The Effects of Health-Related Fitness on School Attendance in New York City 6th-8th Grade Youth

Emily M. D'Agostino
The Graduate Center, City University of New York

Recommended Citation
D'Agostino, Emily M., "The Effects of Health-Related Fitness on School Attendance in New York City 6th-8th Grade Youth" (2016). CUNY Academic Works. https://academicworks.cuny.edu/gc_etds/1561

How does access to this work benefit you? Let us know!
Follow this and additional works at: https://academicworks.cuny.edu/gc_etds

Part of the Community Health and Preventive Medicine Commons, Epidemiology Commons, Health and Physical Education Commons, and the Sports Sciences Commons

This Dissertation is brought to you by CUNY Academic Works. It has been accepted for inclusion in All Dissertations, Theses, and Capstone Projects by an authorized administrator of CUNY Academic Works. For more information, please contact deposit@gc.cuny.edu.
The effects of health-related fitness on school attendance in New York City 6th-8th grade youth

by

Emily M. D’Agostino

A dissertation submitted to the Graduate Faculty in Public Health in partial fulfillment of the requirements for the degree of Doctor of Public Health, The City University of New York

2016
This manuscript has been read and accepted for the Graduate Faculty in Public Health in satisfaction of the dissertation requirement for the degree of Doctor of Public Health.

Denis Nash

Date

Chair of Examining Committee

Denis Nash

Date

Executive Officer

Katarzyna Wyka

Lorna Thorpe

Joseph Egger

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK
Abstract
The effects of health-related fitness on school attendance in New York City 6th-8th grade youth

By
Emily M. D’Agostino

Advisor: Katarzyna Wyka

Background: Only 42% of youth ages 6-11 in the United States meet the World Health Organization’s recommendation for ≥60 minutes of daily moderate to vigorous physical activity. Estimates for adolescents ages 12-19 are even lower, ranging from 8-17%. Literature suggests low levels of youth health-related fitness (fitness) may negatively impact attendance, potentially due to reduced physical and psychosocial wellness.

Nationally, 10-15% of (5-7.5 million) students are chronically absent, meaning that they miss ≥10% of the school year (or ≥20 days of school per year). Moreover, 20-30% of students in high-poverty, urban school districts do not attend school regularly (≥6 days absent per year). To the author’s knowledge, this is the first multi-year study on the longitudinal fitness-attendance relationship, taking into account multilevel student and school factors.

Objectives: To describe differences in student- and school-level attendance, examine the longitudinal causal effects of change in fitness on attendance (days absent and chronic absenteeism), and assess gender effect measure modification in the fitness-attendance relationship in a large and diverse sample of New York City (NYC) middle school students.
Methods: Data were drawn from the NYC Fitnessgram. Six cohorts of NYC public middle school students were followed over 4 years during a seven-year study period (2006/7-2012/13; n=349,381). A 2-level cross-sectional Generalized Linear Mixed Model (GLMM) was used to assess the extent of school-level clustering in attendance, and determine the variation in student attendance accounted for by school-area poverty. A 3-level longitudinal adjusted GLMM, overall and stratified by gender, was used to test the effects of change in fitness (cardiorespiratory, muscular strength and endurance composite score) on attendance (1-year lagged student-specific days absent). Lastly, a 3-level logistic GLMM was used to test the effects of change in fitness on chronic absenteeism (1-year lagged student-specific chronic absenteeism, defined as ≥20 days absent per year).

Results: The Intraclass Correlation Coefficient (ICC) estimate demonstrated a large degree of variance in student attendance explained by schools (11%; p<.001). The percent of between-school variance in student attendance accounted for by school-area poverty was 20% (p<.001). There was a general trend of increasing attendance with increasing fitness across grade levels. Findings from 3-level longitudinal GLMM showed that girls with a large increase in fitness (>20% increase) demonstrated 0.66 fewer days absent per year (95%CI: 0.56, 0.75; p<.001) compared with boys who demonstrated 0.64 fewer days absent per year (95%CI: 0.55, 0.72; p<.005) relative to the reference group (>20% decrease in fitness based on change in composite percentile scores from the year prior). The difference in days absent for the most compared with
least improved fitness groups was greatest in both genders for those attending schools in high- and very high- compared with mid- and low-poverty areas (0.567 vs. 0.249, and 0.522 vs. 0.275 in girls and boys, respectively). Lastly, 3-level logistic GLMM showed that a large increase in fitness (>20%) was associated with a 19% lower probability of chronic absenteeism the following year (95%CI: 17.82, 20.69), relative to students with a >20% decrease in fitness composite percentile scores (p<.001).

**Conclusion:** School-level clustering in attendance was sizeable and statistically significant. School-area poverty accounted for a large proportion of the variance in student attendance at the school-level. An inverse dose-response relationship was found between student change in fitness level and days absent in both genders, with slightly stronger effects in girls, and youth attending schools in high-poverty areas. Further, an inverse dose-response relationship was found between student fitness and 1-year lagged school chronic absenteeism. Study findings suggest cumulative effects of fitness improvement could have a significant impact on child attendance over time. Given over 200,000 NYC students are chronically absent each year, this work suggests that fitness interventions should be examined as a method to promote attendance at the population level.
Acknowledgments

First and foremost, I must thank my husband, Steven D’Agostino, for helping me endure this extremely challenging and equally rewarding journey. He supported me from the very beginning, when I took time off from medical school to determine my true path, all the way through to my defense. I would not have been able to succeed without him. I also would like to thank my children, Mia, Cassandra and Vincent. They have been exceptionally patient, and at times quite engaged in the world of epidemiology. I hope I can continue to inspire them to question how things work. I also must thank my parents, Richard and Susan Bronson, who were consistently supportive of me throughout this doctoral program, and became particularly essential over the last year as they hosted me on countless trips back and forth between Miami and New York City while I completed my dissertation. It was a special gift each visit to have the opportunity for late-night chats and a home-cooked meal after 12 hour days of data analysis and writing. In addition, I would like to thank my committee members, Denis Nash, Lorna Thorpe, Joseph Egger, and in particular, Katarzyna Wyka. Dr. Wyka provided countless hours of support and guidance, and was instrumental in my successful defense. I also would like to thank my mentors, colleagues and friends in the CUNY Doctor of Public Health program, including in particular Sharon Joseph, Catherine Diamond, Charles Platkin, Neal Cohen, Shiro Horiuchi, Betty Levin, Grace Sembajwe, Heidi Jones, Susan Klitzman, and Christina Zarcadoolas, who provided me with confidence and highly treasured advice along the way. Lastly, and perhaps most importantly, I must thank Sophia Day and Kevin Konty, the Fitnessgram Team at NYC Department of Health and Mental Hygiene, who were generous in providing me access to this incredibly valuable dataset, and equally generous in their time and support over the last two years. I hope to have future opportunities to collaborate with them, and continue to learn from them.
Table of Contents

List of Tables ........................................................................................................... xiii

List of Figures ......................................................................................................... xv

Chapter 1: Introduction ......................................................................................... 1

1.1. Youth physical activity in the United States ............................................ 1
1.2. Characterizing health-related fitness ....................................................... 1
1.3 Youth health-related fitness and school attendance .............................. 2
  1.3.1. School-level differences in the fitness-attendance association ........ 5
  1.3.2. Gender modification of the fitness-attendance association ........... 5
  1.3.3. Health-related fitness and chronic absenteeism .......................... 6
1.4. Summary and gaps in current literature .................................................... 7
1.5 Overview of the dissertation ..................................................................... 8
  1.5.1. Overall Goals .................................................................................... 8
  1.5.2. Specific Aims .................................................................................... 8
  1.5.3. Organization of the dissertation .................................................... 10
  1.5.4. Significance of the dissertation .................................................... 11
1.6. Data sources and study population ......................................................... 11
  1.6.1. Primary exposure .......................................................................... 11
  1.6.2. Primary outcomes .......................................................................... 12
  1.6.3. Individual-level variables .............................................................. 12
  1.6.4. School-level variables .................................................................. 14

Chapter 2: Individual- and school-level differences in the fitness-attendance
association in New York City 6-8th grade youth ............................................. 15

2.1. Background ............................................................................................... 15
2.1.1. Physical activity and school outcomes........................................15
2.1.2. Effects of individual and school-level factors on the fitness-attendance association.................................................................16
2.1.3. Reporting attendance at the individual- and school-levels.........16
2.2. Methods....................................................................................17
2.2.1. Data source, collection, management and study population.......17
2.2.2. Primary exposure........................................................................20
2.2.3 Primary outcome..........................................................................20
2.2.4. Individual-level variables............................................................23
2.2.5. School-level variables.................................................................24
2.2.6. Statistical methods.....................................................................24
2.3 Results.........................................................................................26
2.3.1. Sample characteristics...............................................................26
2.3.2. Descriptive trends in attendance at the individual- and school-level.........................................................................................28
2.3.3. Descriptive trends in attendance by fitness, grade, and school-area poverty...........................................................................32
2.3.4. Variation in student attendance accounted for by schools.......35
2.4. Discussion...................................................................................36
2.4.1. Limitations................................................................................39
2.4.2. Conclusion................................................................................41

Chapter 3: Investigating the causal longitudinal effects of fitness and lagged attendance in New York City middle-school youth.................................42
3.1. Background ...........................................................................................................42
  3.1.1. Longitudinal data to study the fitness-attendance relationship……43
  3.1.2. Gender differences in the fitness-attendance association………43

3.2. Methods ...............................................................................................................45
  3.2.1. Data source, collection, management and study population……45
  3.2.2. Primary exposure .......................................................................................45
  3.2.3. Primary outcome .......................................................................................46
  3.2.4. Individual-level variables .........................................................................46
  3.2.5. School-level variables ..............................................................................47
  3.2.6. Statistical methods ...................................................................................48
  3.2.7. Regression diagnostics and sensitivity analyses.................................50

