
run;
proc contents data=final06 varnum;
  title 'Final06';
run;
proc contents data=final07 varnum;
  title 'Final07';
run;
proc contents data=final08 varnum;
  title 'Final08';
run;
proc contents data=Finalmerge varnum;
  title 'Finalmerge';
run;

*****************************************************************************;
**# of missing SIDs and Remove blank SIDs**;
*****************************************************************************;
data finalMERGE; set finalMerge;
  if sidf=.;
    *if sex=1 then sexnew=1;
```
APPENDIX A (continued)

*else if sex=2 then sexnew=0;
run;
******************************************************************************;
**Determine which years have data**;
******************************************************************************;
data finalMERGE; set finalMerge;
year00=0;
   If NOT missing(sidf) then do;
      if bmi00^= then year00=1; else year00=0;
      if bmi01^= then year01=1; else year01=0;
      if bmi02^= then year02=1; else year02=0;
      if bmi03^= then year03=1; else year03=0;
      if bmi04^= then year04=1; else year04=0;
      if bmi05^= then year05=1; else year05=0;
      if bmi06^= then year06=1; else year06=0;
      if bmi07^= then year07=1; else year07=0;
      if bmi08^= then year08=1; else year08=0;
   end;
run;
proc freq;
tables year00 year01 year02 year04 year04 year05 year06 year07 year08;
title '---Number per year----------';
run;
data finalMERGE; set finalMerge;
   valuespres="a012345678";
   valuespres="";
      if year00=1 then valuespres=trim(valuespres) || '0';
      if year01=1 then valuespres=trim(valuespres) || '1';
      if year02=1 then valuespres=trim(valuespres) || '2';
      if year03=1 then valuespres=trim(valuespres) || '3';
      if year04=1 then valuespres=trim(valuespres) || '4';
      if year05=1 then valuespres=trim(valuespres) || '5';
      if year06=1 then valuespres=trim(valuespres) || '6';
      if year07=1 then valuespres=trim(valuespres) || '7';
      if year08=1 then valuespres=trim(valuespres) || '8';
   run;
proc freq;
tables valuespres;
title '---ValuesPresent---------';
run;
data finalMERGE; set finalMerge;
   value=length(valuespres);
run;
proc freq;
tables value;
run;
data finalMERGE; set finalMerge;
exptime=0;
   if year02=1 then exptime=exptime+1;
   if year03=1 then exptime=exptime+1;
   if year04=1 then exptime=exptime+1;
   if year05=1 then exptime=exptime+1;
   if year06=1 then exptime=exptime+1;
APPENDIX A (continued)

if year07=1 then exptime=exptime+1;
if year08=1 then exptime=exptime+1;
run;
proc freq;
tables exptime;
run;

data finalmerge; set finalmerge;
exptime2=0; exptime3=0; exptime4=0; exptime5=0;
exptime6=0; exptime7=0; exptime8=0;
counter=0;
if year02=1 then do;
exptime2=counter+1;
counter=counter+1;
end;
if year03=1 then do;
exptime3=counter+1;
counter=counter+1;
end;
if year04=1 then do;
exptime4=counter+1;
counter=counter+1;
end;
if year05=1 then do;
exptime5=counter+1;
counter=counter+1;
end;
if year06=1 then do;
exptime6=counter+1;
counter=counter+1;
end;
if year07=1 then do;
exptime7=counter+1;
counter=counter+1;
end;
if year08=1 then exptime8=counter+1;
run;
data finalMERGE; set finalMerge;
proc freq;
tables exptime2 exptime3 exptime4 exptime5 exptime6 exptime7 exptime8;
run;
*******************************************************************************
** Choose ethnicity
** Choose the last ethnicity used
** Ethnic error records a change in the ethnicity values among years
*******************************************************************************;

data finalMERGE; set finalMerge;
ethnicerror=0;
if ethnic00~= then ethnic=ethnic00;
if ethnic01~= then ethnic=ethnic01;
if ethnic02~= then ethnic=ethnic02;
if ethnic03~= then ethnic=ethnic03;
if ethnic04~= then ethnic=ethnic04;
APPENDIX A (continued)

if ethnic05~= then ethnic=ethnic05;
if ethnic06~= then ethnic=ethnic06;
if ethnic07~= then ethnic=ethnic07;
if ethnic08~= then ethnic=ethnic08;
if ((ethnic00~= and ethnic01~) and (ethnic00~=ethnic01)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic01~= and ethnic02~) and (ethnic01~=ethnic02)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic02~= and ethnic03~) and (ethnic02~=ethnic03)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic03~= and ethnic04~) and (ethnic03~=ethnic04)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic04~= and ethnic05~) and (ethnic04~=ethnic05)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic05~= and ethnic06~) and (ethnic05~=ethnic06)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic06~= and ethnic07~) and (ethnic06~=ethnic07)) then
  do; ethnicerror=ethnicerror+1; end;
if ((ethnic07~= and ethnic08~) and (ethnic07~=ethnic08)) then
  do; ethnicerror=ethnicerror+1; end;
  drop ethnic00 ethnic01 ethnic02 ethnic03 ethnic04 ethnic05 ethnic06 ethnic07 ethnic08;
run;
proc freq;
tables ethnicerror ethnic;
  title '---Ethnicity Fixes ---------';
run;

******************************************************************************;
** Choose Birthdate ;
** Choose the last birthdate used
** bdateerror records a change in the birthdate values among years
*******************************************************************************;

data finalMERGE; set finalMerge;
bdateerror=0; bdateyes=0;
  format bdate date.;
  if bdate00~= . and bdate00=0 then do bdateyes=1; bdate=bdate00;
end;
  if bdate01~= . and bdate01=0 then do bdateyes=1;
bdate=bdate01;end;
  if bdate02~= . and bdate02=0 then do bdateyes=1; bdate=bdate02;
end;
  if bdate03~= . and bdate03=0 then do bdateyes=1;
bdate=bdate03;end;
  if bdate04~= . and bdate04=0 then do bdateyes=1; bdate=bdate04;
end;
  if bdate05~= . and bdate05=0 then do bdateyes=1; bdate=bdate05;
end;
  if bdate06~= . and bdate06=0 then do bdateyes=1;
bdate=bdate06;end;
  if bdate07~= . and bdate07=0 then do bdateyes=1;
bdate=bdate07;end;
  if bdate08~= . and bdate08=0 then do bdateyes=1;
APPENDIX A (continued)

bdate=bdate08; end;
if ((bdate00~=. and bdate01~=. and bdate00~=0 and bdate01~=0) and
(bdate00~=bdate01)) then do; bdateerror=bdateerror+1; end;
if ((bdate01~=. and bdate02~=. and bdate01~=0 and bdate02~=0) and
(bdate01~=bdate02)) then do; bdateerror=bdateerror+1; end;
if ((bdate02~=. and bdate03~=. and bdate02~=0 and bdate03~=0) and
(bdate02~=bdate03)) then do; bdateerror=bdateerror+1; end;
if ((bdate03~=. and bdate04~=. and bdate03~=0 and bdate04~=0) and
(bdate03~=bdate04)) then do; bdateerror=bdateerror+1; end;
if ((bdate04~=. and bdate05~=. and bdate04~=0 and bdate05~=0) and
(bdate04~=bdate05)) then do; bdateerror=bdateerror+1; end;
if ((bdate05~=. and bdate06~=. and bdate05~=0 and bdate06~=0) and
(bdate05~=bdate06)) then do; bdateerror=bdateerror+1; end;
if ((bdate06~=. and bdate07~=. and bdate06~=0 and bdate07~=0) and
(bdate06~=bdate07)) then do; bdateerror=bdateerror+1; end;
if ((bdate07~=. and bdate08~=. and bdate07~=0 and bdate08~=0) and
(bdate07~=bdate08)) then do; bdateerror=bdateerror+1; end;
if bdate=0 or bdate=. then bdateerror=9;
*drop bdate00 bdate01 bdate02 bdate03 bdate04 bdate05 bdate06
bdate07 bdate08;
run;
proc freq;
tables bdateerror bdateyes /norow nocol missing;
title '--BDATE Errors ---------';
run;

**********************************************************************;
** Choose Sex   ;
** Choose the last birthdate used
** Sexerror records a change in the ethnicity values among years
**********************************************************************;

data finalMERGE; set finalMerge;
* sex will be the last sex;
sexerror=9;
if sex00~=. then do; sex=sex00; sexerror=sexerror-1; end;
if sex01~=. then do; sex=sex01; sexerror=sexerror-1; end;
if sex02~=. then do; sex=sex02; sexerror=sexerror-1; end;
if sex03~=. then do; sex=sex03; sexerror=sexerror-1; end;
if sex04~=. then do; sex=sex04; sexerror=sexerror-1; end;
if sex05~=. then do; sex=sex05; sexerror=sexerror-1; end;
if sex06~=. then do; sex=sex06; sexerror=sexerror-1; end;
if sex07~=. then do; sex=sex07; sexerror=sexerror-1; end;
if sex08~=. then do; sex=sex08; sexerror=sexerror-1; end;
* drop sex00 sex01 sex02 sex03 sex04 sex05 sex06 sex07 sex08;
run;
data finalMERGE; set finalMerge;
if sex=1 then sexnew=1;
if sex=2 then sexnew=0;
run;
proc freq data=finalMerge;
tables sexerror sex sexerror*sex sex00*year00 sex01*year01 sex02*year02
sex03*year03 sex04*year04 sex05*year05 sex06*year06 sex07*year07 sex08*year08
/norow nocol missing;
title '--sex fixes ---------';
run;

*----------------------------------------------------------------------*;
** Choose zip ;
**  Choose the last zip used
*----------------------------------------------------------------------*;

data finalMERGE; set finalMerge;
* zip will be the last zip_code;
  if zip00~=. then do; zip=zip00; end;
    if zip01~=. then do; zip=zip01; end;
    if zip02~=. then do; zip=zip02; end;
    if zip03~=. then do; zip=zip03; end;
    if zip04~=. then do; zip=zip04; end;
    if zip05~=. then do; zip=zip05; end;
    if zip06~=. then do; zip=zip06; end;
    if zip07~=. then do; zip=zip07; end;
    if zip08~=. then do; zip=zip08; end;
  drop zip00 zip01 zip02 zip03 zip04 zip05 zip06 zip07 zip08;
run;
proc freq data=finalMerge;
tables zip /norow nocol missing;
title '--zip fixes ---------';
run;

*----------------------------------------------------------------------*;
** Choose primlang ;
**  Choose the last primlang used
*----------------------------------------------------------------------*;

data finalMERGE; set finalMerge;
* primlang will be the last primlang;
  primlang=0;
    if primlang00~=. then do; primlang=primlang00+0; end;
    if primlang01~=. then do; primlang=primlang01+0; end;
    if primlang02~=. then do; primlang=primlang02+0; end;
    if primlang03~=. then do; primlang=primlang03+0; end;
    if primlang04~=. then do; primlang=primlang04+0; end;
    if primlang05~=. then do; primlang=primlang05+0; end;
    if primlang06~=. then do; primlang=primlang06+0; end;
    if primlang07~=. then do; primlang=primlang07+0; end;
    if primlang08~=. then do; primlang=primlang08+0; end;
  drop primlang00 primlang01 primlang02 primlang03 primlang04
      primlang05 primlang06 primlang07 primlang08;
run;
proc freq data=finalMerge;
tables primlang /norow nocol missing;
title '--Primlang fixes ---------';
run;

*----------------------------------------------------------------------*;
** Choose homelang ;
**  Choose the last primlang used
*----------------------------------------------------------------------;
APPENDIX A (continued)

```plaintext
data finalMERGE; set finalMerge;
  * primlang will be the last primlang;
    if homelang00=. then do; homelang=homelang00; end;
    if homelang01=. then do; homelang=homelang01; end;
    if homelang02=. then do; homelang=homelang02; end;
    if homelang03=. then do; homelang=homelang03; end;
    if homelang04=. then do; homelang=homelang04; end;
    if homelang05=. then do; homelang=homelang05; end;
    if homelang06=. then do; homelang=homelang06; end;
    if homelang07=. then do; homelang=homelang07; end;
    if homelang08=. then do; homelang=homelang08; end;
  drop homelang00 homelang01 homelang02 homelang03 homelang04
    homelang05 homelang06 homelang07 homelang08;
run;
proc freq data=finalMerge;
tables  homelang /norow nocol missing;
title '---homelang fixes ---------';
run;