3.3. Results ...............................................................................................................51
  3.3.1. Sample characteristics ............................................................................51
  3.3.2. Attendance by fitness, grade and gender ..............................................52
  3.3.3. Longitudinal individual-level and school-level clustering of attendance......................................................................................................................56
  3.3.4. Longitudinal causal effects of fitness-change on attendance.............56
  3.3.5. Longitudinal causal effects of fitness-change on attendance by gender.................................................................................................................................57
  3.3.6. Regression diagnostics and sensitivity analysis results .......................58

3.4. Discussion .........................................................................................................61
  3.4.1. Limitations ..............................................................................................64
  3.4.2. Conclusion ..............................................................................................66
Chapter 4: Examining the causal longitudinal effects of fitness and chronic absenteeism in New York City 6th-8th grade youth

4.1. Background

4.1.1. Fitness and attendance in youth

4.2. Methods

4.2.1. Data source, collection, management and study population

4.2.2. Primary exposure

4.2.3. Primary outcome

4.2.4. Individual-level variables

4.2.5. School-level variables

4.2.6. Statistical methods

4.3. Results

4.3.1. Sample characteristics

4.3.3. Chronic absenteeism prevalence rates by demographics and fitness

4.3.4. Cross-sectional fitness-change - chronic absenteeism association by grade

4.3.5. Longitudinal causal effects of fitness-change on chronic absenteeism

4.4. Discussion

4.5.1. Limitations

4.5.2. Conclusion

Chapter 5: Discussion
5.1. Overview of the Dissertation ...............................................................89
5.2. Summary of Findings ........................................................................89
  5.2.1. Chapter 2 ......................................................................................89
  5.2.2. Chapter 3 ......................................................................................91
  5.2.3. Chapter 4 ......................................................................................93
  5.2.4. Overall findings ............................................................................95
5.3. Limitations ..........................................................................................96
5.4. Strengths and public health significance ...........................................98
  5.4.1. Public health significance: Prevention paradox and school attendance ..................................................................................98
5.5. Policy recommendations and future research directions .................99
  5.5.1. Conclusion ....................................................................................101

Appendix ..................................................................................................102

References ...............................................................................................112
List of Tables

Table 2.1. Demographic and fitness-change characteristics of the study population….27

Table 2.2. Days absent per year across student- and school-level demographic and fitness-change characteristics……………………………………………………………….....30

Table 2.3. Mean attendance for New York City public school students in grades 6-8, by fitness-change from the year prior……………………………………………………………………33

Table 2.4. Mean attendance for New York City public school students across school-area poverty (SAP) by level of fitness-change from the previous year…………………………33

Table 2.5. Variability in attendance explained by clustering students in schools………35

Table 3.1. Demographic and fitness-change characteristics of the study population…….51

Table 3.2. Mean in attendance for New York City public school students in grades 6-8, by level of fitness-change from the previous year across gender…………………………54

Table 3.3. Between- and within-school level variances in attendance in empty models and with fitness as the predictor………………………………………………………………55

Table 3.4. Longitudinal causal effects of fitness-change and attendance in New York City public school students in grades 6-8……………………………………………………….57

Table 3.5. Longitudinal causal effects of fitness-change on attendance by gender in New York City public school students in grades 6-8………………………………………………58

Table 4.1. Demographic and fitness-change characteristics of the study population…….76
Table 4.2. Chronic absenteeism overall and across demographic and fitness-change characteristics……………………………………………………………………………………………………..78

Table 4.3. Predicted probability of chronic absenteeism in New York City public school students grades 6-8, by level of fitness-change from the previous year……………81

Table 4.4. Overall effects of fitness-change on predicted probability of chronic absenteeism in New York City public school students in grades 6-8…………………81
List of Figures

Figure 2.1. Sample selection flow chart .................................................................19

Figure 2.2. Days absent across fitness-change categories .................................29

Figure 2.3. Mean days absent by grade across fitness-change categories ..........34

Figure 2.4. Mean days absent by school-area poverty across fitness-change categories .34

Figure 3.1. Mean days absent by grade across fitness-change categories in girls ....53

Figure 3.2. Mean days absent by grade across fitness-change categories in boys .......53

Figure 4.1. Predicted probability of chronic absenteeism by fitness-change and grade ...80
Chapter 1: Introduction

1.1. Youth physical activity in the United States

Both the National Association for Sport and Physical Education (NASPE) and the World Health Organization (WHO) recommend children have ≥60 minutes of moderate to vigorous physical activity every day (MVPA) to promote health-related fitness (fitness).\(^1,2\) However, only 42% of children in the United States (US) ages 6-11 years meet these recommendations, and estimates for adolescents range from 8-17%.\(^3,4\) These levels are far lower than Western European countries, where 97% and 62-82% of 9 and 15 year olds, respectively, are estimated to meet international physical activity recommendations.\(^5\) Moreover, physical activity levels have decreased over the last several decades,\(^6,7\) with steeper declines from childhood to adolescence in the US compared with other nations.\(^8\) Low levels of youth physical activity in the US are of particular concern given extensive research demonstrating the benefits of physical activity on children’s health\(^3,9\) and academic performance,\(^3,10-12\) potentially acting through pathways involving enhanced cognition and memory,\(^12-18\) or improvements in both physical and psychosocial wellness.\(^19-27\) In this sense, attendance may mediate the improved physical and psychosocial wellness-academics hypothesized pathway, although this has not been formally tested.

1.2. Characterizing health-related fitness

Health-related fitness is defined as, “A state of being that reflects a person’s ability to perform specific forms of physical activity/exercise or functions, and is related to present
and future health outcomes.\textsuperscript{28(p.3-2)} This construct was first introduced by the Task Force on Youth Fitness in 1977 in response to youth fitness assessments promoted by the American Alliance for Health, Physical Education, Recreation and Dance (AAHPER/AAHPERD), which related physical fitness to sport performance rather than functional health. Fitness is also the basis for the Fitnessgram, the most widely used criterion-referenced fitness assessment, developed by the Cooper Institute for Aerobics Research (CIAR/IAR) and employed globally in school and research settings.\textsuperscript{28-30} The five components of fitness include cardiorespiratory endurance, muscular endurance, muscular strength, body composition, and flexibility.\textsuperscript{31} Body composition is typically measured using BMI as a proxy. While BMI is a good predictor of health-related fitness,\textsuperscript{2,29,31} it cannot be characterized as a fitness test, or physical ability, in and of itself. In contrast, cardiorespiratory endurance, muscular endurance, muscular strength, and flexibility are assessed using physical ability tests, such as timed runs, and other cardiorespiratory, strength, isokinetic, and flexibility tests.\textsuperscript{31} In this sense, body composition versus other fitness components should be treated as conceptually distinct.

### 1.3. Youth health-related fitness and school attendance

Recent literature has suggested fitness may increase attendance,\textsuperscript{32-34} similar to findings on the association of fitness and workplace attendance in adult populations.\textsuperscript{35-41} For example, cardiorespiratory fitness and physical activity in adults are shown to be positively associated with workplace attendance.\textsuperscript{37,38,42} Moreover, interventions targeting improvements in adult fitness have demonstrated an increase in workplace attendance.\textsuperscript{35,36,41,43} Causal mechanisms may include reduced risk of cardiovascular
disease, insulin sensitivity, hypertension, and metabolic syndrome, perhaps contributing to reduced absence related to illness.\textsuperscript{37} Research on the association of global health and attendance in youth populations has shown similar findings.\textsuperscript{44-47} Likewise, improvements in diet and physical activity in children may reduce negative health effects and psychosocial problems associated with overweight and obesity.\textsuperscript{39} Given these findings, it is plausible that improvements in fitness may predict increased attendance in youth.

A large body of research demonstrates a positive or neutral association between youth fitness and educational outcomes,\textsuperscript{3,10-12,19,22,23,32-34,48-54} and further that children’s academic performance is not negatively impacted when instructional time is reassigned to aerobic physical activity programs. For example, in the review performed by the Center for Disease Control’s (CDC) from 2010 on 50 studies examining the association between school-based physical activity and several indicators of academic performance (course grades, academic behavior, cognitive skills and attitudes) found positive associations in half of the studies reviewed, nonsignificant associations in 48% and negative associations in 1.5% (effect sizes not shown).\textsuperscript{3} The authors concluded physical activity may promote, and does not appear to adversely affect academic achievement. Similarly, a meta-analysis of 59 studies from 1947-2009 concluded that school-based physical activity programs in children have a significant positive association with cognitive performance based on school grades or standardized test performance, with a greater effect size for experimental and quasi-experimental versus observational designs (mean effect size=0.35, 95\%CI: 0.27-0.43 vs. 0.32, 95\%CI: 0.26-0.37, for experimental and observational studies, respectively).\textsuperscript{12}
To the author’s knowledge, only 5 studies have examined the relationship between fitness and attendance among youth, and their findings are consistent in demonstrating a positive association. 19,22,32-34 Mohar, in a cross-sectional study of primary and middle school children (n=324) demonstrated a significant inverse association between MVPA (based on accelerometry) and mean days absent (M=6.99 (SD=0.42), M=3.90 (SD=2.50) and M=3.34 (SD=.25) days per year for the lowest, middle and highest physical activity tertiles, respectively). 34 Likewise, Shannonhouse in a non-randomized controlled trial found that a game-based physical activity intervention (n=96) increased attendance (M=8.82 (SD=6.78) vs. M=12.03 (SD=7.86) days absent per year, for experimental vs. control groups, respectively; p=.056.). 33,55 Kristjánsson also found a significant positive association between individual student physical activity and attendance (β±SE=— .15±0.0.024), where self-reported frequency of physical activity (1 = almost never, 2 = less than once a week, 3 = once a week, 4 = 2-3 times a week, 5 = 4-5 times a week, and 6 = almost every day) reduced the frequency of self-reported skipped classes by .15 on a scale of 1-5 (response format: 1 = almost never, 2 = less than once a month, 3 = every month, 4 = every week, and 5 = almost daily). 19 Moreover, Welk et al. in a large cross-sectional study found moderate positive correlations between cardiovascular health-related fitness and attendance (r=.38) measured at the school level. 32 Lastly, Blom et al.’s cross-sectional study demonstrated students with greater fitness had a lower odds of ≥8 absences per year (OR=3.31, 95%CI:1.51-7.28, for students with 6 vs. ≤5 Healthy Fitness Zones achieved, p=.0093). 22 Although mounting evidence show a positive link between fitness and attendance, only 1 study on the fitness-attendance relationship drew from longitudinal data, 33 and no studies examined effects over multiple years. Given that
temporality of exposure and outcome are not known in cross-sectional studies, longitudinal data spanning multiple years can more accurately assess a causal fitness-attendance association. In sum, analyses based on longitudinal research are needed to address the potential causal pathway between fitness and attendance.