********************************************;
** Fix school lunch  ;
**  if there is data on both sides of a year then this fills in the
**  if the two previous years or two following years are the same then
** schluncherror records that a value was fixed
********************************************;

data finalMERGE; set finalMerge;
  * Fixing blank school lunches will be the last value;
    schluncherror=0;
    if (lunch00f=. and year00=1) then do;
      if ((lunch01=lunch02 and year01=1)and (year02=1)) then do;
        lunch00f=lunch01; schluncherror=schluncherror+1; end;
      end;
    if (lunch01=. and year01=1) then do;
      if ((lunch00f=lunch02 and year00=1)and (year02=1)) then do;
        lunch01=lunch00f; schluncherror=schluncherror+1; end;
      else if ((lunch02=lunch03 and year02=1)and (year03=1)) then do;
        lunch01=lunch02; schluncherror=schluncherror+1; end;
      else if (lunch00f=. and year00=1) then do;
        lunch01=lunch00f; schluncherror=schluncherror+1; end;
      end;
    if (lunch02=. and year02=1) then do;
      if ((lunch01=lunch03 and year01=1)and (year03=1)) then do;
        lunch02=lunch01; schluncherror=schluncherror+1; end;
      else if ((lunch03=lunch04 and year03=1)and (year04=1)) then do;
        lunch02=lunch03; schluncherror=schluncherror+1; end;
      else if ((lunch01=lunch00f and year01=1)and (year00=1)) then do;
        lunch02=lunch01; schluncherror=schluncherror+1; end;
      end;
    if (lunch03=. and year03=1) then do;
      if ((lunch02=lunch04 and year02=1)and (year04=1)) then do;
        lunch03=lunch02; schluncherror=schluncherror+1; end;
      ```
else if ((lunch04=lunch05 and year04=1) and (year05=1)) then do;
  lunch03=lunch04; schluncherror=schluncherror+1; end;
else if ((lunch01=lunch02 and year01=1) and (year02=1)) then do;
  lunch03=lunch02; schluncherror=schluncherror+1; end;
end;
if (lunch04=. and year04=1) then do;
  if ((lunch03=lunch05 and year03=1) and (year05=1)) then do;
    lunch04=lunch03; schluncherror=schluncherror+1; end;
  else if ((lunch02=lunch03 and year02=1) and (year03=1)) then do;
    lunch04=lunch03; schluncherror=schluncherror+1; end;
  end;
if (lunch05=. and year05=1) then do;
  if ((lunch04=lunch06 and year04=1) and (year06=1)) then do;
    lunch05=lunch04; schluncherror=schluncherror+1; end;
  else if ((lunch02=lunch03 and year02=1) and (year03=1)) then do;
    lunch05=lunch04; schluncherror=schluncherror+1; end;
  end;
if (lunch06=. and year06=1) then do;
  if ((lunch05=lunch07 and year05=1) and (year07=1)) then do;
    lunch06=lunch05; schluncherror=schluncherror+1; end;
  else if ((lunch04=lunch05 and year04=1) and (year05=1)) then do;
    lunch06=lunch05; schluncherror=schluncherror+1; end;
  end;
if (lunch07=. and year07=1) then do;
  if ((lunch06=lunch08 and year06=1) and (year08=1)) then do;
    lunch07=lunch06; schluncherror=schluncherror+1; end;
  else if ((lunch05=lunch06 and year05=1) and (year06=1)) then do;
    lunch07=lunch06; schluncherror=schluncherror+1; end;
  end;
if (lunch08=. and year08=1) then do;
  if ((lunch07=lunch06 and year07=1) and (year06=1)) then do;
    lunch08=lunch07; schluncherror=schluncherror+1; end;
run;
proc freq data=finalMerge;
tables schluncherror lunch00f*year00 lunch01*year01 lunch02*year02 lunch03*year03 lunch04*year04 lunch05*year05 lunch06*year06 lunch07*year07 lunch08*year08 / norow nocol missing;
title 'lunch fixes' ;
run;
------------------------------------------------------------------------;
** Fix Grade ;
** if there is data on both sides of a grade then this fills in the middle data
** if the two previous years or two following years are the same then the missing is applied
** gradeerror records that a value was fixed
------------------------------------------------------------------------;
APPENDIX A (continued)

```sas
data finalMERGE; set finalMerge;
  gradeerror=0;
  if (grade00=. and year00=1) then do;
    if ((grade01=. and year01=1) and (grade01<97 and grade01~=0)) then do; grade00=grade01-1; gradeerror=gradeerror+1; end;
    end;
  if (grade01=. and year01=1) then do;
    if ((grade02=. and year02=1) and (grade02<97 and grade02~=0)) then do; grade01=grade02-1; gradeerror=gradeerror+1; end;
    else if ((grade00=. and year00=1) and (grade00<97 and grade00~=0)) then do; grade01=grade00+1; gradeerror=gradeerror+1; end;
  end;
if (grade02=. and year02=1) then do;
  if ((grade03f=. and year03=1) and (grade03f<97 and grade03f~=0)) then do; grade02=grade03f-1; gradeerror=gradeerror+1; end;
  else if ((grade02=. and year02=1) and (grade02<97 and grade02~=0)) then do; grade02=grade01+1; gradeerror=gradeerror+1; end;
  end;
if (grade03f=. and year03=1) then do;
  if ((grade04f=. and year04=1) and (grade04f<97 and grade04f~=0)) then do; grade03f=grade04f-1; gradeerror=gradeerror+1; end;
  else if ((grade02=. and year02=1) and (grade02<97 and grade02~=0)) then do; grade03f=grade02+1; gradeerror=gradeerror+1; end;
  end;
if (grade04f=. and year04=1) then do;
  if ((grade05f=. and year05=1) and (grade05f<97 and grade05f~=0)) then do; grade04f=grade05f-1; gradeerror=gradeerror+1; end;
  else if ((grade03f=. and year03=1) and (grade03f<97 and grade03f~=0)) then do; grade04f=grade03f+1; gradeerror=gradeerror+1; end;
  end;
if (grade05f=. and year05=1) then do;
  if ((grade06f=. and year06=1) and (grade06f<97 and grade06f~=0)) then do; grade05f=grade06f-1; gradeerror=gradeerror+1; end;
  else if ((grade04f=. and year04=1) and (grade04f<97 and grade04f~=0)) then do; grade05f=grade04f+1; gradeerror=gradeerror+1; end;
  end;
if (grade06f=. and year06=1) then do;
  if ((grade07=. and year07=1) and (grade07<97 and grade07~=0)) then do; grade06f=grade07-1; gradeerror=gradeerror+1; end;
  else if ((grade05f=. and year05=1) and (grade05f<97 and grade05f~=0)) then do; grade06f=grade05f+1; gradeerror=gradeerror+1; end;
  end;
if (grade07=. and year07=1) then do;
  if ((grade08=. and year08=1) and (grade08<97 and grade08~=0)) then do; grade07=grade08-1; gradeerror=gradeerror+1; end;
  else if ((grade06f=. and year06=1) and (grade06f<97 and grade06f~=0)) then do; grade07=grade06f+1; gradeerror=gradeerror+1; end;
  end;
if (grade08=. and year08=1) then do;
  if ((grade07=. and year07=1) and (grade07<97 and grade07~=0)) then do; grade08=grade07+1; gradeerror=gradeerror+1; end;
  end;
run;
proc freq data=finalMerge;
```

tables gradeerror grade00*year00 grade01*year01 grade02*year02 grade03f*year03 grade04f*year04 grade05f*year05 grade06f*year06 grade07*year07 grade08*year08 /norow nocol missing;
title '--Grade fixes ---------';
run;
***********************************************************************;
** Fix Age ;
** Fix the age by calculating from Bdate or subtracting or adding one
from previous year
** ageerror records a change
***********************************************************************;
data finalMERGE; set finalMerge;
ageerror=0;
   if (age00=. and year00=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate00,tdate00)- (day(tdate00) < day(bdate00))) / 12);
      ageerror=ageerror+1; end;
      else if (age01=.) then end age00=age01-1;
   end;
   if (age01=. and year01=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate01,tdate01)- (day(tdate01) < day(bdate01))) / 12);
      ageerror=ageerror+1; end;
      else if (age02=.) then end age01=age02-1;
   end;
   if (age02=. and year02=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate02,tdate02)- (day(tdate02) < day(bdate02))) / 12);
      ageerror=ageerror+1; end;
      else if (age03=.) then end age02=age03-1;
   end;
   if (age03=. and year03=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate03,tdate03)- (day(tdate03) < day(bdate03))) / 12);
      ageerror=ageerror+1; end;
      else if (age04=.) then end age03=age04-1;
   end;
   if (age04=. and year04=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate04,tdate04)- (day(tdate04) < day(bdate04))) / 12);
      ageerror=ageerror+1; end;
      else if (age05=.) then end age04=age05-1;
   end;
   if (age05=. and year05=1) then do;
      if( bdate~= . and bdate=0)then do; age00
         = floor((intck('month',bdate05,tdate05)- (day(tdate05) < day(bdate05))) / 12);
      ageerror=ageerror+1; end;
      else if (age06=.) then end age05=age06-1;end;
APPENDIX A (continued)

```
if (age06=. and year06=1) then do;
  if (bdate=. and bdate=0) then do; age00
    = floor((intck('month', bdate06, tdate06) - (day(tdate06) < day(bdate06))) / 12);
    ageerror=ageerror+1; end;
  else if (age07=. then age06=age07-1;end;
if (age07=. and year07=1) then do;
  if (bdate=. and bdate=0) then do; age00
    = floor((intck('month', bdate07, tdate07) - (day(tdate07) < day(bdate07))) / 12);
    ageerror=ageerror+1; end;
  else if (age08=. then age07=age08-1;end;
if (age08=. and year08=1) then do;
  if (bdate=. and bdate=0) then do; age00
    = floor((intck('month', bdate08, tdate08) - (day(tdate08) < day(bdate08))) / 12);
    ageerror=ageerror+1; end;
end;
run;
```

```
proc freq data=finalMerge;
  tables ageerror age00*year00 age01*year01 age02*year02 age03*year03
  age04*year04 age05*year05 age06*year06 age07*year07 age08*year08;
  title '---Age Fixes ---
run;
```

```
proc contents data=Finalmerge varnum;
  title 'Finalmerge ---
run;
```

```
proc freq data=Finalmerge ;
  tables year00 year01 year02 year03 year04 year05 year06 year07 year08;
  title 'Finalmerge ---
run;
```

```
proc freq data=Finalmerge ;
  tables school00 school01 school02 school03 school04 school05 school06
  school07 school08;
  title 'Finalmerge ---
run;
```

```
*proc freq data=Finalmerge;
*tables lunch00f lunch01 lunch02 lunch03 lunch04 lunch05 lunch06
  lunch07 lunch08;
*title 'Finalmerge ---
run;
```

```
*Assess weight status prevalence, cross sectionally;
**************************************************************;
```

```
data finalmerge; set finalmerge;
*ethnicity; * I changed ethnicnew to ethincategory;
  if Ethnic=. then ethniccategory=.;
  else if ethnic in (1) then ethniccategory=1; *white;
  else if ethnic in (2) then ethniccategory=2; *Black or
```
APPENDIX A (continued)