1.3.1. School-level differences in the fitness-attendance association

Little is also known about the impact of school-level factors on the fitness-attendance relationship. School contextual factors, such as area poverty and the built environment are shown in the literature to be associated with children’s tendency to participate in school- and neighborhood-based physical activity. Neighborhood factors, for example, may contribute to opportunities for safe, attractive, and accessible physical activity. Likewise, school contextual effects may impact attendance. Indeed, community norms and attitudes may inform parental decisions to permit children’s school absences for family or work obligations. Similarly, perceptions of neighborhood safety may influence attendance. However, only 1 paper was identified on the fitness-academics association that adjusted for school-area poverty. No studies were identified on the association of fitness and attendance considering contextual factors as potential antecedents or confounders. In this sense, research is warranted on the effects of school contextual factors on the fitness-attendance relationship.

1.3.2. Gender modification of the fitness-attendance association

Numerous studies demonstrate low self-esteem in adolescent girls is significantly associated with both lower physical activity levels and attendance, attributed in part
to perceived weight status and self-appearance. Gender may serve as an effect measure modifier (EMM) in the association of fitness and attendance, perhaps with self-esteem acting as a more important antecedent for girls than boys. In fact, some literature applied Social Cognitive Theory (SCT) to support the hypothesis that psychosocial factors in girls serve as antecedents to fitness, such as by precluding students’ tendency to participate in physical activity. In this sense, gender may act as an EMM in the fitness-attendance relationship due to differing influences of esteem or other factors on the relationship.

To the author’s knowledge, no studies have specifically addressed gender as an EMM in the fitness-attendance relationship. However, six studies examined whether gender modifies the fitness-academics association, four of which found stronger effects for females. For example, Bezold found girls with a substantial increase in fitness relative to peers (0.36 percentile points per year vs. the reference group) showed the largest increase in academic ranking (1.06 percentile points per year). One study found no significant differences by gender, and 1 study found stronger effects for boys, although the study sample was younger (elementary school-aged children).

1.3.3. Health-related fitness and chronic absenteeism

Chronic absenteeism rates remain high. Nationally, 10-15% of (5-7.5 million) students are chronically absent, meaning they miss ≥10% of the school year (or ≥20 days of school per year). Chronic absenteeism rates increase with increasing student age, and are strongly associated with student race/ethnicity, and socioeconomic status.
Moreover, chronic absenteeism is shown to reduce academic performance, and has long-term effects on graduation rates. Reducing chronic absenteeism may also diminish racial/ethnic disparities in academic achievement. For example, Musser et al. found that moving from chronic absenteeism to average attendance was associated with a 17% and 26% decrease in the achievement gap between non-Hispanic white and minority 4th grade students on English and math standardized tests, respectively. Given findings demonstrating a positive association between fitness and attendance, it is plausible that children’s fitness may similarly predict reduced chronic absenteeism. To the author’s knowledge, no studies have addressed the causal effects of fitness on chronic absenteeism in youth.

1.4. Summary and gaps in current literature

The literature to date hypothesizes fitness improves cognition and/or psychosocial health in youth, which may promote attendance. Most research demonstrates a positive association between fitness and attendance. However, the bulk of research on fitness and attendance comprises cross-sectional studies, which are unable to support causal hypotheses given temporality of exposure and outcome are not known. Likewise, unmeasured confounding is probable in the majority of studies reviewed above, which mainly adjusted only for gender, individual SES, age, and race/ethnicity. Indeed, these studies did not account for additional potential confounders, including psychosocial factors, drug and alcohol use, family structure, and the school and neighborhood built environment, particularly in light of research suggesting school-level factors play a large role. Also, none of these studies looked at individual student data nested in schools or
neighborhoods, making it difficult to disentangle confounding due to individual compared with area-level factors. Finally, no studies have formally examined gender as a modifier in the fitness-attendance association, or the longitudinal, causal effects of fitness on chronic absenteeism, taking contextual factors into account.

1.5. Overview of the dissertation

1.5.1. Overall Goals

The purpose of this dissertation was to prospectively examine the effects of change in fitness on subsequent attendance, overall and by gender, using longitudinal, individual-level data, and taking contextual factors into account. It also aimed to explore the longitudinal, causal effects of change in fitness on chronic absenteeism. Findings from this study can inform school-based policy targeting attendance. The study drew from the NYC Fitnessgram data on approximately 350,000 individual students, comprising 6 cohorts followed over 4 consecutive years (grades 5-8, 2006/07-2012/2013; see Appendix A). Methodologically, this study utilized state-of-the-art analytic techniques to maximize the strengths of the dataset. Multilevel longitudinal data analysis was used to examine repeated outcome measures on individuals clustered within schools and accommodate neighborhood/contextual covariates. Specifically, both 3-level linear and logistic adjusted models were used to causally examine the fitness-attendance relationship (days absent and chronic absenteeism, respectively), after accounting for potential confounders including change in BMI, and sociodemographic factors.

1.5.2. Specific Aims
**Aim 1.** To characterize individual-level and between-school variation in student health-related fitness and attendance, using the NYC Fitnessgram dataset (2006/07-2012/2013).

*Hypothesis 1a.* One-year lagged days absent per year will increase with decreasing fitness (based on aerobic capacity, muscular strength, and endurance tests), increasing grade levels, and increasing school-area poverty.

*Hypothesis 2b.* Between-school variability and school level factors (school-area poverty) will account for a small but significant proportion of total variability in one-year lagged attendance (days absent per year).

**Aim 2.** To analyze the causal effects of change in health-related fitness on subsequent attendance in 6 cohorts of NYC Department of Education (DOE) middle school students followed consecutively over 4 years (grades 5-8).

*Hypothesis 2a.* Higher positive change in individual-level fitness (cardiorespiratory and muscular endurance, and muscular strength fitness composite percentile scores) will negatively predict 1-year lagged individual-level days absent per year after accounting for potential individual- and school-level confounders, as well as accounting for clustering by individual and school, and time-dependent interactions.

*Hypothesis 2b.* Gender will modify the relationship between change in fitness and 1-year lagged attendance. Fitness will be a stronger predictor of attendance in females compared with males.
**Aim 3.** To analyze the causal effects of change in fitness on subsequent chronic absenteeism in 6 cohorts of NYC DOE middle school students followed consecutively over 4 years (grades 5-8).

*Hypothesis 3a.* School chronic absenteeism will decrease with increasing fitness and decreasing grade levels.

*Hypothesis 3b.* Higher positive change in individual-level fitness (categorical variable represented by cardiorespiratory and muscular endurance, and muscular strength fitness composite percentile scores) will predict lower probability in one-year lagged individual-level chronic absenteeism (chronically absent represented by \( \geq 20 \) days absent per year) after accounting for potential individual- and school-level confounders, as well as accounting for clustering by individual and school, and time-dependent interactions.

1.5.3. *Organization of the dissertation*

This dissertation includes five chapters. Individual-level and between-school variation in fitness and attendance using the NYC Fitnessgram dataset (2006/7-2012/13) are described in Chapter 2 (Aim 1). Chapter 3 comprises analyses examining the causal effects of change in fitness on subsequent attendance in 6 cohorts of NYC DOE middle school students followed consecutively over 4 years (Aim 2). The results of analyses on gender as an effect measure modifier in the fitness-attendance causal effects are also presented in Chapter 3. In Chapter 4, findings are presented from analyses testing the longitudinal, causal effects of change in fitness on chronic absenteeism in NYC middle school students followed consecutively over 4 years (Aim 3). Summarized findings from
Chapters 2 through 4, and public health policy recommendations targeting both student fitness and attendance are discussed in Chapter 5.

1.5.4. Significance of the dissertation

This dissertation presents the first study to the author’s knowledge that prospectively examined the effects of change in fitness on attendance in youth, drawing from multilevel, repeated measures data. Findings from this dissertation also stand to offer strong evidence in support of public health interventions which promote opportunities for youth physical activity, including ≥60 minutes of physical activity per day for 6-17 year olds, and quality physical education before, during, and after school. Schools in the US report pressure to replace physical education and other opportunities for physical activity with non-physical instructional time due in part to an increasing emphasis on high-stakes testing.\(^76\). In 2006, for example, <10% of US middle schools provided daily physical education to students in all grades. Elucidating the longitudinal, causal effects of fitness on attendance thus may provide evidence in support of public health policy targeting expansion of school-based physical activity programs.

1.6. Data sources and study population

1.6.1. Primary exposure

Fitness data were drawn from the NYC Fitnessgram dataset jointly managed by NYC DOE and Department of Health and Mental Hygiene (DOHMH), and comprise annual fitness assessments collected by NYC DOE for approximately 870,000 public school students per year (grades K-12) starting in 2006-07. The Fitnessgram is demonstrated to
have both strong reliability and validity.\textsuperscript{28,29} Morrow et al. also demonstrated reliability and validity of the Fitnessgram across testing sites.\textsuperscript{77}

Individual student Fitnessgram data from multiple years were linked by a unique identifier. Given no prior research has examined the fitness-attendance association over multiple years, the decision to use a 1-year lag is based on prior work from longitudinal data demonstrating fitness may promote academic performance,\textsuperscript{52} and also in an effort to maximize the size of the analytic sample while including repeated fitness measures from students during the middle school period.

1.6.2. Primary outcomes

The primary outcomes were days absent per year, measured at year-end, and chronic absenteeism (defined as ≥20 days absent per year). Attendance data was also drawn from the NYC Fitnessgram dataset. Attendance information is collected at year-end and linked to Fitnessgram data by unique student identifiers. Student admission and discharge dates are included in the dataset if they occurred prior to or post-end of the school year.

1.6.3. Individual-level variables

Demographic variables included gender, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian or Pacific Islander, Native American, and other (multiracial or parent refused), and place of birth (NYC, United States (not NYC), or foreign born). This information was based on DOE demographic surveys administered annually to parents and linked to Fitnessgram data by unique student identifiers.
Change in obesity status (obese to not obese, consistently not obese, consistently obese, not obese to obese) was also included as a potential confounder in the models given the literature which generally supports a positive association between obesity status and attendance.\textsuperscript{55,78-83} Obesity status was separated from other fitness variables given aerobic capacity, muscular strength and endurance components of fitness have been shown to predict youth academic performance across all weight categories.\textsuperscript{84} Moreover, while BMI is a good predictor of health-related fitness,\textsuperscript{2,29,31} it cannot be characterized as a fitness test, or physical ability, in and of itself. In contrast, cardiorespiratory endurance, muscular endurance, muscular strength, and flexibility are assessed using physical ability tests, such as timed runs, and other cardiorespiratory, strength, isokinetic, and flexibility tests.\textsuperscript{31} In this sense, body composition versus other fitness components should be treated as conceptually distinct. Change in obesity status was therefore controlled for in the analyses in order to ensure that any effects of fitness on attendance were due to aerobic capacity, muscular strength and endurance components of fitness, and not obesity status.

Height and weight are collected on NYC students during routine physical education classes as part of the Fitnessgram assessment. Age- and gender-specific BMI for children are computed using the following formula: $BMI = \frac{\text{weight (kg)}}{\text{(height (m))}^2}$. Obesity status was defined according to CDC growth chart-derived norms for gender and age in months based on a historical reference population, and used to compute the BMI percentile for each child.\textsuperscript{28} Obesity was defined as having a BMI $\geq 95^{th}$ percentile for youth in the same gender and age in months group.\textsuperscript{85,86}
1.6.4. School-level variables

A categorical school-area poverty variable was based on percentage of households in the school zip code living below the federal poverty threshold (low (<10%), medium (10%-20%), high (20%-30%), and very high (>30%) area poverty) drawing from the American Community Survey (ACS) 2007-2013. Area poverty data were linked to individual student Fitnessgram records based on school zip code.

Inclusion criteria for this study were active enrollment status in a NYC public school for ≥2 consecutive years while in grades 6-8 during the study period (2006/7-2012/13) in districts 1-32 (i.e. the schools that have Fitnessgram measurements; n=457,397). Students were excluded (n=6,225) if they were enrolled for less than n-5 days to ensure a consistent period of observation across school years with different total instructional days per year, where n is the maximum number of days enrolled across all students each year (range: 292-297 days). Next, students were excluded if they did not take the Fitnessgram test for ≥2 consecutive years (n=56,464). Students who attended schools with poor quality fitness data were also excluded from the analysis (n=350). Lastly, students were excluded if they changed schools (to be able to account for school clustering in the analysis; n=44,977; See Appendix B for demographic profile of students who did not meet inclusion criteria). After the above exclusions, the final sample of 6-8th graders comprised 349,381 unique students nested in 624 schools (51% female, 77% NYC born, 38% Hispanic, 28% Non-Hispanic black, and 17% non-Hispanic white; 177,281, 220,769, and 186,135 student-years contributed 6th, 7th and 8th grade data, respectively).
Chapter 2: Individual- and school-level differences in the fitness-attendance association in New York City 6-8th grade youth

2.1. Background

2.1.1. Physical activity and school outcomes

In the United States (US), 42% of children ages 6-11 meet National Association for Sport and Physical Education (NASPE) and World Health Organization (WHO) recommendations for children to have ≥60 minutes of moderate to vigorous physical activity every day (MVPA). Estimates for adolescents ages 12-19 are even lower, ranging from 8-17%. Moreover, physical activity levels have decreased over the last several decades, with steeper declines from childhood to adolescence in the US compared with other nations. These national trends are evident in New York City (NYC), where 43%, 35% and 20% of youth ages 6-10, 11-12, and 14-18 meet physical activity guidelines.

Low levels of youth physical activity in the US are of particular concern given extensive research demonstrating the benefits of physical activity on children’s health and academic performance, potentially acting through pathways involving enhanced cognition and memory, or improvements in both physical and psychosocial wellness. Likewise, recent literature has suggested youth health-related fitness (fitness) may increase student attendance, similar to findings on the association of fitness and workplace attendance in adult populations. Specifically, cardiorespiratory fitness and physical activity in adults are shown to be positively associated with
workplace attendance. Furthermore, interventions targeting improvements in adult fitness have demonstrated an increase in workplace attendance.

2.1.2. Effects of individual and school-level factors on the fitness-attendance association

Individual-level factors associated with the fitness-attendance relationship include gender, individual household socioeconomic status, age or grade level, obesity status, race/ethnicity, home language, and place of birth. However, little is known about the impact of school contextual effects on the fitness-attendance association. School contextual factors, such as area poverty and the built environment are shown in the literature to be associated with children’s tendency to participate in school- and neighborhood-based physical activity. For example, neighborhood factors have been shown to contribute to opportunities for safe, attractive, and accessible physical activity. Likewise, school contextual effects may impact student attendance. Community norms and attitudes may inform parental decisions to permit children’s school absences for family or work obligations. Similarly, perceptions of neighborhood safety may influence student attendance. Given the potential for neighborhoods to affect both physical activity and attendance, it is important to evaluate contextual factors as possible antecedents or confounders in this relationship. However, no papers were identified on the fitness-attendance association that also included school-area measures in their analyses.

2.1.3. Reporting attendance at the individual- and school-levels

Based on education reports and the scientific literature, attendance data are typically
aggregated at the school-level.\textsuperscript{32,73,74,89-91} For example, the NYC DOE reported an average daily student attendance rate of 92\% for all students, and 95\%, 94\% and 93\% for 6\textsuperscript{th}, 7\textsuperscript{th} and 8\textsuperscript{th} grade students, respectively, citywide in 2014-15.\textsuperscript{89,92-94} Information may be lost on account of not including analyses based on student-level data. For example, school-aggregated attendance prevalence rates may not fully capture the extent of chronic absenteeism, nor accurately present nuanced patterns of attendance across students compared with figures based on student-level data. Indeed, a school can have an average 90\% daily attendance rate, while 40\% of its students are chronically absent (i.e. missing ≥20 days per year), depending on the composition of the student population present on a given day.\textsuperscript{72}

This study aimed to characterize individual-level and between-school variation in fitness and student attendance in middle school students using the NYC Fitnessgram dataset (2006/7-2012/13). It was hypothesized that student attendance would increase with increasing fitness, decreasing grade, and decreasing school-area poverty. It was also hypothesized that schools and school-area poverty would account for a small but significant proportion of total variability in student attendance levels.

\textbf{2.2. Methods}

\textbf{2.2.1. Data source, collection, management and study population}

Data were drawn from the NYC Fitnessgram dataset jointly managed by NYC Department of Education (DOE) and Department of Health and Mental Hygiene (DOHMH), and comprised of annual fitness assessments collected by DOE for
approximately 870,000 NYC public school students per year (grades K-12) starting in 2006-07. The NYC DOE is the largest school district in the US, serving approximately 1.1 million students in over 1,800 schools and including approximately 170,000 middle school students in any given year. NYC schools are mandated to have ≥85% of Fitnessgram measurements on their students each year. Individual student fitness data from multiple years are linked in the dataset by a unique student identifier.

Attendance data were also drawn from the NYC Fitnessgram dataset. Attendance information is collected at year-end and linked to Fitnessgram data by unique student identifiers. Student admission and discharge dates are included in the dataset if they occurred prior to or post-end of the school year.

Inclusion criteria for this study were active enrollment status in a NYC public school for ≥2 consecutive years while in grades 6-8 during the study period (2006/7-2012/13) in districts 1-32 (i.e. the schools that have Fitnessgram measurements; n=457,397; Figure 2.1 for sample selection flow chart). Students were excluded (n=6,225) if they were enrolled for less than n-5 days to ensure a consistent period of observation across school years with different total instructional days per year, where n is the maximum number of days enrolled across all students each year (range: 292-297 days). Next, students were excluded if they did not take the Fitnessgram test for ≥2 consecutive years (n=56,464). Students who attended schools with poor quality fitness data were also excluded from the analysis (n=350). Lastly, students were excluded if they changed schools (to be able to
Figure 2.1. Sample Selection Flow

NYC DOE students, enrolled in grades 6, 7, or 8 between 2006/07-2012/13, districts 1-32 (n=457,397)

Number Excluded

- Enrolled <(n-5) days, where n=maximum days enrolled across students per year (n=6,225)

Enrolled ≥(n-5) days, where n=maximum days enrolled across students per year (n=451,172)

Missing Fitnessgram measurements for ≥2 consecutive years while in 5th-8th grades between 2006/07-2012/13 (n=56,464)

Fitnessgram measurements for at minimum two consecutive years (n=394,708)

Students from schools with poor fitness data quality between 2006/07-2012/13 (n=350)

Attended schools with acceptable fitness data quality (n=394,358)

Students who changed schools in 6th-8th grades between 2006/07-2012/13 (n=44,977)

Did not change schools in 6th-8th grades (n=349,381)
account for school clustering in the analysis; n=44,977). After the above exclusions, the final sample of 6-8th graders was comprised of 349,381 unique students nested in 624 schools (51% female, 77% NYC born, 38% Hispanic, 28% Non-Hispanic black, and 17% non-Hispanic white; 177,281, 220,769, and 186,135 student-years contributed 6th, 7th and 8th grade data, respectively).