African American;
    else if ethnic in (3) then ethniccategory=3; *Asian;
    else if ethnic in (33, 34) then ethniccategory=4; *latino;
    else if ethnic in (4, 5, 6, 7, 8, 10, 11, 12, 14, 16, 17, 35, 36, 38, 39, 40, 49, 51) then ethniccategory=5; *other;
*Lunch Status;
    if lunch00f=. then lunchcategory00=.;
    else if lunch00f=1 or lunch00f=2 then lunchcategory00=1; /* free/reduced */
    else if lunch00f=3 or lunch00f=4 then lunchcategory00=0; /* paid */
    else if lunch01=. then lunchcategory01=.;
    else if lunch01=1 or lunch01=2 then lunchcategory01=1; /* free/reduced */
    else if lunch02=. then lunchcategory02=.;
    else if lunch02=1 or lunch02=2 then lunchcategory02=1; /* free/reduced */
    else if lunch03=. then lunchcategory03=.;
    else if lunch03=1 or lunch03=2 then lunchcategory03=1; /* free/reduced */
    else if lunch04=. then lunchcategory04=.;
    else if lunch04=1 or lunch04=2 then lunchcategory04=1; /* free/reduced */
    else if lunch05=. then lunchcategory05=.;
    else if lunch05=1 or lunch05=2 then lunchcategory05=1; /* free/reduced */
    else if lunch06=. then lunchcategory06=.;
    else if lunch06=1 or lunch06=2 then lunchcategory06=1; /* free/reduced */
    else if lunch07=. then lunchcategory07=.;
    else if lunch07=1 or lunch07=2 then lunchcategory07=1; /* free/reduced */
    else if lunch08=. then lunchcategory08=.;
    else if lunch08=1 or lunch08=2 then lunchcategory08=1; /* free/reduced */
    else if lunch09=. then lunchcategory09=.;
    else if lunch09=1 or lunch09=2 then lunchcategory09=1; /* free/reduced */
    else if lunch10=. then lunchcategory10=.;
    else if lunch10=1 or lunch10=2 then lunchcategory10=1; /* free/reduced */
    else if lunch11=. then lunchcategory11=.;
    else if lunch11=1 or lunch11=2 then lunchcategory11=1; /* free/reduced */
    else if lunch12=. then lunchcategory12=.;
    else if lunch12=1 or lunch12=2 then lunchcategory12=1; /* free/reduced */
    else if lunch13=. then lunchcategory13=.;
    else if lunch13=1 or lunch13=2 then lunchcategory13=1; /* free/reduced */
    else if lunch14=. then lunchcategory14=.;
    else if lunch14=1 or lunch14=2 then lunchcategory14=1; /* free/reduced */
    else if lunch15=. then lunchcategory15=.;
    else if lunch15=1 or lunch15=2 then lunchcategory15=1; /* free/reduced */
    else if lunch16=. then lunchcategory16=.;
    else if lunch16=1 or lunch16=2 then lunchcategory16=1; /* free/reduced */
    else if lunch17=. then lunchcategory17=.;
    else if lunch17=1 or lunch17=2 then lunchcategory17=1; /* free/reduced */
    else if lunch18=. then lunchcategory18=.;
    else if lunch18=1 or lunch18=2 then lunchcategory18=1; /* free/reduced */
    else if lunch19=. then lunchcategory19=.;
    else if lunch19=1 or lunch19=2 then lunchcategory19=1; /* free/reduced */
    else if lunch20=. then lunchcategory20=.;
    else if lunch20=1 or lunch20=2 then lunchcategory20=1; /* free/reduced */
    else if lunch21=. then lunchcategory21=.;
    else if lunch21=1 or lunch21=2 then lunchcategory21=1; /* free/reduced */
    else if lunch22=. then lunchcategory22=.;
    else if lunch22=1 or lunch22=2 then lunchcategory22=1; /* free/reduced */
    else if lunch23=. then lunchcategory23=.;
    else if lunch23=1 or lunch23=2 then lunchcategory23=1; /* free/reduced */
    else if lunch24=. then lunchcategory24=.;
    else if lunch24=1 or lunch24=2 then lunchcategory24=1; /* free/reduced */
    else if lunch25=. then lunchcategory25=.;
    else if lunch25=1 or lunch25=2 then lunchcategory25=1; /* free/reduced */
    else if lunch26=. then lunchcategory26=.;
    else if lunch26=1 or lunch26=2 then lunchcategory26=1; /* free/reduced */
    else if lunch27=. then lunchcategory27=.;
    else if lunch27=1 or lunch27=2 then lunchcategory27=1; /* free/reduced */
    else if lunch28=. then lunchcategory28=.;
    else if lunch28=1 or lunch28=2 then lunchcategory28=1; /* free/reduced */
    else if lunch29=. then lunchcategory29=.;
    else if lunch29=1 or lunch29=2 then lunchcategory29=1; /* free/reduced */
    else if lunch30=. then lunchcategory30=.;
    else if lunch30=1 or lunch30=2 then lunchcategory30=1; /* free/reduced */
    proc freq;
        tables ethniccategory;
        run;
    data finalmerge; set finalmerge;
    *YEAR 00 Weight status prevalence, 5 categories; 
        if bmipct00=. then weight_status00=.;
        else if bmipct00<5 then weight_status00=1;
        else if 5<=bmipct00<85 then weight_status00=2;
        else if 85<=bmipct00<95 then weight_status00=3;
        else if bmipct00>=95 then weight_status00=4;
    *Creating bivariate BMI variable;
APPENDIX A (continued)

if bmipct00= . then obese00= .; 
else if bmipct00<95 then obese00= 0; 
else if bmipct00>=95 then obese00= 1; 
*delete improbable weight values; 
*if bmiz00 le -4 or bmiz00 ge 5 then delete; 

*Indicator Variable; 
if bmi00 ne 0 then ind1= 1; 
else ind1= 0; 

*YEAR 01  Weight status prevalence, 5 categories; 
if bmipct01= . then weight_status01= .; 
else if bmipct01<5 then weight_status01= 1; 
else if 5<=bmipct01<85 then weight_status01= 2; 
else if 85<=bmipct01<95 then weight_status01= 3; 
else if bmipct01>=95 then weight_status01= 4; 
*Creating bivariate BMI variable; 
if bmipct01= . then obese01= .; 
else if bmipct01<95 then obese01= 0; 
else if bmipct01>=95 then obese01= 1; 
*delete improbable weight values; 
*if bmiz01 le -4 or bmiz01 ge 5 then delete; 

*Indicator Variable; 
if bmi01 ne 0 then ind2= 1; 
else ind2= 0; 

*YEAR 02  Weight status prevalence; 
if bmipct02= . then weight_status02= .; 
else if bmipct02<5 then weight_status02= 1; 
else if 5<=bmipct02<85 then weight_status02= 2; 
else if 85<=bmipct02<95 then weight_status02= 3; 
else if bmipct02>=95 then weight_status02= 4; 
*Creating bivariate BMI variable; 
if bmipct02= . then obese02= .; 
else if bmipct02<95 then obese02= 0; 
else if bmipct02>=95 then obese02= 1; 
*delete improbable weight values; 
*if bmiz02 le -4 or bmiz02 ge 5 then delete; 

*Indicator Variable; 
if bmi02 ne 0 then ind3= 1; 
else ind3= 0; 

*YEAR 03  Weight status prevalence; 
if bmipct03= . then weight_status03= .; 
else if bmipct03<5 then weight_status03= 1; 
else if 5<=bmipct03<85 then weight_status03= 2; 
else if 85<=bmipct03<95 then weight_status03= 3; 
else if bmipct03>=95 then weight_status03= 4; 
*Creating bivariate BMI variable; 
if bmipct03= . then obese03= .; 
else if bmipct03<95 then obese03= 0; 
else if bmipct03>=95 then obese03= 1; 
*delete improbable weight values; 
*if bmiz03 le -4 or bmiz03 ge 5 then delete; 

*Indicator Variable; 
if bmi03 ne 0 then ind4= 1; 
else ind4= 0; 

*YEAR 04  Weight status prevalence;
APPENDIX A (continued)

if bmipct04=. then weight_status04=.;
else if bmipct04<5 then weight_status04=1;
else if 5<=bmipct04<85 then weight_status04=2;
else if 85<=bmipct04<95 then weight_status04=3;
else if bmipct04>=95 then weight_status04=4;

*Creating bivariate BMI variable;
if bmipct04=. then obese04=.;
else if bmipct04<95 then obese04=0;
else if bmipct04>=95 then obese04=1;
*delete improbable weight values;
*if bmiz04 le -4 or bmiz04 ge 5 then delete;

*Indicator Variable;
if bmi04 ne 0 then ind5=1;
else ind5=0;

*YEAR 05  Weight status prevalence;
if bmipct05=. then weight_status05=.;
else if bmipct05<5 then weight_status05=1;
else if 5<=bmipct05<85 then weight_status05=2;
else if 85<=bmipct05<95 then weight_status05=3;
else if bmipct05>=95 then weight_status05=4;

*Creating bivariate BMI variable;
if bmipct05=. then obese05=.;
else if bmipct05<95 then obese05=0;
else if bmipct05>=95 then obese05=1;
*delete improbable weight values;
*if bmiz05 le -4 or bmiz05 ge 5 then delete;

*Indicator Variable;
if bmi05 ne 0 then ind6=1;
else ind6=0;

*YEAR 06  Weight status prevalence;
if bmipct06=. then weight_status06=.;
else if bmipct06<5 then weight_status06=1;
else if 5<=bmipct06<85 then weight_status06=2;
else if 85<=bmipct06<95 then weight_status06=3;
else if bmipct06>=95 then weight_status06=4;

*Creating bivariate BMI variable;
if bmipct06=. then obese06=.;
else if bmipct06<95 then obese06=0;
else if bmipct06>=95 then obese06=1;
*delete improbable weight values;
*if bmiz06 le -4 or bmiz06 ge 5 then delete;

*Indicator Variable;
if bmi06 ne 0 then ind7=1;
else ind7=0;

*YEAR 07  Weight status prevalence;
if bmipct07=. then weight_status07=.;
else if bmipct07<5 then weight_status07=1;
else if 5<=bmipct07<85 then weight_status07=2;
else if 85<=bmipct07<95 then weight_status07=3;
else if bmipct07>=95 then weight_status07=4;

*Creating bivariate BMI variable;
if bmipct07=. then obese07=.;
else if bmipct07<95 then obese07=0;
else if bmipct07>=95 then obese07=1;
APPENDIX A (continued)

*delete improbable weight values;
  *if bmi07 le -4 or bmiz07 ge 5 then delete;
*Indicator Variable;
  if bmi07 ne 0 then ind8=1;
  else ind8=0;
*YEAR 08      Weight status prevalence;
  if bmipct08= . then weight_status08= . ;
  else if 5<=bmipct08<85 then weight_status08=2;
  else if 85<=bmipct08<95 then weight_status08=3;
  else if bmipct08>=95 then weight_status08=4;
*Creating bivariate BMI variable;
  if bmipct08= . then obese08= . ;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
*delete improbable weight values;
  *if bmiz08 le -4 or bmiz08 ge 5 then delete;
*Indicator Variable;
  if bmi08 ne 0 then ind9=1;
  else ind9=0;
*Coding zip code;
data finalmerge; set finalmerge;
  proc freq data=finalmerge;
  tables zip;
  run;
  proc freq;
  title '---Weight Status x ethnic Cross Tables --------';
  tables ethinccategory*weight_status00 ethinccategory*weight_status01
  ethinccategory*weight_status02 ethinccategory*weight_status03
  ethinccategory*weight_status04 ethinccategory*weight_status05
  ethinccategory*weight_status06 ethinccategory*weight_status07
  ethinccategory*weight_status08;
  run;
  data finalmerge; set finalmerge;
  *Coding Zip codes;
  zipcode=0;
  if zip=2138 or zip='02138' then zipcode=1;
  else if zip=2139 or zip='02139' then zipcode=2;
  else if zip=2140 or zip='02140' then zipcode=3;
  else if zip=2141 or zip='02141' then zipcode=4;
  else if zip=2142 or zip='02142' then zipcode=5;
  else if zip=02173 then zipcode=6;
 data finalmerge; set finalmerge;
  proc freq data=finalmerge;
  tables zipcode;
  title 'zipcode';
  run;
  proc contents data=Finalmerge varnum;
  title 'Finalmerge - FINAL after adding categorical variables';
  run;
  data finalmerge; set finalmerge;
  proc freq;
  tables weight_status00 weight_status04 weight_status08;
  run;
APPENDIX A (continued)