2.2.2. Primary exposure

The primary exposure was a categorical variable representing age- and gender-specific change in a fitness composite percentile scores based on mean performance on aerobic capacity (Progressive Aerobic Cardiovascular Endurance Run (PACER)), muscle strength and endurance (curl-up and push-up) tests relative to peers (categorized as: >20% increase, 10-20% increase, <10% change, 10-20% decrease, and >20% decrease from the year prior consistent with prior longitudinal research on fitness and academic outcomes and drawing from the NYC Fitnessgram dataset\(^5\)). The PACER comprises an aerobic capacity test for which individuals must run back and forth across a 20-meter space at a specified pace which increases incrementally. The pushup and curlup (i.e. sit-up) also are set to a specified pace. For all three assessments, students are asked to complete as many repetitions as possible.\(^{28,93}\) Students pass the tests based on whether they achieve a score within age- and gender-specific Healthy Fitness Zones (Appendix C).

2.2.3. Primary outcome

Several variables pertaining to student attendance available in the NYC Fitnessgram dataset were examined in order to ensure a consistent period of observation across
students. Initially, **days present** per year was considered as a potential variable to account for students’ total days of attendance per year; however, univariate descriptive analyses revealed that a large proportion of students had values reflecting more days present (approximately 20 days per year) than the maximum possible number of days in the school year for three years of data (2006/7-2008/9). We speculated that inconsistent approaches across schools may have existed at this time, leading some schools to enter attendance data for the entire calendar year (i.e. including summer). In 2010, however, the NYC Mayor’s Interagency Task Force on Chronic Absenteeism was founded, which prioritized improved attendance data collection and reporting by schools, perhaps providing an explanation for improved data quality of the days present variable after this time.⁷³

In lieu of days present per year, a variable in the Fitnessgram dataset representing total **days enrolled** was explored as an inclusion criterion to ensure students were observed for a consistent period of time. The students’ values for this variable were calculated by Fitnessgram data specialists based on the sum of days from admission to discharge date. The first and last days of school were applied to this calculation if admission and discharge dates occurred prior to and post-end of school year, respectively. Although this variable is defined in part based on period of enrollment at a single school, students were clustered by school in this analysis, therefore limiting the population to those students who did not switch schools during the study period was not problematic. The days enrolled variable was therefore used as one of the study exclusion criteria.
The days absent variable was then examined in detail to ensure its appropriateness as an outcome variable. Based on univariate descriptive analyses, the range of values for days absent across students for all years was plausible, with no values exceeding total days in the school year. Next, mean attendance rates across students and across years were calculated and compared with the reported NYC DOE attendance rates (a categorical variable indicating students’ percent of days in attendance that year (<10%, 10-19%, 20-29%, 30-39%, 40-49%, 50-59%, 60-69%, 70-79%, 80-89%, 90-99%, no show, or perfect attendance)) and presented by NYC DOE as an accurate measure of attendance.\textsuperscript{95,96} Specifically, a new variable was created based on the student’s total days absent using the Fitnessgram dataset divided by the total days per year (categorized into units of 10% to align with the DOE attendance rate variable, and similarly including no show and perfect attendance categories). DOE attendance rate categories were then compared with the newly created Fitnessgram attendance rate variable. Across all years (2006/7-2012/13), DOE and Fitnessgram attendance rate categories matched on average 98% (range: 97.9\%-99.3\%). An additional average of 1% of Fitnessgram attendance rate values matched within 10% across the DOE attendance rates (range: 0.6\%-2.4\%).

Finally, year-specific rates of chronic and severe absenteeism for NYC students found in the literature were compared and found to be consistent with values generated based on the Fitnessgram days absent variable.\textsuperscript{72,73,89,97,98} For example, in 2011/12, the rate of chronic and severe absenteeism (defined as being absent 10% (≥20 days) and 20% (≥40 days) per year, respectively)\textsuperscript{89} for NYC 6-8\textsuperscript{th} graders using the Fitnessgram days absent variable and data from this study’s analytic population were 19% and 6%, respectively.
Similarly, chronic and severe absenteeism rates for the same year published by the NYC Mayor’s Interagency Task Force on Truancy, Chronic Absenteeism & School Engagement were 19% and 6%, respectively.\textsuperscript{73} Over time, chronic absenteeism rates for middle school students calculated based on the Fitnessgram days absent variable were found to have decreased approximately 1% per year, from 26% in 2006/7 to 18% in 2012/13. Similarly, trends in chronic absenteeism rates published in the literature show rates of chronic absenteeism decreased by approximately 1% per year during this period.\textsuperscript{72-74,97,98} Based on the above analyses, it was determined that the Fitnessgram days absent variable (a discrete variable; See Appendix D for univariate descriptive plot of days absent variable for the study population) accurately represented student attendance and therefore was appropriate as the primary outcome variable in this analysis.

\textbf{2.2.4. Individual-level variables}

Demographic variables included gender, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian or Pacific Islander, Native American, and other (multiracial or parent refused), and place of birth (NYC, United States (not NYC), or foreign born). This information was based on DOE demographic surveys administered annually to parents and linked to Fitnessgram data by unique student identifiers. Demographic survey data from year one was applied to all other years of data for the same individual.
Change in obesity status (obese to not obese, consistently not obese, consistently obese, not obese to obese) was also included. Height and weight were collected on NYC students during routine physical education classes as part of the Fitnessgram assessment. Age- and gender-specific BMI for children were computed using the following formula:

\[
BMI = \left( \frac{\text{weight (kg)}}{\text{height (m)}} \right)^2
\]

Obesity status was defined according to CDC growth chart-derived norms for gender and age in months based on a historical reference population, and used to compute the BMI percentile for each child.\textsuperscript{28} Obesity was defined as having a BMI $\geq 95^{th}$ percentile for youth in the same gender and age in months group.\textsuperscript{85,86}

2.2.5. School-level variables

A categorical school-area poverty variable was based on average percentage of households in the school zip code living below the federal poverty threshold (low (<10%), medium (10%-20%), high (20%-30%), and very high (>30%) area poverty) drawing from the American Community Survey (ACS) 2007-2013. Area poverty data was linked to individual student Fitnessgram records based on school zip code.

2.2.6. Statistical methods

Descriptive statistics were computed to summarize sample characteristics. Trends in attendance (days absent) were examined at the individual-level and aggregated at the school-level across demographic and fitness-change categories (>20% increase, 10-20% increase, <10% change, 10-20% decrease, and >20% decrease from the year prior).

Trends in attendance (days absent) by fitness-change categories were then examined across grade level (6\textsuperscript{th}-8\textsuperscript{th}) and by school-area poverty (low, medium, high and very high,
based on percent of population living below the federal poverty level (<10%, 10%-20%, 20%-30%, and >30%, respectively)).

Next, mixed models methodology was used to assess between-school variability in student attendance. Given children spend many hours per day interacting within schools and therefore share the same physical surroundings, routines, peers, and teachers/staff, ignoring within-school clustering introduces violations of assumptions of independence of standard errors. In most cases this leads to an underestimate of standard errors, which results in inaccurate beta coefficient estimates and increased Type I (false positive) errors. In this sense, mixed models are particularly relevant to determining the extent to which variance in student outcomes is explained by clustering units (i.e. schools), and examine the effects of group-level factors (e.g. school-area poverty) on person-specific outcomes (e.g. attendance).

First, an unconditional, 2-level Generalized Linear Mixed Model (GLMM) with a random intercept was fit to assess the overall variability in attendance across schools. The unconditional model is represented as: \( \text{Attendance}_{ij} = \beta_{00} + \mu_{0j} + \epsilon_{ij} \), where \( \epsilon_{ij} \sim N(0, \sigma^2) \) and \( \mu_{0j} \sim N(0, \tau^2) \). The Intraclass Correlation (ICC) was calculated as the between-school variance divided by the sum of the between-school and within-school variances. The school-level ICC estimate represented the effect of school clustering, and was calculated as the ratio of the between school variance/(between + within school variance), represented as: \( \sigma^2_{\text{between school}} / (\sigma^2_{\text{between school}} + \sigma^2_{\text{within school}}) \).
Next, a 2-level conditional GLMM model was fit with a random intercept and school-area poverty (SAP) as the predictor, represented as: Attendance\(_{ij}\) = \(\beta_{00} + \beta_{01}(\text{SAP})_j + \mu_{0j} + \epsilon_{ij}\). This model assessed overall variability in attendance across schools explained by contextual school-area poverty based on the average percentage of households in the school zip code living below the federal poverty threshold (described above) and drawing from the ACS 2007-2013. The ICC for this model was calculated as the difference between the unconditional between-school variance and conditional between-school variance, divided by the unconditional between-school variance, represented as:

\[
(\sigma^2_{\text{unconditional between school}} - \sigma^2_{\text{conditional between school}}) / \sigma^2_{\text{unconditional between school}}.
\]

This ICC indicated the percent of between-school explainable variance in student attendance accounted for by school-area poverty. In these analyses, students contributed 1-3 years of fitness-change data. In this sense, students may have contributed fitness-change data for 5-6\(^{th}\), 6\(^{th}\)-7\(^{th}\), and/or 7\(^{th}\)-8\(^{th}\) grades (n=349,381 unique students; 675,318 observations). All analyses were performed using PROC MIXED in SAS 9.4 software (Cary, NC).

### 2.3. Results

#### 2.3.1. Sample characteristics

Table 2.1 presents demographic, fitness, and school characteristics of the study population. The population was comprised of slightly more females (n=177,355; 51%) and mostly Hispanic and Non-Hispanic Blacks (n=134,453; 38% and n=99,363; 28%, respectively). Most students were English-speaking (n=197,727; 57%) and most were born in NYC (n=269,251; 77%).
Across all years and grades, 37% of students (n=253,161) had <10% change in fitness, 20% (n=134,753) had a >20% increase in fitness, and 12% (n=82,117) had a 10-20% increase in fitness based on the difference in composite percentile scores from the year prior. Nineteen percent of students (n=126,115) had a >20% decrease in fitness, and 12% had a 10-20% decrease (n=79,172, 12%) in fitness from the year prior. Also, most students were consistently not obese across all years (n=504,762; 73%), followed by consistently obese (n=119,235; 17%), those who changed from obese to not obese (n=36,029; 5%), and not obese to obese (n=27,273; 4%).

There were 624 schools included in the analysis. The average school size was 559.91 students (SD=713.34), including 365 (58%) small schools (<400 students per school). Twenty six percent (n=89,407) and 22% (n=78,510) of students attended a school in high- or very high-poverty areas, respectively.

| Table 2.1. Demographic and fitness-change characteristics of the study population (N=349,381) |
|---------------------------------|------|-----|
| Gender                         | n    | %   |
| Female                         | 177,355 | 51  |
| Male                           | 172,026 | 49  |
| Race/Ethnicity                 |      |     |
| American Indian                | 1222  | <1  |
| Asian and/or Pacific Islander  | 58,295 | 17  |
| Hispanic                       | 134,453 | 38  |
| Non-Hispanic Black             | 99,363 | 28  |
2.3.2. Descriptive trends in attendance at the individual- and school-level

Table 2.2 presents descriptive attendance trends across student- and school-level characteristics. At the student-level, mean days absent per year from highest to lowest were 11.91 (SD=12.79), 11.14 (SD=12.16), 10.71 (SD=11.88), 10.29 (SD=11.27) and 10.26 (SD=11.15) for students who had a decrease of >20%, decrease of 10-20%, <10%...
change, 10-20% increase, and >20% increase, respectively, in fitness composite percentile scores from the year prior (Figure 2.2).

**Figure 2.2. Days absent across fitness-change categories**

Mean days absent per year were highest among females (M=10.98 (SD=11.67)), and American Indian, Hispanic and Non-Hispanic Blacks race/ethnic groups (M=13.00 (SD=14.80), M=12.60 (SD=12.85), and M=12.28 (SD=13.05), respectively). Mean days absent were also highest among students who spoke English in the home (M=11.89 (SD=12.08)) and who were born in NYC (M=11.36 (SD=12.08)).

Students who attended small schools had higher mean days absent compared with those who attended non-small schools (M=11.96 (SD=12.34) vs. M=10.25 (SD=11.07), respectively). Also, the widest range in days absent by demographic factors was found across students attending schools in in very high-, high-, medium-, and low-poverty areas (M=13.10 (SD=13.31), M=11.13 (SD=11.69), M=9.49 (SD=10.31), and M= 8.51 (SD=9.15), respectively).
<table>
<thead>
<tr>
<th></th>
<th>Student-Level*</th>
<th>School-Level*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>10.98</td>
<td>11.67</td>
</tr>
<tr>
<td>Male</td>
<td>10.09</td>
<td>10.97</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>13.00</td>
<td>14.80</td>
</tr>
<tr>
<td>Asian and/or Pacific Islander</td>
<td>5.50</td>
<td>7.65</td>
</tr>
<tr>
<td>Hispanic</td>
<td>12.60</td>
<td>12.85</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>12.28</td>
<td>13.05</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>10.02</td>
<td>9.65</td>
</tr>
<tr>
<td><strong>Language Spoken at Home</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>11.89</td>
<td>12.08</td>
</tr>
<tr>
<td>Spanish</td>
<td>10.93</td>
<td>11.13</td>
</tr>
<tr>
<td>Other Language</td>
<td>5.98</td>
<td>7.44</td>
</tr>
<tr>
<td><strong>Place of Birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYC</td>
<td>11.36</td>
<td>11.68</td>
</tr>
<tr>
<td>US (Not NYC)</td>
<td>11.03</td>
<td>11.67</td>
</tr>
<tr>
<td>Foreign</td>
<td>11.07</td>
<td>13.75</td>
</tr>
<tr>
<td><strong>Change in Fitness (all years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>10.26</td>
<td>11.15</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>10.29</td>
<td>11.27</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>10.71</td>
<td>11.88</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>11.14</td>
<td>12.16</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>11.91</td>
<td>12.79</td>
</tr>
<tr>
<td><strong>Grade</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grade 6</td>
<td>10.24</td>
<td>11.04</td>
</tr>
<tr>
<td>Grade 7</td>
<td>10.92</td>
<td>12.45</td>
</tr>
<tr>
<td>Grade 8</td>
<td>13.14</td>
<td>14.47</td>
</tr>
<tr>
<td><strong>School-Area Poverty</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Area Poverty</td>
<td>8.51</td>
<td>9.15</td>
</tr>
<tr>
<td>Medium Area Poverty</td>
<td>9.49</td>
<td>10.31</td>
</tr>
<tr>
<td>High Area Poverty</td>
<td>11.13</td>
<td>11.69</td>
</tr>
<tr>
<td>Very High Area Poverty</td>
<td>13.10</td>
<td>13.31</td>
</tr>
<tr>
<td>------------------------</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td><strong>School Size</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small Schools (&lt;400 students)</td>
<td>11.96</td>
<td>12.34</td>
</tr>
<tr>
<td>Non-Small Schools (≥400 students)</td>
<td>10.25</td>
<td>11.07</td>
</tr>
</tbody>
</table>

*N*missing Place of Birth* = 72; *N*missing Area Poverty* = 7; *N*missing or > 1 race/ethnicity* = 177. *School-level columns account for school-clustering; Student-level columns do not account for school-clustering.*

Similar to trends found at the individual-level, when attendance data was aggregated at the school-level, mean days absent were highest in students with the greatest decrease (>20%) in fitness (M=12.65 (SD=13.21)) and lowest in students with the greatest increase (>20%) in fitness composite percentile scores from the year prior (M=10.78 (SD=11.54); Table 2.2). Similar to the student-level, mean days absent per year at the school-level were highest among females (M=11.23 (SD=11.51)) and students who spoke English in the home (M=12.04 (SD=11.93)). Mean days absent were also highest at the school-level in students born in NYC (M=11.39 (SD=11.55)). At the school-level, students attending schools in very high, high, medium and low poverty areas had the most to least mean days absent per year (M=13.10 (SD=12.86), M=11.37 (SD=11.57), M=9.83 (SD=10.22), and M=8.90 (SD=9.26), respectively).

Unlike that described above for student-level attendance, mean days absent aggregated at the school-level were highest in Hispanic students (M=13.27 (SD=13.18)), followed by American Indian and Non-Hispanic Blacks (M=12.85 (SD=14.43) and M=12.80 (SD=13.26), respectively). Lastly, there was no difference at the school-level in the mean days absent per year across small and non-small schools (though standard deviations differed; M=11.76 (SD=11.89) and M=11.76 (SD=10.99), respectively).
2.3.3. Descriptive trends in attendance by fitness, grade, and school-area poverty

Overall, attendance decreased with increasing grade in NYC middle school students (2006/7-2012/13) at both the student- and school-levels, with a large decrease in attendance from 7th to 8th grades (Table 2.3; Figure 2.3).

There was a general trend of decreasing attendance (increasing days absent) with decreasing fitness across increasing grade levels. For example, mean days absent for students with the greatest increase (>20%) in fitness were 9.56 (SD=10.11), 9.85 (SD=10.81) and 11.87 (SD=12.73), for students in 6, 7 and 8th grades, respectively. In contrast, mean days absent for students with the greatest decrease (>20%) in fitness were 10.62 (SD=11.27), 11.57 (SD=12.62), and 13.87 (SD=14.32), for students in 6, 7 and 8th grades, respectively (Table 2.3).

As shown in Table 2.4 and Figure 2.4, attendance was found to decrease with decreasing fitness and increasing school-area poverty. For example, students with the greatest decrease (>20%) in fitness who attended schools in the lowest compared with highest area poverty had a mean days absent of 10.11 (SD=10.52) vs. 14.04 (SD=14.74) respectively. In contrast, students with the greatest increase (>20%) in fitness who attended schools in the lowest compared with highest area poverty had a mean days absent of 8.63 (SD=9.07) vs. 12.44 (SD=13.07) respectively.
Table 2.3. Mean attendance\(^c\) for New York City public school students in grades 6-8, by fitness change from the year prior (N=349,381; 675,318 observations\(^a\))

<table>
<thead>
<tr>
<th>Change in Fitness(^b)</th>
<th>Grade 6</th>
<th></th>
<th>Grade 7</th>
<th></th>
<th>Grade 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
<td>Mean (SD)</td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>53790</td>
<td>9.55(10.14)</td>
<td>45562</td>
<td>9.83(10.82)</td>
<td>35401</td>
<td>11.89(12.75)</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>23683</td>
<td>9.71(10.37)</td>
<td>31937</td>
<td>9.72(10.98)</td>
<td>26497</td>
<td>11.50(12.24)</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>62129</td>
<td>10.01(10.88)</td>
<td>101514</td>
<td>10.07(11.67)</td>
<td>89518</td>
<td>11.93(12.65)</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>21463</td>
<td>10.18(10.91)</td>
<td>31515</td>
<td>10.62(11.90)</td>
<td>26194</td>
<td>12.55(13.28)</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>43524</td>
<td>10.64(11.28)</td>
<td>46875</td>
<td>11.57(12.64)</td>
<td>35916</td>
<td>13.89(14.37)</td>
</tr>
</tbody>
</table>

\(^a\)Observations account for 1-3 years of fitness-change per student. \(^b\)Based on change in fitness composite percentile scores from the year prior. \(^c\)Tabulated mean estimates.