*create formats;
proc format;

value weightf
1="less than 5th percentile (underweight)"
2="5th to less than 85th percentile (healthy weight)"
3="85th to less than 95th percentile (overweight)"
4="equal to or greater than 95th percentile (obese)";
value ethnic
1='White'
2='Black or African American'
3='Asian'
4='Latino'
5='other';
value lunch
1='Free/reduced'
0='Paid';
value sex
1='male'
2='female';
run;
*apply formats*;

data longbmi.finalmerge; set finalmerge;
format ethniccategory ethnic. sex sex. ;
format lunchcategory00 lunch. lunchcategory01 lunch. lunchcategory02 lunch. lunchcategory03 lunch. lunchcategory04 lunch. lunchcategory05 lunch. lunchcategory06 lunch. lunchcategory07 lunch. lunchcategory08 lunch. ;
format weight_status00 weightf. weight_status01 weightf. weight_status02 weightf. weight_status03 weightf. weight_status04 weightf. weight_status05 weightf. weight_status06 weightf. weight_status07 weightf. weight_status08 weightf. ;
proc freq;
tables lunchcategory00 lunchcategory04 lunchcategory08;
run;
data longbmi.finalmerge; set finalmerge;
proc freq;
tables ethnic*weight_status00 ethnic*weight_status01 ethnic*weight_status02 ethnic*weight_status03 ethnic*weight_status04 ethnic*weight_status05 ethnic*weight_status06 ethnic*weight_status07 ethnic*weight_status08;
run;
proc freq;
tables weight_status00 age00 age00*weight_status00 sex*weight_status00 ethniccategory*weight_status00 lunchcategory00*weight_status00 lunchcategory01*weight_status00 sex*lunchcategory00*weight_status00;
title 'Year 00 Weight Status Cross Tables----------';
run;
data finalmerge; set finalmerge;
proc freq;
tables weight_status01 age01 age01*weight_status01 sex*weight_status01 ethniccategory*weight_status01 lunchcategory01*weight_status01 ethniccategory*sex*weight_status01 age01*sex*weight_status01 sex*lunchcategory01*weight_status01;
APPENDIX A (continued)

run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status02 age02 age02*weight_status02
sex*weight_status02 ethniccategory*weight_status02
lunchcategory02*weight_status02 ethniccategory*sex*weight_status02
age02*sex*weight_status02 sex*lunchcategory02*weight_status02;
title 'Year 02 Weight Status Cross Tables --------
run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status03 age03 age03*weight_status03
sex*weight_status03 ethniccategory*weight_status03
lunchcategory03*weight_status03 ethniccategory*sex*weight_status03
age03*sex*weight_status03 sex*lunchcategory03*weight_status03;
title 'Year 03 Weight Status Cross Tables --------
run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status04 age04 age04*weight_status04
sex*weight_status04 ethniccategory*weight_status04
lunchcategory04*weight_status04 ethniccategory*sex*weight_status04
age04*sex*weight_status04 sex*lunchcategory04*weight_status04;
title 'Year 04 Weight Status Cross Tables --------
run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status05 age05 age05*weight_status05
sex*weight_status05 ethniccategory*weight_status05
lunchcategory05*weight_status05 ethniccategory*sex*weight_status05
age05*sex*weight_status05 sex*lunchcategory05*weight_status05;
title 'Year 05 Weight Status Cross Tables --------
run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status06 age06 age06*weight_status06
sex*weight_status06 ethniccategory*weight_status06
lunchcategory06*weight_status06 ethniccategory*sex*weight_status06
age06*sex*weight_status06 sex*lunchcategory06*weight_status06;
title 'Year 06 Weight Status Cross tables --------
run;

data finalmerge; set finalmerge;
proc freq;
tables weight_status07 age07 age07*weight_status07
sex*weight_status07 ethniccategory*weight_status07
lunchcategory07*weight_status07 ethniccategory*sex*weight_status07
age07*sex*weight_status07 sex*lunchcategory07*weight_status07;
title 'Year 07 Weight Status Cross Tables --------

run;

data finalmerge; set finalmerge;
proc freq;
  tables weight_status08 age08 age08*weight_status08 sex*weight_status08 ethniccategory*weight_status08 lunchcategory08*weight_status08 ethniccategory*sex*weight_status08 age08*sex*weight_status08 sex*lunchcategory08*weight_status08; title '---Year 08 Weight Status Cross Tables -------';
run;
************************;
*Create Time variable;
*************************
data time; set finalmerge;
proc freq;
  tables tdate00 tdate01 tdate02 tdate03 tdate04 tdate05 tdate06 tdate07 tdate08;
run;
data time; set finalmerge;
input date date7.;
sasdate=date;
format date mmddyy8.;
datalines;
15APR00 15APR01 12APR02 15APR02 15APR03 01MAR04 15MAR05 13MAR06 15MAR06 15APR07 15MAR08 ;
run;
proc print;
run;
*data finalmerge; *set finalmerge;
* stdytime = 99;
* if (year00=1) then stdytime=0;
* else if year01=1 then stdytime=1;
  else if year02=1 then stdytime=2;
  else if year03=1 then stdytime=3;
  else if year04=1 then stdytime=4;
  else if year05=1 then stdytime=5;
  else if year06=1 then stdytime=6;
  else if year07=1 then stdytime=7;
  else if year08=1 then stdytime=8;
  else if year09=1 then stdytime=9;
  run;
* proc freq data=Finalmerge ;
tables stdytime;
*title 'Finalmerge - stdytime';
APPENDIX A (continued)

*run;

*if tdate00=. then time=.;
*else if tdate01=. then time=.;
*else if tdate02=. then time=.;
*else if tdate03=. then time=.;
*else if tdate04=. then time=.;
*else if tdate05=. then time=.;
*else if tdate06=. then time=.;
*else if tdate07=. then time=.;
*else if tdate08=. then time=.;
*else if tdate00=14715 then time=0; *T1 2000;
*else if tdate01=15080 then time=1; *T2 2001;
* else if tdate02=15442 or 15445 then time=2; *T3 2002;
* else if tdate03=15810 then time=3; *T4 2003;
* else if tdate04=16131 then time=4; *T5 2004;
* else if tdate05=16510 then time=5; *T6 2005;
* else if tdate06=16873 or 16875 then time=6; *T7 2006;
* else if tdate07=17271 then time=7; *T8 2007;
* else if tdate08=17606 then time=8; *T9 2008;
*run;
*proc freq data=finalmerge;
*tables time;
*run
******************************************************************************

**************;
* Create HLM Files;
* ;
* The following un-merges the previously merged table FinalMerge;
* by creating tables Level1HLMYear00 to Level1HLMYear08;
* and Level2HLM;
******************************************************************************

******************************************************************************;
** Unmerge Data, Create Level 1 HLM Table ;
**
**************;
data final;
set finalMerge;
if proc freq;
tables weight_status00 weight_status08;
run;
data finalMerge;
set finalMerge;
if value=. then numyear=.;
else if value=2 then numyear=1;
else if value=3 then numyear=2;
else if value=4 then numyear=3;
else if value=5 then numyear=4;
APPENDIX A (continued)

```plaintext
else if value=6 then numyear=5;
else if value=7 then numyear=6;
else if value=8 then numyear=7;
else if value=9 then numyear=8;
else if value=10 then numyear=9;
run;
proc freq;
tables numyear;
run;
data finalMerge; set finalMerge;
if value>=8;
run;
data Level1HLMyear00;
set finalMerge;
numyear1=(value=2);
numyear2=(value=3);
numyear3=(value=4);
numyear4=(value=5);
numyear5=(value=6);
numyear6=(value=7);
numyear7=(value=8);
numyear8=(value=9);
numyear9=(value=10);
run;
proc freq;
tables numyear1 numyear2 numyear3 numyear4 numyear5 numyear6 numyear7 numyear8 numyear9;
run;
data Level1HLMyear00;
set finalMerge (keep= sidf lunchcategory00 tdate00 age00 grade00 bmi00 bmiz00 bmipct00 year00
rename= (tdate00=tdate
age00=age
grade00=grade
bmi00=bmi
bmiz00=bmiz
bmipct00=bmipct
lunchcategory00=lunchbi));
cumexp=0;
if year00=0 then delete;
drop year00; year=0;
run;

data Level1HLMyear01;
set finalMerge (keep= sidf lunchcategory01 tdate01 age01 grade01 bmi01 bmiz01 bmipct01 year01
rename= (tdate01=tdate
age01=age
grade01=grade
bmi01=bmi
bmiz01=bmiz
bmipct01=bmipct
lunchcategory01=lunchbi));
```

APPENDIX A (continued)

cumexp=0;
    if year01=0 then delete;
    drop year01; year=1;
run;

data Level1HLMyear02;
set finalMerge (keep= sidf lunchcategory02 tdate02 age02 grade02 bmi02 bmiz02 bmipct02 exptime2 year02
    rename= (tdate02=tdate
    age02=age
    grade02=grade
    bmi02=bmi
    bmiz02=bmiz
    bmipct02=bmipct
    lunchcategory02=lunchbi));

cumexp=exptime2; drop exptime2;
    if year02=0 then delete;
    drop year02; year=2;
run;

data Level1HLMyear03;
set finalMerge (keep= sidf lunchcategory03 tdate03 age03 Grade03f bmi03 bmiz03 bmipct03 exptime3 year03
    rename= (tdate03=tdate
    age03=age
    grade03f=grade
    bmi03=bmi
    bmiz03=bmiz
    bmipct03=bmipct
    lunchcategory03=lunchbi));

cumexp=exptime3; drop exptime3;
    if year03=0 then delete;
    drop year03; year=3;
run;

data Level1HLMyear04;
set finalMerge (keep= sidf lunchcategory04 tdate04 age04 Grade04f bmi04 bmiz04 bmipct04 exptime4 year04
    rename= (tdate04=tdate
    age04=age
    grade04f=grade
    bmi04=bmi
    bmiz04=bmiz
    bmipct04=bmipct
    lunchcategory04=lunchbi));

cumexp=exptime4; drop exptime4;
    if year04=0 then delete;
    drop year04; year=4;
run;

data Level1HLMyear05;
set finalMerge (keep= sidf lunchcategory05 tdate05 age05 Grade05f bmi05 bmiz05 bmipct05 exptime5 year05
    rename= (tdate05=tdate
    age05=age
    grade05f=grade
    bmi05=bmi
    bmiz05=bmiz
    bmipct05=bmipct
    lunchcategory05=lunchbi));
APPENDIX A (continued)

```
bmi05=bmi
bmiz05=bmiz
bmipct05=bmipct
lunchcategory05=lunchbi));
cumexp=exptime5; drop exptime5;
if year05=0 then delete;
drop year05; year=5;
run;
data Level1HLMyear06;
set finalMerge (keep= sidf lunchcategory06 tdate06 age06 Grade06f
bmi06 bmiz06 bmipct06 exptime6 year06
rename= (tdate06=tdate
age06=age
grade06f=grade
bmi06=bmi
bmiz06=bmiz
bmipct06=bmipct
lunchcategory06=lunchbi));
cumexp=exptime6; drop exptime6;
if year06=0 then delete;
drop year06; year=6;
run;
data Level1HLMyear07;
set finalMerge (keep= sidf lunchcategory07 tdate07 age07 Grade07
bmi07 bmiz07 bmipct07 exptime7 year07
rename= (tdate07=tdate
age07=age
grade07=grade
bmi07=bmi
bmiz07=bmiz
bmipct07=bmipct
lunchcategory07=lunchbi));
cumexp=exptime7; drop exptime7;
if year07=0 then delete;
drop year07; year=7;
run;
data Level1HLMyear08;
set finalMerge (keep= sidf lunchcategory08 tdate08 age08 Grade08
bmi08 bmiz08 bmipct08 exptime8 year08
rename= (tdate08=tdate
age08=age
grade08=grade
bmi08=bmi
bmiz08=bmiz
bmipct08=bmipct
lunchcategory08=lunchbi));
cumexp=exptime8; drop exptime8;
if year08=0 then delete;
drop year08; year=8;
run;
data Level2aHLM;
```
APPENDIX A (continued)

set finalMerge (keep= sidf ethniccategory primlang sex zipcode year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime);
if sex=1 then sex2=1;
else if sex=2 then sex2=0;
ethnic1=(ethniccategory=1);
ethnic2=(ethniccategory=2);
ethnic3=(ethniccategory=3);
ethnic4=(ethniccategory=4);
zip1=(zipcode=1);
zip2=(zipcode=2);
zip3=(zipcode=3);
zip4=(zipcode=4);
run;
proc freq;
tables zipcode;
run;
data Level2b;
set Level2aHLM (keep= sidf ethniccategory primlang sex2 zipcode);
if sex2=1;
run;
data Level2g;
set Level2aHLM (keep= sidf ethniccategory primlang sex2 zipcode);
if sex2=0;
run;
data longbmi.Level2HLM;
set Level2aHLM (keep= sidf ethnic1 ethnic2 ethnic3 ethnic4 zip1 zip2 zip3 zip4 year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime primlang sex2 zipcode);
run;
proc freq;
tables ethnic1 ethnic2 ethnic3 ethnic4 sex2;
run;
data longbmi.Level2BAAG;
set Level2aHLM (keep= sidf ethnic2 year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime primlang sex2 zipcode);
if sex2=0;
if ethnic2=1;
run;
data longbmi.Level2BAAB;
set Level2aHLM (keep= sidf ethnic2 year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime primlang sex2 zipcode);
if sex2=1;
if ethnic2=1;
run;
data longbmi.Level2WG;
set Level2aHLM (keep= sidf ethnic1 year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime primlang sex2 zipcode);
if sex2=0;
if ethnic1=0;
run;
data longbmi.Level2WB;
set Level2aHLM (keep= sidf ethnic1 year00 year01 year02 year03 year04 year05 year06 year07 year08 exptime primlang sex2 zipcode);
if sex2=1;
APPENDIX A (continued)