Table 2.4. Mean attendance\(^a\) for New York City public school students across school-area poverty (SAP)\(^b\) by level of fitness-change from the previous year (N=349,381; 675,318 observations)\(^c\)

| Change in Fitness\(^d\) | Low SAP | | Medium SAP | | High SAP | | Very High SAP | |
|-------------------------|---------|-------|------------|-------|---------|-------|---------|
|                         | n       | Mean(SD) | n       | Mean(SD) | n       | Mean(SD) | n       | Mean(SD) |
| >20% Increase           | 24074   | 8.63(9.07) | 47620   | 9.35(10.13) | 32443   | 10.74(11.62) | 30611   | 12.44(13.07) |
| 10-20% Increase         | 15397   | 8.64(9.00) | 29104   | 9.38(10.39) | 20278   | 10.72(11.48) | 17335   | 12.78(13.56) |
| <10% Change             | 49070   | 8.74(9.37) | 88654   | 9.53(10.57) | 63218   | 11.44(12.37) | 52197   | 13.70(14.54) |
| 10-20% Decrease         | 14251   | 9.32(10.05) | 26807   | 10.07(11.05) | 20664   | 11.88(12.85) | 17438   | 13.39(13.98) |
| >20% Decrease           | 20384   | 10.11(10.52) | 40590   | 10.82(11.90) | 34237   | 12.34(12.82) | 30894   | 14.04(14.74) |

\(^a\) Days absent per year. \(^b\)Based on percent of population living below the federal poverty level (<10%, 10%-20%, 20%-30%, and >30%, respectively from the American Community Survey). \(^c\)Based on change in fitness composite percentile scores from the year prior. \(^d\)Tabulated mean estimates.
Figure 2.3. Mean days absent by grade across fitness-change\(^a\) categories\(^b\)

\(\text{Days absent per year} \quad \text{Grade 6} \quad \text{Grade 7} \quad \text{Grade 8} \)

\(\begin{array}{|c|c|c|c|}
\hline
\text{Change in fitness} & \text{>20% Decrease} & \text{10-20% Decrease} & \text{<10% Change} \\
\hline
\text{Grade 6} & 10 & 8 & 6 \\
\text{Grade 7} & 12 & 10 & 8 \\
\text{Grade 8} & 14 & 12 & 10 \\
\hline
\end{array} \)

\(^a\)Based on change in fitness composite percentile scores from the year prior. \(^b\) Based on tabulated mean estimates.

Figure 2.4. Mean days absent by school-area poverty across fitness-change\(^a\) categories\(^b\)

\(\text{Days absent per year} \quad \text{Low SAP} \quad \text{Medium SAP} \quad \text{High SAP} \quad \text{Very High SAP} \)

\(\begin{array}{|c|c|c|c|c|}
\hline
\text{Change in fitness} & \text{>20% Decrease} & \text{10-20% Decrease} & \text{<10% Change} & \text{10-20% Increase} \\
\hline
\text{Low SAP} & 4 & 2 & 0 & 6 \\
\text{Medium SAP} & 6 & 4 & 2 & 0 \\
\text{High SAP} & 8 & 6 & 4 & 2 \\
\text{Very High SAP} & 10 & 8 & 6 & 4 \\
\hline
\end{array} \)

\(^a\)Based on change in fitness composite percentile scores from the year prior. \(^b\)Tabulated mean estimates.
2.3.4. Variation in student attendance accounted for by schools

The overall mean days absent across all schools ($n_{\text{schools}}=624$) was 11.85 days per year (unconditional model intercept). Table 2.5 shows the between- and within-school variances in student attendance. Between-student variation in attendance was found to be much higher compared with between-school variation in attendance. The ICC estimate, however, demonstrated a sizable degree of variance in student attendance explained by schools (11%, $p<.001$).

When school-area poverty was added to the model (conditional model), the ICC estimate diminished but remained substantial and significant (9%, $p<.001$; Table 2.5). The percent of between-school variance in student attendance accounted for by school-area poverty was 20% ($\frac{(16.2-13.0)}{16.2}$).

| Table 2.5. Variability in attendance\(^a\) explained by clustering students in schools |
|---------------------------------|----------------|----------------|----------------|----------------|
|                                |                |                |                |                |
| **Unconditional model\(^{bc}\)** |                | 95% CI         |                |                |
| Between (schools)              | 16.2           | 14.34          | 18.45          |                |
| Within (residual)              | 128.5          | 128.07         | 128.93         |                |
| ICC (proportion of variance at the school-level) | 11.2%          |                |                |                |
| **School-level poverty-added model\(^{cd}\)** |                |                | 95% CI         |                |
| Between (schools)              | 13.0           | 11.49          | 14.81          |                |
| Within (residual)              | 128.5          | 127.6          | 128.46         |                |
| ICC (proportion of variance at the school-level) | 9.2%           |                |                |                |

\(^a\)Days absent over a 1-year period. \(^b\)Random intercept model with no predictors. \(^c\)Random intercept model with school-area poverty predictor based on percent of population living below the federal poverty level (<10%, 10%-20%, 20%-30%, and >30%, respectively, from the American Community Survey. \(^d\)p<.0001.
2.4. Discussion

A large body of research demonstrates a positive or neutral association between youth fitness and academics, although limited research exists on the fitness-attendance relationship, or the impact of school contextual effects on that relationship. This study presents the first paper to the author’s knowledge that descriptively examines multilevel attendance and fitness trends, drawing from a large and diverse urban sample of several student cohorts and spanning multiple years. The present study found a larger decrease in attendance (increasing days absent) from 6th to 8th grades in students with decreasing fitness, perhaps indicative of an interaction between fitness and time. Similarly, trends in attendance by fitness and school-area poverty showed a larger decrease in attendance with decreasing fitness in settings where school-area poverty is high compared to where it is low. In sum, descriptive analyses showed the expected trends in attendance by individual- and school-level factors. Future work should address the causal fitness-attendance relationship, including drawing from repeated measures analyses and considering interactive effects.

These findings further contribute to a literature which shows that physical activity benefits youth academic performance and attendance. Given this body of work, it is plausible that children’s physical activity may predict academic performance through a pathway involving improved fitness, followed by improved attendance, and resulting in higher academic performance. Attendance in fact is well documented in the literature to predict academic performance. For example, a one standard deviation increase in the days a student is absent is shown to be associated with a statistically significant 0.45 and 0.39 standard deviation
change in school grade point average for elementary and middle school students, respectively (p<.01). Further work should formally examine the role of attendance as a mediator linking children’s health-related fitness to academic performance, as schools in the US report pressure to replace physical education and other opportunities for physical activity with non-physical instructional time due in part to an increasing emphasis on high-stakes testing. Elucidating attendance’s role as a mediator in the fitness-academics association thus will lend further support for why public health policy targeting fitness may promote student academic success, and furthermore can help justify increasing the number of schools that provide daily physical education to students in all grades.

This study also found that school-level clustering in attendance was sizeable and statistically significant, demonstrating that taking school-level clustering into account is important in analyses drawing from student data. Children naturally spend large portions of their day clustered in classrooms, grades, and schools. Given this, they are exposed to common social and physical factors, such as the influence of particular peer groups, teachers and administrative staff, and shared access to resources (including physical spaces, classroom materials, and also student-teacher ratios). Moreover, children who attend the same school share daily school-neighborhood exposures. Indeed, in this analysis, school-area poverty accounted for a large proportion of the variance in student attendance at the school-level. This paper therefore contributes to a body of literature which suggests that student attendance is shaped not only by student-specific factors, but also the characteristics of the schools they attend. For example, prior work demonstrates that recreational resources in the areas surrounding schools predicts youth fitness. Similarly, the literature has shown that children living in employment- and
income-deprived areas have lower attendance. Future work should further examine the effects of area poverty, and include the effects of other school-level factors such as the built environment on the fitness-attendance relationship.

Lastly, this study found differences between individual- and school-aggregated attendance estimates, with standard deviations generally larger at the school-level. While in general, the literature reports on average attendance at the school-level, these figures do not reveal individual student patterns of absence, and do not permit tracking changes in attendance at the individual level. This is particularly important given individual-level factors may be important in predicting attendance, and may interact with school-level factors to influence attendance. In this sense, work documenting student attendance prevalence rates should include figures based on student-level data. Moreover, more nuanced research is necessary to examine individual- and school-level factors associated with attendance, and to assess the full extent of chronic absenteeism when student-level data is considered.