if ethnic1=0;
run;
proc freq;
tables ethnic1 ethnic2 ethnic3 ethnic4;
run;
data longbmi.Level1HLM;
   set Level1HLMyear00 Level1HLMyear01 Level1HLMyear02 Level1HLMyear03 Level1HLMyear04 Level1HLMyear05 Level1HLMyear06 Level1HLMyear07 Level1HLMyear08;
run;
proc freq;
tables cumexp;
run;
data count; set longbmi.level1hlm;
proc freq;
tables grade;
run;
*data longbmi.level1hlm; *set longbmi.Level1HLM;
*agesq=age*age;
proc freq;
tables grade;
run;
data longbmi.level1fl; set longbmi.level1hlm;
if grade=0 then time=0;
else if grade=1 then time=1;
else if grade=2 then time=2;
else if grade=3 then time=3;
else if grade=4 then time=4;
else if grade=5 then time=5;
else if grade=6 then time=6;
else if grade=7 then time=7;
else if grade=8 then time=8;
run;
proc freq;
tables time grade lunchbi;
run;
proc sort data=longbmi.level1fl;
by sidf;
run;
proc sort data=longbmi.level1hlm;
by SIDf;
run;
data longbmi.level1fl; set longbmi.level1hlm;
if time=0 then timesq=0;
if time=1 then timesq=1;
if time=2 then timesq=4;
if time=3 then timesq=9;
if time=4 then timesq=16;
if time=5 then timesq=25;
if time=6 then timesq=36;
if time=7 then timesq=49;
if time=8 then timesq=64;
run;
data longbmi.Level2gir; set Level2g;
APPENDIX A (continued)

run;
data longbmi.Level2boy;
set Level2b;
run;
proc contents data=longbmi.Level1HLM varnum;
title 'Level1HLM';
run;
proc contents data=longbmi.Level2HLM varnum;
title 'Level2HLM';
run;
proc sort data=longbmi.level1hlm;
by SIDf;
count+1;
run;
run;
data longbmi.level1fl; set longbmi.level1fl;
interact=time*cumexp;
intersq=timesq*cumexp;
run;
data longbmi.level1f; set longbmi.level1f;
proc freq;
tables bmipct;
run;
data longbmi.level1f; set longbmi.level1f;
run;
run;
data longbmi.level1fl; set longbmi.level1fl;
if cumexp=6 or cumexp=7;
run;
data longbmi.Level104; set longbmi.level1fl;
set Level1HLMyear00 Level1HLMyear01 Level1HLMyear02
Level1HLMyear03 Level1HLMyear04;
run;
proc sort data=longbmi.Level104;
by SIDf;
run;
run;
data longbmi.Level104 varnum;
title 'Level1FinalHLM';
run;

*Create level 3 cross classification table: school;
data Level3HLMyear00;
set finalMerge (keep= sidf school00
rename= (school00=school));
run;

data Level3HLMyear01;
set finalMerge (keep= sidf school01
rename= (school01=school));
run;

data Level3HLMyear02;
set finalMerge (keep= sidf school02
rename= (school02=school));
run;
APPENDIX A (continued)

data Level3HLMyear03;
set finalMerge (keep=sidf school03
rename= (school03=school));
run;
data Level3HLMyear04;
set finalMerge (keep=sidf school04
rename= (school04=school));
run;
data Level3HLMyear05;
set finalMerge (keep=sidf school05
rename= (school05=school));
run;
data Level3HLMyear06;
set finalMerge (keep=sidf school06
rename= (school06=school));
run;
data Level3HLMyear07;
set finalMerge (keep=sidf school07
rename= (school07=school));
run;
data Level3HLMyear08;
set finalMerge (keep=sidf school08
rename= (school08=school));
run;
data longbmi.school;
set Level3HLMyear00 Level3HLMyear01 Level3HLMyear02
Level3HLMyear03 Level3HLMyear04 Level3HLMyear05 Level3HLMyear06
Level3HLMyear07 Level3HLMyear08;
run;
proc sort data=longbmi.school;
by SIDf;
run;
proc contents data=school varnum;
title 'Level3HLM';;
run;
*Create level 3 cross classification table: neighborhood;
data neighborhood;
set finalmerge(keep= sidf zipcode);
run;
data format; set finalmerge;
*********************************************************************;
** Unmerge Data, Create Level 1 HLM Table until 2004 only ;
**
*********************************************************************;
data Level1HLMyear00;
set finalMerge (keep= year00 sidf lunchcategory00 tdate00 age00
school00 grade00 bmi00 bmiz00 bmipct00 obese00
rename= (tdate00=tdate
age00=age
school00=school
grade00=grade
bmi00=bmi
APPENDIX A (continued)

data Level1HLMyear01;
  set finalMerge (keep= year00 sidf lunchcategory00 tdate01 age01 school01 grade01 bmi01 bmiz01 bmipct01 obese01)
    rename= (tdate01=tdate age01=age school01=school grade01=grade bmi01=bmi bmiz01=bmiz bmipct01=bmipct lunchcategory01=lunchbi obese01=obese));
  if year00=0 then delete;
  drop year00; year=0;
run;
data Level1HLMyear01;
  set finalMerge (keep= year01 sidf lunchcategory01 tdate01 age01 school01 grade01 bmi01 bmiz01 bmipct01 obese01)
    rename= (tdate01=tdate age01=age school01=school grade01=grade bmi01=bmi bmiz01=bmiz bmipct01=bmipct lunchcategory01=lunchbi obese01=obese));
  if year01=0 then delete;
  drop year01; year=1;
run;
data longbmi.Level1HLM2001;
  set Level1HLMyear00 Level1HLMyear01;
  if grade=0 then time=0;
  else if grade=1 then time=1;
  else if grade=2 then time=2;
  else if grade=3 then time=3;
  else if grade=4 then time=4;
  else if grade=5 then time=5;
  else if grade=6 then time=6;
  else if grade=7 then time=7;
  else if grade=8 then time=8;
run;
data longbmi.Level1HLM2001; set longbmi.Level1HLM2001;
  if time=0 then timesq=0;
  if time=1 then timesq=1;
  if time=2 then timesq=4;
  if time=3 then timesq=9;
  if time=4 then timesq=16;
  if time=5 then timesq=25;
  if time=6 then timesq=36;
  if time=7 then timesq=49;
  if time=8 then timesq=64;
run;
proc sort data=longbmi.level1hlm2001;
  by SIDf;
run;
data longbmi.Level2HLM;
  set Level2aHLM (keep= sidf ethnic1 ethnic2 ethnic3 ethnic4 primlang sex2 zipcode);
run;
data Level1HLMyear02;
APPENDIX A (continued)

    set finalMerge (keep= year02 sidf lunchcategory02 tdate02 age02 school02 grade02 bmi02 bmiz02 bmipct02 obese02
        rename= (tdate02=tdate age02=age school02=school grade02=grade bmi02=bmi bmiz02=bmiz bmipct02=bmipct
        lunchcategory02=lunchbi obese02=obese));

    if year02=0 then delete;
    drop year02; year=2;
    run;
    data Level1HLMyear03;
        set finalMerge (keep= year03 sidf lunchcategory03 tdate03 age03 school03 Grade03f bmi03 bmiz03 bmipct03 obese03
            rename= (tdate03=tdate age03=age school03=school grade03f=grade bmi03=bmi bmiz03=bmiz bmipct03=bmipct
            lunchcategory03=lunchbi obese03=obese));

    if year03=0 then delete;
    drop year03; year=3;
    run;
    data Level1HLMyear04;
        set finalMerge (keep= year04 sidf lunchcategory04 tdate04 age04 school04 Grade04f bmi04 bmiz04 bmipct04 obese04
            rename= (tdate04=tdate age04=age school04=school grade04f=grade bmi04=bmi bmiz04=bmiz bmipct04=bmipct
            lunchcategory04=lunchbi obese04=obese));

    if year04=0 then delete;
    drop year04; year=4;
    run;
    data longbmi.Level1HLM2004;
        set Level1HLMyear00 Level1HLMyear01 Level1HLMyear02 Level1HLMyear03 Level1HLMyear04;
        run;
    data longbmi.Level1HLM2004; set longbmi.Level1HLM2004;
        if grade=0 then time=0;
        else if grade=1 then time=1;
        else if grade=2 then time=2;
        else if grade=3 then time=3;
        else if grade=4 then time=4;
else if grade=5 then time=5;
else if grade=6 then time=6;
else if grade=7 then time=7;
else if grade=8 then time=8;
run;

proc sort data=longbmi.level1fl;
by sidf;
run;

data longbmi.Level1HLM2004; set longbmi.Level1HLM2004;
if time=0 then timesq=0;
if time=1 then timesq=1;
if time=2 then timesq=4;
if time=3 then timesq=9;
if time=4 then timesq=16;
if time=5 then timesq=25;
if time=6 then timesq=36;
if time=7 then timesq=49;
if time=8 then timesq=64;
run;

proc sort data=longbmi.level1hlm2004;
by SIDf;
run;

data longbmi.Level1HLM2005;
set Level1HLMyear00 Level1HLMyear01 Level1HLMyear02
Level1HLMyear03 Level1HLMyear04 Level1HLMyear05;
run;

proc sort data=longbmi.level1hlm2005;
by SIDf;
run;

*create formats;
proc format;

    value weightf
        1="less than 5th percentile (underweight)"
        2="5th to less than 85th percentile (healthy weight)"
        3="85th to less than 95th percentile (overweight)"
        4="equal to or greater than 95th percentile (obese)"
    value ethnic
        1='White'
        2='Black or African American'
        3='Asian'
        4='Latino'
        5='other';

    value lunch
        1='Free/reduced'
        0='Paid';

    value sex
        1='male'
        0='female';

run;

*apply formats*;

data format2; set format;

    format weight_status08 weightf. ethniccategory ethnic. sexnew sex.
     lunchcategory08 lunch.;
run;
APPENDIX A (continued)

```plaintext
proc freq;
tables weight_status00 weight_status08;
run;
*data obese00; *set obese00;
*proc freq;
  *tables year00*lunchcategory00 year00*age00 sexnew*year00
  sex*year00 obese00 sexnew*year01 sexnew*year02 sexnew*year03 sexnew*year04
  sexnew*year05 sexnew*year06 sexnew*year07 sexnew*year08;
  *title 'Year 00 Weight Status Cross Tables'-----';
*run;
**********************************************************************************************
*************
*2000 Weight status prevalence, bivariate variable: obese v others;  
**********************************************************************************************

data obese00; set format2;
  if bmipct00= . then obese00= .;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
  if ethniccategory NE 5;

  if year00=1 then year=1;
  else if year01=1 then year=2;
  else if year02=1 then year=3;
  else if year03=1 then year=4;
  else if year04=1 then year=5;
  else if year05=1 then year=6;
  else if year06=1 then year=7;
  else if year07=1 then year=8;
  else if year08=1 then year=9;
  else year=0;
run;
proc freq;
tables obese00;
run;
proc logistic descending;
  class ethniccategory(ref='White')/param=ref;
  model obese00= sexnew ethniccategory lunchcategory00 age00 year01
  year02 year03 year04 year05 year06 year07 year08;
run;
proc logistic descending;
  class ethniccategory(ref='White')/param=ref;
  class numyear(ref='1')/param=ref;
  model obese00= sexnew ethniccategory lunchcategory00 age00 numyear;
run;
*Weight status prevalence, separate ethnic variables;

data obese00; set format2;
  ethnic1=(ethniccategory=1);
  ethnic2=(ethniccategory=2);
  ethnic3=(ethniccategory=3);
  ethnic4=(ethniccategory=4);
run;
proc logistic descending;
  model obese00= sexnew ethnic1 ethnic2 ethnic3 ethnic4 lunchcategory00
  age00;
```

APPENDIX A (continued)

run;
*no Asian;
proc logistic descending;
model obese00=ethnic3 ethnic1 ethnic2 ethnic4 sexnew lunchcategory00 age00;
run;
*Weight status prevalence,bivariate variable: obese v others, excluding underweight;
data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if 5<=bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
run;
proc logistic descending;
  class ethniccategory(ref='White')/param=ref;
  model obese00= sexnew ethniccategory lunchcategory00 age00;
run;
data obese00; set format2;
  *Weight status prevalence,bivariate variable: obese v others, excluding "others" race/ethnic category and "age" since not significant above;
if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory ne 5;
run;
proc logistic descending;
  class ethniccategory(ref='White')/param=ref;
  model obese00= sexnew ethniccategory lunchcategory00;
run;
proc logistic
  class ethniccategory(ref='White')/param=ref;
  model obese00= sexnew ethniccategory lunchcategory00;
run;
  *Weight status prevalence,bivariate variable: obese v others excluding "others" race/ethnic category and "age" since not significant and bivariate ethnic- white and latino/black;
data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory=1 then ethnicbi=0;
else if ethniccategory=2 or ethniccategory=4 then ethnicbi=1;
run;
proc logistic descending;
  model obese00= sexnew ethnicbi lunchcategory00;
run;
  * Modeling separate race/ethnic curves White;
data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory=1;
APPENDIX A (continued)

```plaintext
run;
proc freq;
tables obese00;
run;
proc logistic descending;
  model obese00= sexnew lunchcategory00 age00;
run;
*Black;
data obese00; set format2;
if bmipct00= . then obese00= . ;
  else if bmipct00<95 then obese00= 0 ;
  else if bmipct00>=95 then obese00= 1 ;
if ethniccategory=2;
run;
proc freq;
tables obese00;
run;
proc logistic descending;
  model obese00= sexnew lunchcategory00 age00;
run;
*Latino;
data obese00; set format2;
if bmipct00= . then obese00= . ;
  else if bmipct00<95 then obese00= 0 ;
  else if bmipct00>=95 then obese00= 1 ;
if ethniccategory=4;
run;
proc freq;
tables obese00;
run;
proc logistic descending;
  model obese00= sexnew lunchcategory00 age00;
run;
*Asian;
data obese00; set format2;
if bmipct00= . then obese00= . ;
  else if bmipct00<95 then obese00= 0 ;
  else if bmipct00>=95 then obese00= 1 ;
if ethniccategory=3;
run;
proc freq;
tables obese00;
run;
proc logistic descending;
  model obese00= sexnew lunchcategory00 age00;
run;
*Modeling separate sex curves;
*boys;
data obese00; set format2;
if bmipct00= . then obese00= . ;
  else if bmipct00<95 then obese00= 0 ;
  else if bmipct00>=95 then obese00= 1 ;
if sexnew=1;
if ethniccategory ne 5;
```

APPENDIX A (continued)

PROC LOGISTIC DESCENDING;
  CLASS ETHNICCATEGORY(REF='WHITE')/PARAM=REF;
  MODEL OBES00=ETHNICCATEGORY LUNCHCATEGORY00 AGE00;
RUN;
*WHITE BOYS;
  DATA OBES00; SET FORMAT2;
  IF BMIPCT00=. THEN OBES00=.;
  ELSE IF BMIPCT00<95 THEN OBES00=0;
  ELSE IF BMIPCT00>=95 THEN OBES00=1;
  IF ETHNICCATEGORY =1;
  IF SEXNEW=1;
  PROC LOGISTIC DESCENDING;
  MODEL OBES00=LUNCHCATEGORY00;
RUN;
*BLACK BOYS;
  DATA OBES00; SET FORMAT2;
  IF BMIPCT00=. THEN OBES00=.;
  ELSE IF BMIPCT00<95 THEN OBES00=0;
  ELSE IF BMIPCT00>=95 THEN OBES00=1;
  IF ETHNICCATEGORY =2;
  IF SEXNEW=1;
  PROC LOGISTIC DESCENDING;
  MODEL OBES00=LUNCHCATEGORY00;
RUN;
*LATINO BOYS;
  DATA OBES00; SET FORMAT2;
  IF BMIPCT00=. THEN OBES00=.;
  ELSE IF BMIPCT00<95 THEN OBES00=0;
  ELSE IF BMIPCT00>=95 THEN OBES00=1;
  IF ETHNICCATEGORY =4;
  IF SEXNEW=1;
  PROC LOGISTIC DESCENDING;
  MODEL OBES00=LUNCHCATEGORY00;
RUN;
*ASIAN BOYS;
  DATA OBES00; SET FORMAT2;
  IF BMIPCT00=. THEN OBES00=.;
  ELSE IF BMIPCT00<95 THEN OBES00=0;
  ELSE IF BMIPCT00>=95 THEN OBES00=1;
  IF ETHNICCATEGORY =3;
  IF SEXNEW=1;
  PROC LOGISTIC DESCENDING;
  MODEL OBES00=LUNCHCATEGORY00;
RUN;
*GIRLS;
  DATA OBES00; SET FORMAT2;
  IF BMIPCT00=. THEN OBES00=.;
  ELSE IF BMIPCT00<95 THEN OBES00=0;
  ELSE IF BMIPCT00>=95 THEN OBES00=1;
  IF SEXNEW=0;
  IF ETHNICCATEGORY NE 5;
  PROC LOGISTIC DESCENDING;
  CLASS ETHNICCATEGORY(REF='WHITE')/PARAM=REF;
  MODEL OBES00=ETHNICCATEGORY LUNCHCATEGORY00 AGE00;
run;
*White girls;
  data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory =1;
if sexnew=0;
  proc logistic descending;
  model obese00=lunchcategory00 age00;
run;
*Black girls;
  data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory =2;
if sexnew=0;
  proc logistic descending;
  model obese00=lunchcategory00;
run;
*Latino girls;
  data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory =4;
if sexnew=0;
  proc logistic descending;
  model obese00=lunchcategory00;
run;
*Asian girls;
  data obese00; set format2;
  if bmipct00=. then obese00=.;
  else if bmipct00<95 then obese00=0;
  else if bmipct00>=95 then obese00=1;
if ethniccategory =3;
if sexnew=0;
  proc logistic descending;
  model obese00=lunchcategory00;
run;
***********************************************************************
*************
*2001 Weight status prevalence, bivariate variable: obese v others;
***********************************************************************
*************;
  data obese01; set format2;
  if bmipct01=. then obese01=.;
  else if bmipct01<95 then obese01=0;
  else if bmipct01>=95 then obese01=1;
if ethniccategory ne 5;
run;
  proc freq;
  tables obese01;
run;
proc logistic descending;
    class ethniccategory(ref='White')/param=ref;
    model obese01= sexnew ethniccategory lunchcategory01 age01;
run;
*Weight status prevalence, bivariate variable: obese v others, excluding Asian;

data obese01; set format2;
    if ethniccategory ne 3;
    proc logistic descending;
    class ethniccategory(ref='White')/param=ref;
    model obese01= sexnew ethniccategory lunchcategory01 age01;
run;
*Modeling separate sex curves;
*boys;

data obese01; set format2;
    if bmipct01=. then obese01=.;
    else if bmipct01<95 then obese01=0;
    else if bmipct01>=95 then obese01=1;
    if sexnew=1;
    if ethniccategory ne 5;
    proc logistic descending;
    class ethniccategory(ref='White')/param=ref;
    model obese01= ethniccategory lunchcategory01 age01;
run;
*Girls;

data obese01; set format2;
    if bmipct01=. then obese01=.;
    else if bmipct01<95 then obese01=0;
    else if bmipct01>=95 then obese01=1;
    if sexnew=0;
    if ethniccategory ne 5;
    proc logistic descending;
    class ethniccategory(ref='White')/param=ref;
    model obese01= ethniccategory lunchcategory01 age01;
run;
*Weight status prevalence, bivariate variable: obese v others, excluding Asian, using separate dummy variables;

data obese01; set format2;
    ethnic1=(ethniccategory=1);
    ethnic2=(ethniccategory=2);
    ethnic3=(ethniccategory=3);
    ethnic4=(ethniccategory=4);
    proc logistic descending;
    model obese01= sexnew ethnic1 ethnic2 ethnic4 lunchcategory01 age01;
run;
* *******************************************************
*******
***********
*2002 Weight status prevalence, bivariate variable: obese v others;
***************************************************************************************
***********;

data obese02; set format2;
    if bmipct02=. then obese02=.;
APPENDIX A (continued)

else if bmipct02<95 then obese02=0;
else if bmipct02>=95 then obese02=1;

if ethniccategory ne 5;
run;
proc freq;
tables obese02;
run;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese02= sexnew ethniccategory lunchcategory02 age02;
run;
*boys;
data obese02; set format2;
if bmipct02=. then obese02=.;
else if bmipct02<95 then obese02=0;
else if bmipct02>=95 then obese02=1;
if sexnew=1;
if ethniccategory ne 5;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese02=ethniccategory lunchcategory02 age02;
run;
*Girls;
data obese02; set format2;
if sexnew=0;
if ethniccategory ne 5;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese02=ethniccategory lunchcategory02 age02;
run;

***************************************************************
******
*2003 Weight status prevalence,bivariate variable: obese v others;
***********************************************************************
*****
data obese03; set format2;
if bmipct03=. then obese03=.;
else if bmipct03<95 then obese03=0;
else if bmipct03>=95 then obese03=1;
if ethniccategory ne 5;
run;
proc freq;
tables obese03;
run;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese03= sexnew ethniccategory lunchcategory03 age03;
run;
*boys;
data obese03; set format2;
APPENDIX A (continued)

if bmipct03=. then obese03=.;
   else if bmipct03<95 then obese03=0;
   else if bmipct03>=95 then obese03=1;
if sexnew=1;
if ethniccategory ne 5;
proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obese03=ethniccategory lunchcategory03 age03;
  run;
*Girls;

data obese03; set format2;
if bmipct03=. then obese03=.;
   else if bmipct03<95 then obese03=0;
   else if bmipct03>=95 then obese03=1;
if sexnew=0;
if ethniccategory ne 5;
proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obese03=ethniccategory lunchcategory03 age03;
  run;

**********************************************************************
*************
*2004 Weight status prevalence, bivariate variable: obese v others;
**********************************************************************;

************;

data obese04; set format2;
if bmipct04=. then obese04=.;
   else if 5<=bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if ethniccategory NE 5;
run;
 proc freq;
tables obese04;
run;
 proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obese04= sexnew ethniccategory lunchcategory04 age04 year01 year02 year03 year04 year05 year06 year07 year08;
  run;
 proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obese04= sexnew ethniccategory lunchcategory04 age04 year01 year02 year03 year04 year05 year06 year07 year08;
  run;

data obese04; set format2;
*Weight status prevalence, bivariate variable: obese v other, excluding
"others" race/ethnic category and "age" since not significant above;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if ethniccategory ne 5;
run;
 proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
APPENDIX A (continued)

model obese04= sexnew ethniccategory lunchcategory04;
run;

*Weight status prevalence, bivariate variable: obese v other excluding "others" race/ethnic category and "age" since not significant and bivariate ethnic- white and latino/black;

data obese04; set format2;
  if bmipct04=. then obese04=.
  else if bmipct04<95 then obese04=0;
  else if bmipct04>=95 then obese04=1;
if ethniccategory=1 then ethnicbi=0;
else if ethniccategory=2 or ethniccategory=4 then ethnicbi=1;
run;

proc logistic descending;
  model obese04= sexnew ethnicbi lunchcategory04;
  run;

* Modeling separate race/ethnic curves
  White;

data obese04; set format2;
  if bmipct04=. then obese04=.
  else if bmipct04<95 then obese04=0;
  else if bmipct04>=95 then obese04=1;
if ethniccategory=1;
run;

proc freq;
  tables obese04;
run;

proc logistic descending;
  model obese04= sexnew lunchcategory04 age04;
  run;

*Black;

data obese04; set format2;
  if bmipct04=. then obese04=.
  else if bmipct04<95 then obese04=0;
  else if bmipct04>=95 then obese04=1;
if ethniccategory=2;
run;

proc freq;
  tables obese04;
run;

proc logistic descending;
  model obese04= sexnew lunchcategory04 age04;
  run;

*Latino;

data obese04; set format2;
  if bmipct04=. then obese04=.
  else if bmipct04<95 then obese04=0;
  else if bmipct04>=95 then obese04=1;
if ethniccategory=4;
run;

proc freq;
  tables obese04;
run;

proc logistic descending;
  model obese04= sexnew lunchcategory04 age04;
run;
APPENDIX A (continued)

run;
*Asian;
data obese04; set format2;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if ethniccategory=3;
proc freq;
tables obese04;
run;
proc logistic descending;
   model obese04= sexnew lunchcategory04 age04;
run;
   *Modeling separate sex curves;
   *boys;
data obese04; set format2;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if sexnew=1;
if ethniccategory ne 5;
proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obese04=ethniccategory lunchcategory04 age04;
run;
   *White boys;
   data obese04; set format2;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if sexnew=1;
proc logistic descending;
   model obese04=lunchcategory04 age04;
run;
   *Black boys;
   data obese04; set format2;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if ethniccategory =2;
if sexnew=1;
proc logistic descending;
   model obese04=lunchcategory04;
run;
   *Latino boys;
   data obese04; set format2;
if bmipct04=. then obese04=.;
   else if bmipct04<95 then obese04=0;
   else if bmipct04>=95 then obese04=1;
if ethniccategory =4;
if sexnew=1;
proc logistic descending;
   model obese04=lunchcategory04;
run;
*Asian boys;
  data obese04; set format2;
  if bmipct04=. then obese04=.;
    else if bmipct04<95 then obese04=0;
    else if bmipct04>=95 then obese04=1;
  if ethniccategory =3;
  if sexnew=1;
  proc logistic descending;
    model obese04=lunchcategory04;
  run;

*GIRLS;
  data obese04; set format2;
  if sexnew=0;
    if ethniccategory ne 5;
    proc logistic descending;
      class ethniccategory(ref='White')/param=ref;
      model obese04=ethniccategory lunchcategory04 age04;
    run;

  *White girls;
    data obese04; set format2;
    if sexnew=0;
      if ethniccategory =1;
      if sexnew=0;
      proc logistic descending;
        model obese04=lunchcategory04;
      run;

  *Black girls;
    data obese04; set format2;
    if sexnew=0;
      if ethniccategory =2;
      if sexnew=0;
      proc logistic descending;
        model obese04=lunchcategory04;
      run;

  *Latino girls;
    data obese04; set format2;
    if sexnew=0;
      if ethniccategory =4;
      if sexnew=1;
      proc logistic descending;
        model obese04=lunchcategory04;
      run;

  *Asian girls;
    data obese04; set format2;
APPENDIX A (continued)

if bmiPct04=. then obese04=.;
   else if bmiPct04<95 then obese04=0;
   else if bmiPct04>=95 then obese04=1;
if ethnicCategory =3;
   if sexNew=1;
   proc logistic descending;
      model obese04=lunchCategory04;
   run;

***********************************************************************
*************
*2005 Weight status prevalence,bivariate variable: obese v others;
***********************************************************************
*************;
data obese05; set format2;
   if bmiPct05=. then obese05=.;
      else if 5=bmiPct05<95 then obese05=0;
      else if bmiPct05>=95 then obese05=1;
   if ethnicCategory NE 5;
   run;
   proc freq;
      tables obese05;
   run;
   proc logistic descending;
      class ethnicCategory(ref='White')/param=ref;
      model obese05= sexNew ethnicCategory lunchCategory05 age05;
   run;

***********************************************************************
*************
*2006 Weight status prevalence,bivariate variable: obese v others;
***********************************************************************
*************;
data obese06; set format2;
   if bmiPct06=. then obese06=.;
      else if 5=bmiPct06<95 then obese06=0;
      else if bmiPct06>=95 then obese06=1;
   if ethnicCategory NE 5;
   run;
   proc freq;
      tables obese06;
   run;
   proc logistic descending;
      class ethnicCategory(ref='White')/param=ref;
      model obese06= sexNew ethnicCategory lunchCategory06 age06;
   run;

***********************************************************************
*************
*2007 Weight status prevalence,bivariate variable: obese v others;
***********************************************************************
*************;
data obese07; set format2;
   if bmiPct07=. then obese07=.;
APPENDIX A (continued)

```plaintext
else if 5<=bmipct07<95 then obese07=0;
else if bmipct07>=95 then obese07=1;
if ethniccategory NE 5;
run;
proc freq;
tables obese07;
run;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese07= sexnew ethniccategory lunchcategory07 age07;
run;
***********************************************************************
**********
*2008 Weight status prevalence, bivariate variable: obese v others;
***********************************************************************
**********;
data obese08; set format2;
if bmipct08=. then obese08=.;
   else if 5<=bmipct08<95 then obese08=0;
       else if bmipct08>=95 then obese08=1;
if ethniccategory NE 5;
if year00=1 then year=1;
else if year01=1 then year=2;
else if year02=1 then year=3;
else if year03=1 then year=4;
else if year04=1 then year=5;
else if year05=1 then year=6;
else if year06=1 then year=7;
else if year07=1 then year=8;
else if year08=1 then year=9;
else year=0;
run;
proc freq;
tables obese08 ethniccategory;
run;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
class year(ref='1')/param=ref;
model obese08= sexnew ethniccategory lunchcategory08 age08 year01
   year02 year03 year04 year05 year06 year07 year08;
run;
data obese08; set format2;
*Weight status prevalence, bivariate variable: obese v other, excluding
"others" race/ethnic category and "age" since not significant above;
if bmipct08=. then obese08=.;
   else if bmipct08<95 then obese08=0;
       else if bmipct08>=95 then obese08=1;
if ethniccategory ne 5;
run;
proc logistic descending;
class ethniccategory(ref='White')/param=ref;
model obese08= sexnew ethniccategory lunchcategory08;
run;
*Weight status prevalence, bivariate variable: obese v other excluding
```
"others" race/ethnic category and "age" since not significant and bivariate ethnic- white and latino/black;

```sas
data obese08; set format2;
  if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
  if ethniccategory=1 then ethnicbi=0;
  else if ethniccategory=2 or ethniccategory=4 then ethnicbi=1;
run;
proc logistic descending;
  model obese08= sexnew ethnicbi lunchcategory08;
  run;
* Modeling separate race/ethnic curves White;
data obese08; set format2;
  if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
  if ethniccategory=1;
  if year00=1 then year=1;
  else if year01=1 then year=2;
  else if year02=1 then year=3;
  else if year03=1 then year=4;
  else if year04=1 then year=5;
  else if year05=1 then year=6;
  else if year06=1 then year=7;
  else if year07=1 then year=8;
  else if year08=1 then year=9;
  else year=0;
run;
proc freq;
  tables obese08;
run;
proc logistic descending;
  class year(ref='1')/param=ref;
  model obese08= sexnew lunchcategory08 age08 year;
  run;
*Black;
```

```sas
data obese08; set format2;
  if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
  if ethniccategory=2;
  if year00=1 then year=1;
  else if year01=1 then year=2;
  else if year02=1 then year=3;
  else if year03=1 then year=4;
  else if year04=1 then year=5;
  else if year05=1 then year=6;
  else if year06=1 then year=7;
  else if year07=1 then year=8;
  else if year08=1 then year=9;
  else year=0;
run;
```
*Latino;*

```sas
data obese08; set format2;
if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
if ethniccategory=4;
run;
proc freq;
tables obese08;
run;
proc logistic descending;
  model obese08= sexnew lunchcategory08 age08;
run;
*Asian;*
```

```sas
data obese08; set format2;
if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
if ethniccategory=3;
run;
proc freq;
tables obese08;
run;
proc logistic descending;
  model obese08= sexnew lunchcategory08 age08;
run;
```

*Modeling separate sex curves; boys;*

```sas
data obese08; set format2;
if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
if sexnew=1;
if ethniccategory ne 5;
proc logistic descending;
  class ethniccategory(ref='White')/param=ref;
  model obese08=ethniccategory lunchcategory08 age08;
run;
*White boys;*
```

```sas
data obese08; set format2;
if bmipct08=. then obese08=.;
  else if bmipct08<95 then obese08=0;
  else if bmipct08>=95 then obese08=1;
if ethniccategory =1;
if sexnew=1;
proc logistic descending;
  model obese08=lunchcategory08 age08;
```

```sas
```
APPENDIX A (continued)

run;
*Black boys;
  data obese08; set format2;
  if bmipct08= then obese08=.;
    else if bmipct08<95 then obese08=0;
    else if bmipct08>=95 then obese08=1;
  if ethniccategory =2;
  if sexnew=1;
  proc logistic descending;
  model obese08=lunchcategory08 age08;
run;
*Latino boys;
  data obese08; set format2;
  if bmipct08= then obese08=.;
    else if bmipct08<95 then obese08=0;
    else if bmipct08>=95 then obese08=1;
  if ethniccategory =4;
  if sexnew=1;
  proc logistic descending;
  model obese08=lunchcategory08 age08;
run;
*Asian boys;
  data obese08; set format2;
  if bmipct08= then obese08=.;
    else if bmipct08<95 then obese08=0;
    else if bmipct08>=95 then obese08=1;
  if ethniccategory =3;
  if sexnew=1;
  proc logistic descending;
  model obese08=lunchcategory08 age08;
run;
*GIRLS;
  data obese08; set format2;
  if bmipct08= then obese08=.;
    else if bmipct08<95 then obese08=0;
    else if bmipct08>=95 then obese08=1;
  if sexnew=0;
  if ethniccategory ne 5;
  proc logistic descending;
    class ethniccategory(ref='White')/param=ref;
    model obese08=ethniccategory lunchcategory08 age08;
  run;
*White girls;
  data obese08; set format2;
  if bmipct08= then obese08=.;
    else if bmipct08<95 then obese08=0;
    else if bmipct08>=95 then obese08=1;
  if ethniccategory =1;
  if sexnew=0;
  proc logistic descending;
  model obese08=lunchcategory08 age08;
run;
*Black girls;
  data obese08; set format2;
if bmipct08=. then obese08=.;
else if bmipct08<95 then obese08=0;
else if bmipct08>=95 then obese08=1;
if ethniccategory =2;
if sexnew=0;
proc logistic descending;
   model obese08=lunchcategory08 age08;
run;
*Latino girls;
   data obese08; set format2;
   if bmipct08=. then obese08=.;
   else if bmipct08<95 then obese08=0;
   else if bmipct08>=95 then obese08=1;
   if ethniccategory =4;
   if sexnew=1;
   proc logistic descending;
      model obese08=lunchcategory08 age08;
run;
*Asian girls;
   data obese08; set format2;
   if bmipct08=. then obese08=.;
   else if bmipct08<95 then obese08=0;
   else if bmipct08>=95 then obese08=1;
   if ethniccategory =3;
   if sexnew=1;
   proc logistic descending;
      model obese08=lunchcategory08 age08;
run;
***********************************************************************
**********
Looking at obese and overweight combined
***********************************************************************

**********;
   data obandov00; set format2;
   if bmipct00=. then obov00=.;
   else if bmipct00<85 then obov00=0;
   else if bmipct00>=85 then obov00=1;
run;
proc logistic descending;
   class ethniccategory(ref='White')/param=ref;
   model obOV00= sexnew ethniccategory lunchcategory00 age00;
run;
***********************************************************************
**********
BMI trajectories
***********************************************************************
**********;
APPENDIX B

CAMBRIDGE PUBLIC SCHOOLS
159 THORNDIKE STREET CAMBRIDGE, MASSACHUSETTS 02141

July 1, 2008

Ms. Stephanie Shapiro Berkson
Dr. Virginia Rall Chomitz
c/o Institute for Community Health
119 Windsor Street
Cambridge, MA 02139

Dear Ms. Shapiro Berkson and Dr. Chomitz,

As a follow up to our original meeting in August 2007 and to your request to conduct a research project in the Cambridge Public Schools (CPS), please be advised that;

Your research proposal, Informing Obesity Prevention Strategies: Results from Long-term BMI Surveillance in a Multi-ethnic, Urban School System, has been approved [effective 8/07]. Included in the approval of this project, please be reminded that I have also granted your request to work with appropriate district administrators to access CPS height and weight data sets for your research, conditional upon your sending written copies of your findings to me. These findings could be very helpful to our staff.

Best wishes for a successful endeavor.

Sincerely,

Caroline L. Turk, Ed.D.
Deputy Superintendent
Teaching and Learning

The Cambridge School Department is an equal opportunity / affirmative action employer.
APPENDIX C

UNIVERSITY OF ILLINOIS
AT CHICAGO

Office for the Protection of Research Subjects (OPRS)
Office of the Vice Chancellor for Research (MC 672)
203 Administrative Office Building
1737 West Poll Street
Chicago, Illinois 60612 7227

Exemption Granted

September 5, 2008

Stephanie Shapiro Berkson
Community Health Sciences
160 E. Berkeley St., #517
Boston, MA 02218
Phone: (312) 996-9460 / Fax: (312) 355-3700

RE: Research Protocol # 2008-0712
“Informing Obesity Prevention Strategies: Results from Eight Years of BMI Surveillance
in a Multi-ethnic, Urban School System Undergoing Ecological Physical Activity and
Nutrition Intervention”

Dear Ms. Shapiro Berkson:

Your Claim of Exemption was reviewed on September 1, 2008 and it was determined that your research protocol meets the criteria for exemption as defined in the U. S. Department of Health and Human Services Regulations for the Protection of Human Subjects [(45 CFR 46.101(b)]. You may now begin your research.

Your research may be conducted at UIC and the Cambridge Public School District, utilizing existing de-identified data only.

The specific exemption category under 45 CFR 46.101(b) is:

(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

You are reminded that investigators whose research involving human subjects is determined to be exempt from the federal regulations for the protection of human subjects still have
APPENDIX C (continued)

responsibilities for the ethical conduct of the research under state law and UIC policy. Please be aware of the following UIC policies and responsibilities for investigators:

1. **Amendments** You are responsible for reporting any amendments to your research protocol that may affect the determination of the exemption and may result in your research no longer being eligible for the exemption that has been granted.

2. **Record Keeping** You are responsible for maintaining a copy all research related records in a secure location in the event future verification is necessary, at a minimum these documents include: the research protocol, the claim of exemption application, all questionnaires, survey instruments, interview questions and/or data collection instruments associated with this research protocol, recruiting or advertising materials, any consent forms or information sheets given to subjects, or any other pertinent documents.

3. **Final Report** When you have completed work on your research protocol, you should submit a final report to the Office for Protection of Research Subjects (OPRS).

Please be sure to:

- Use your research protocol number (listed above) on any documents or correspondence with the IRB concerning your research protocol.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact me at (312) 355-2908 or the OPRS office at (312) 996-1711. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Charles W. Hoehne  
Assistant Director, IRB # 2  
Office for the Protection of Research Subjects

Enclosure(s): None

cc: Bernard Turnock, Community Health Sciences, M/C 923  
Noel Chavez, Community Health Sciences, M/C 923
May 26, 2010

Stephanie Shapiro Berkson
Community Health Sciences
160 E. Berkeley St., #517
Boston, MA 02218
Phone: (312) 996-9460 / Fax: (312) 355-3700

RE: Protocol # 2008-0712

“Informing Obesity Prevention Strategies: Results from Eight Years of BMI Surveillance in a Multi-ethnic, Urban School System Undergoing Ecological Physical Activity and Nutrition Intervention”

Dear Ms. Shapiro Berkson:

The OPRS staff/members of Institutional Review Board (IRB) #2 have reviewed this amendment to your research, and have determined that your research protocol continues to meet the criteria for exemption as defined in the U. S. Department of Health and Human Services Regulations for the Protection of Human Subjects [(45 CFR 46.101(b)].

The specific exemption category under 45 CFR 46.101(b) continues to be:
(4) Research involving the collection or study of existing data, documents, records, pathological specimens, or diagnostic specimens, if these sources are publicly available or if the information is recorded by the investigator in such a manner that subjects cannot be identified, directly or through identifiers linked to the subjects.

You may now implement the amendment in your research.

Please note the following information about your approved amendment:

**Exemption Period:** May 26, 2010 – May 25, 2013

**Amendment Approval Date:** May 26, 2010

**Amendment:**
APPENDIX C (continued)

Summary: UIC Amendment #1 dated April 28, 2010 and submitted to OPRS on May 5, 2010 is an investigator-initiated amendment adding the following aim: To assess any longitudinal dose-response trends in the incidence, prevalence rates of at risk for overweight and overweight status among cohorts of socio-demographically diverse elementary school students exposed to system-wide ecological nutrition and physical activity intervention with three incremental levels of intensity.

You are reminded that investigators whose research involving human subjects is determined to be exempt from the federal regulations for the protection of human subjects still have responsibilities for the ethical conduct of the research under state law and UIC policy. Please be aware of the following UIC policies and responsibilities for investigators:

1. **Amendments** You are responsible for reporting any amendments to your research protocol that may affect the determination of the exemption and may result in your research no longer being eligible for the exemption that has been granted.

2. **Record Keeping** You are responsible for maintaining a copy all research related records in a secure location in the event future verification is necessary, at a minimum these documents include: the research protocol, the claim of exemption application, all questionnaires, survey instruments, interview questions and/or data collection instruments associated with this research protocol, recruiting or advertising materials, any consent forms or information sheets given to subjects, or any other pertinent documents.

3. **Final Report** When you have completed work on your research protocol, you should submit a final report to the Office for Protection of Research Subjects (OPRS).

Please be sure to:

→ Use your research protocol number (2008-0712) on any documents or correspondence with the IRB concerning your research protocol.

We wish you the best as you conduct your research. If you have any questions or need further help, please contact me at (312) 355-2908 or the OPRS office at (312) 996-1711. Please send any correspondence about this protocol to OPRS at 203 AOB, M/C 672.

Sincerely,

Charles W. Hoehne, CIP
Assistant Director, IRB # 2
Office for the Protection of Research Subjects

Enclosure(s): None

cc: Bernard Turnock, Community Health Sciences, M/C 923
    Noel Chavez, Community Health Sciences, M/C 923
NAME: Stephanie Shapiro Berkson

EDUCATION:

2002-2012  University of Illinois at Chicago, Chicago, IL
            School of Public Health
            Ph.D. candidate, Community Health Sciences
            Graduation: spring, 2012
            Dissertation title: Longitudinal School-based BMI Surveillance: Informing Obesity
            Prevention Strategies

1999-2001  University of Illinois at Chicago, Chicago, IL
            School of Public Health
            MPH, Community Health Sciences and specialization in Health Education and Promotion

1993-1995  Boston University, Boston, MA
            Post-baccalaureate Pre-medical Program

1989-1993  Brandeis University, Waltham, MA
            B.A. in Psychology and Spanish Language

1992  University of Madrid, Complutence, Madrid, Spain
            Junior semester abroad

RESEARCH, TEACHING, AND WORK EXPERIENCE:

Department of Community Health Sciences  January, 2009–present
University of Illinois at Chicago, Chicago, IL
            Lecturer: Design, plan, and teach the online section of Community Health Sciences Public Health
            Planning and Evaluation, CHSC 433.

Institute for Community Health  March, 2006–July, 2011
Cambridge Health Alliance, Boston, MA
            Research Associate: Evaluated healthy weight community and school-based initiatives using a community
            participatory approach. Wrote grants, prepared annual grant reports, and managed student interns.

Department of Epidemiology/Biostatistics  January, 2001–June, 2005
University of Illinois at Chicago, Chicago, IL
            Public Education Coordinator: Created, implemented, managed, and evaluated the Chicago Project for
            Violence Prevention’s CeaseFire social marketing campaign. Created and managed CeaseFire website and
            toll-free number (600 calls/year), wrote and distributed an electronic newsletter, and pitched CeaseFire to
            the media. Directed the Chicago Public Schools Service-Learning Program component, coordinated the
            volunteer program, and planned CeaseFire’s special events.

Safe Haven Program  May, 2000–March, 2001
Northwest Austin Council, Chicago, IL
            Research Assistant: Conducted the Safe Haven program evaluation by interviewing the children involved
            with the program. Coded the children’s responses.
Department of Neonatology, University of Chicago Hospitals, Chicago, IL
March, 2000–June, 2000
Research Assistant: Collected obstetric level and network affiliation data from all of the hospitals in Illinois with obstetric services, designed and entered all data into Excel database.

Department of Cardiology, Children’s Memorial Hospital, Chicago, IL
June, 1999–March, 2000
Telephone counselor: Counseled adolescents in their efforts to stop smoking as part of a pilot smoking cessation research study. Presented and summarized cases at weekly research meetings.

Department of Pulmonary Medicine, University of Chicago, Chicago, IL
Clinical Research Coordinator: Directed the design and implementation of two multi-center clinical research projects studying the effectiveness of high frequency chest wall compressions in ventilated patients and the effect of anabolic steroids in critically ill patients. Recruited subjects, collected data, and maintained the databases.

Department of Pulmonary Medicine, University of Chicago, Chicago, IL
Research Technologist: Performed laboratory research related to the molecular effects of asthma on mammalian and human cells.

Interpreter Services, Massachusetts General Hospital, Boston, MA
January, 1995–August, 1995
On-Call Interpreter: Responsible for interpreting Spanish, off-hours, for any Spanish-speaking patient with a language barrier. Participated in establishing this on-call, off-hour interpreter service.

Pediatric Hematology/Oncology Clinic, Massachusetts General Hospital, Boston, MA
October, 1994–August, 1995
Unit Assistant: Assisted pediatric oncologists, nurses, and administrative assistants in their daily activities.

AWARDS AND HONORS:

2002 Gads Hill Center Board Member of the Year Award
2001 Paul Q. Peterson Scholarship Award
2000 Albert Schweitzer Fellowship
1999 Gads Hill Center Outstanding ESL Tutor of the Year Award
1997 Gads Hill Center Volunteer Commitment Award
1995 Massachusetts General Hospital Over 200 Volunteer Hour Award
1992 Brandeis University Varsity Swim Team Captain

INTERNSHIPS:

Health Promoter Program Fall semester, 2003
Centro San Bonifacio, Chicago, IL
Intern: Co-wrote a health promoter training guide. Covered various health related topics such as heart disease, asthma, A.I.D.S., and car safety.

Community Health Assessment Fall semester, 2000
Gads Hill Center, Chicago, IL
Intern: Conducted preliminary research and developed a community assessment focusing on nutritional needs of the Pilsen community.
The Center for Childhood Safety  
Children’s Memorial Hospital, Chicago, IL  
Intern: Conducted evaluation studies for three car seat safety programs as well as developed an SPSS database.

FELLOWSHIP:

The Albert Schweitzer Fellowship  
Project: Lights, Camera, A.C.T.I.O.N. Program  
Site: Gads Hill Center, Chicago, IL  
Program Creator/Director: Designed and implemented a violence prevention program for children in grades five through nine consisting of a twenty-session intervention followed by the development and production of four public service announcements (PSAs) regarding violence prevention messages. One of these PSAs was nationally and locally aired. Wrote program curriculum manual.

TEACHING ASSISTANTSHIPS:

Behavioral Sciences in Public Health, CHSC 401  
University of Illinois at Chicago, Chicago, IL  
Teaching Assistant: Prepared and conducted lectures focusing on the social determinants of health, maintained Blackboard website, answered student questions, and graded assignments.

Public Health Planning and Evaluation, CHSC 433  
University of Illinois at Chicago, Chicago, IL  
Teaching Assistant: Wrote chapter quizzes for professor’s textbook, maintained Blackboard website, facilitated course lectures, answered student questions, and graded assignments.

Public Health Concepts and Practice, CHSC 400  
University of Illinois at Chicago, Chicago, IL  
Teaching Assistant: Facilitated course modules, answered student questions, and graded all assignments.

PEER-REVIEWED PUBLICATIONS:


MANUSCRIPTS IN PREPERATION:


PEER REVIEWED PRESENTATIONS:


Ndukwu, Ikeadi M., Shapiro, S., Nam, A.J., and Schumm, P. Comparison of High-Frequency Chest Wall Oscillation (HFCWO) and Manual Chest Physiotherapy (mCPT) in Long-Term Acute Care Hospital (LTAC) Ventilator-Dependent Patients. Presentation accepted to the International American Thoracic Society Conference, 1999.


PROFESSIONAL ORGANIZATION MEMBERSHIP:

<table>
<thead>
<tr>
<th>Organization</th>
<th>Membership</th>
</tr>
</thead>
<tbody>
<tr>
<td>American Association of Public Health</td>
<td>1999–present</td>
</tr>
<tr>
<td>Food and Environment Working Group</td>
<td>2006–present</td>
</tr>
<tr>
<td>American Evaluation Association (Listserv)</td>
<td>2006–present</td>
</tr>
<tr>
<td>Social Marketing Listserv</td>
<td>2002–present</td>
</tr>
</tbody>
</table>

COMMUNITY SERVICE:

<table>
<thead>
<tr>
<th>Role</th>
<th>Organization</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leader</td>
<td>Ohel Tsedik and Greater Boston Interfaith Community (GBIO), Temple Israel</td>
<td>9/10–present</td>
</tr>
<tr>
<td>Advisory Board Member</td>
<td>Gads Hill Center</td>
<td>7/05–present</td>
</tr>
<tr>
<td>Big Sister</td>
<td>Waltham Group Big Sister/Big Brother Program, Brandeis University</td>
<td>2/91–present</td>
</tr>
<tr>
<td>Executive Board Member</td>
<td>Gads Hill Center</td>
<td>10/00–7/05</td>
</tr>
<tr>
<td>Board Member</td>
<td>South Loop Dog Park Action Committee (P.A.C.)</td>
<td>6/03–6/05</td>
</tr>
<tr>
<td>Travel Club</td>
<td>Gads Hill Center</td>
<td>2/02–present</td>
</tr>
<tr>
<td>English as a Second Language Instructor</td>
<td>Gads Hill Center</td>
<td>9/97–12/00</td>
</tr>
<tr>
<td>Tutor</td>
<td>Gads Hill Community Center</td>
<td>9/96–6/99</td>
</tr>
<tr>
<td>Interpreter</td>
<td>University of Chicago Hospitals, Thomas Jefferson University Hospital, Massachusetts General Hospital</td>
<td>12/92–8/95</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>