Strengths of this study include drawing from a rich multilevel data, and a large and diverse study sample of approximately 350,000 individual students comprising 6 cohorts each followed 4 consecutive years during a seven-year study period (2006/07-2012/2013). Only 5 studies have been identified by the author that examine the specific fitness-attendance association, and only 1 study draws from data on individual students followed over time (12 months). Likewise, to the author’s knowledge, this is the only paper to simultaneously examine individual- and school-level factors that may affect the fitness-attendance relationship.
Furthermore, this study drew from the NYC Fitnessgram, which comprises an objective physical test battery that is demonstrated to have both strong reliability and validity.\textsuperscript{28,29}

2.4.1. Limitations
This analysis may be limited in that a large number of NYC students were not included due to insufficient period of school enrollment, moving schools or not having Fitnessgram tests for $\geq 2$ consecutive years (see Appendix B for excluded population demographics). As expected, students who do not meet inclusion criteria would be more likely to have lower attendance (higher number of days absent) given psychosocial and family factors associated with moving and long-term absences, potentially misrepresenting descriptive trends in attendance. To determine whether these exclusions impacted findings, sensitivity analyses were performed. When inclusion criteria were widened to include students excluded due to enrollment criteria, descriptive estimates remained similar. To assess whether estimates were different in small ($<400$ students) compared with non-small ($\geq 400$ students) schools, analyses were stratified by school size, and effects remained similar in both groups.

There is potential for systematic bias and differential measurement error given the Fitnessgram dataset includes data not collected for research purposes. Data was not available on many student and school-level factors, including psychosocial measures, drug and alcohol use, family structure, and individual household poverty. Potential sources of systematic bias may have led to exposure and outcome misclassification. For example, there may have been variation across Fitnessgram testing sites in testing protocol that impacted performance. However, prior research has not detected a learning effect when fitness tests are repeated.\textsuperscript{105} Moreover, despite
variability across administer at different schools, testing protocols are designed to promote consistency across administer, including manuals, video-based training and site-visits, as well as use of calibrated scales. Further, Morrow et al. demonstrated reliability and validity of the Fitnessgram across testing sites. It is also possible that schools may report attendance differently (e.g. time of day absences are reported), although the prioritization of improved attendance data collection described above, initiated in 2010 with the inception of the NYC Mayor’s Interagency Task Force on Chronic Absenteeism would likely minimize misclassification of outcome.

This study also may be limited in its application of school-based poverty as a proxy for individual household poverty. Following recent guidelines from the NYC DOHMH, an area-based poverty measure, school-area poverty, was used in lieu of individual student-level meal code status. Area-based poverty may better capture health disparities resulting from socioeconomic differences across individuals, particularly in NYC where great disparities exist in health resources and opportunities across different neighborhoods. Given some students may attend schools in different neighborhoods from where they reside, area-based socioeconomic factors may differ between school and home areas potentially having different influences on the fitness-attendance association. While a more accurate poverty measure may have been to use home-area poverty, home address or zip code was not available in the NYC Fitnessgram dataset for the majority of the analytic population. Future research should be devoted to examining the relationship between school and home zip code to better address the impact of employing different poverty measures in fitness and student attendance research.
In addition, it should be noted that the models presented here are likely mispecified given normal distributional assumptions of linear mixed models are not met. While the outcome (days absent) is count-data and its distribution is Poisson (see Appendix D), the ICC definition is not well defined for Poisson models. Given this, multilevel linear models were used.

2.4.2. Conclusion

Descriptive trend analyses suggested increasing fitness may be associated with increasing days absent. Fitness may contribute to global improvements on a wellness continuum, with potential positive effects on student attendance. Given nearly 70% of NYC children ages 6-8 report not taking part in any structured physical activity program, and less than half report engaging in physical education >1 hour per week, further research is warranted to examine the potential causal effects of fitness on attendance by conducting longitudinal, repeated measures multilevel analyses controlling for both individual and school-level factors. Similarly, additional analyses are needed to further explore fitness-interactions in the fitness-attendance relationship. If fitness is shown to be causally related to attendance, this work stands to offer strong evidence in support of public health interventions which promote opportunities for youth physical activity, including ≥60 minutes of physical activity per day for 6-17 year olds, and quality physical education before, during, and after school.
Chapter 3: Investigating the causal longitudinal effects of fitness and lagged attendance in New York City middle-school youth

3.1. Background

Extensive research demonstrates the benefits of youth physical activity and health-related fitness (fitness) on academic outcomes,\textsuperscript{3,10-12} potentially acting through pathways involving enhanced cognition and memory,\textsuperscript{12-18} or improvements in both physical and psychosocial wellness.\textsuperscript{19-27} However, only 42\% of children in the United States (US) aged 6-11 years old meet international physical activity recommendations compared with 97\% and 62-82\% of 9 and 15 year olds in Western European countries.\textsuperscript{3,45} Moreover, declines in physical activity are steeper from childhood to adolescence in the US compared with other nations.\textsuperscript{8} Evidence for these declines are reflected in New York City (NYC) specifically, where 40\%, and 20\% of youth ages 6-12 and 14-18, respectively, meet physical activity guidelines.\textsuperscript{87,88}

Another key predictor of academic performance is school attendance,\textsuperscript{55,99-102} which may mediate the observed fitness-academic achievement association, consistent with research on adults demonstrating fitness predicts higher work attendance.\textsuperscript{37,38,42} Hypothesized causal mechanisms include the contribution of fitness to global health improvements that advance adults along a wellness continuum.\textsuperscript{35-38,40-43,110} For example, it is shown that individuals with higher fitness have lower risk of cardiovascular disease, insulin sensitivity, hypertension, and metabolic syndrome, perhaps contributing to reduced absences related to illness.\textsuperscript{37} Improvements in diet and physical activity may similarly reduce negative health effects and psychosocial problems associated with overweight and obesity.\textsuperscript{39} In this sense, fitness may positively predict
attendance in youth, potentially by contributing to enhanced physical and psychosocial wellness, as well as global improvements on a wellness continuum.\textsuperscript{35-38,40-43,110}

3.1.1. \textit{Longitudinal data to study the fitness-attendance relationship}

To date, limited research has examined the fitness-attendance association. To the author’s knowledge, only 1 longitudinal study and 4 cross-sectional studies have addressed the specific relationship between youth fitness and attendance.\textsuperscript{19,22,32-34} Findings from this literature are consistent in demonstrating a positive association between fitness and attendance.\textsuperscript{19,22,32-34} However, these studies draw predominantly from cross-sectional data, and do not account for a wide range of potential confounders, including contextual factors. Further, the bulk of research on fitness and attendance are unable to support causal hypotheses given temporality of exposure and outcome are not known. Nuanced research in this area that draws from individual-level measures collected over multiple years and includes school-level factors is necessary to better inform school-based physical activity programs targeting increased attendance.

3.1.2. \textit{Gender differences in the fitness-attendance association}

It is also important to investigate potential gender effect measure modification (EMM) in the fitness-attendance association in light of findings suggesting that it may be important. Specifically, numerous studies demonstrate low self-esteem in adolescent girls is significantly associated with both lower physical activity levels\textsuperscript{21} and attendance,\textsuperscript{25,26} attributed in part to perceived weight status and self-appearance.\textsuperscript{24,62,63} Some literature drawing from Social Cognitive Theory (SCT)\textsuperscript{64-71} supports the hypothesis that psychosocial factors serve as antecedents to fitness, by precluding students’ tendency to participate in physical activities. In
this sense, lower self-esteem in young women may explain gender differences in the fitness-attendance relationship.

To the author’s knowledge, no prior studies have specifically addressed differences in the fitness-attendance relationship by gender. Six studies examined whether gender modifies the fitness-academics association, four of which found stronger effects for females.\textsuperscript{27,48,50,52} For example, Bezold found girls with a substantial increase in fitness relative to peers (0.36 percentile points per year more than the reference group) showed the largest increase in academic ranking (1.06 percentile points per year).\textsuperscript{52} One study found no significant differences by gender,\textsuperscript{49} and 1 paper found stronger effects for boys, although the study sample was younger (elementary school-aged) children.\textsuperscript{23}

The purpose of this study was to analyze the causal longitudinal effects of change in fitness on subsequent attendance in 6 cohorts of NYC Department of Education (DOE) middle school students followed consecutively over 4 years (fitness-change from grades 5-6, 6-7 and 7-8 paired with days absent per year for grades 6, 7, and 8, respectively) during a seven-year study period (2006/7-2012/13). It was hypothesized that change in fitness (cardiorespiratory, muscular endurance, and muscular strength fitness composite percentile scores) from the year prior would positively predict subsequent attendance (days absent per year) after accounting for potential individual- and school-level confounders, as well as clustering by individual and school, and time-dependent interactions (see Appendix E for Directed Acyclic Graph). It was further hypothesized that gender would modify the relationship between change in fitness and
1-year lagged attendance, and that fitness would be a stronger predictor of attendance in females compared with males.

3.2. Methods

3.2.1. Data source, collection, management and study population

Data were drawn from the NYC Fitnessgram dataset jointly managed by NYC DOE and Department of Health and Mental Hygiene (DOHMH), and comprised of annual fitness assessments collected by DOE for approximately 870,000 NYC public school students per year (grades K-12) starting in 2006-07.

The Fitnessgram is demonstrated to have both strong reliability and validity.\textsuperscript{28,29} NYC schools are mandated to have $\geq 85\%$ of Fitnessgram measurements on their students each year.\textsuperscript{93} Individual student Fitnessgram data from multiple years are linked in the dataset by a unique identifier. Inclusion criteria for this study included enrollment in a NYC public school for $\geq 2$ consecutive years while in grades 6-8 during the study period (2006/07-2012/13) while attending a school that collected Fitnessgram measurements (see Chapter 2 for a more in-depth discussion of sample selection criteria). The final sample of 6-8\textsuperscript{th} graders was comprised of 349,381 unique students (51\% female, 77\% NYC born, 38\% Hispanic, 28\% Non-Hispanic black, and 17\% non-Hispanic white; 177,281, 220,769, and 186,135 student-years contributed 6\textsuperscript{th}, 7\textsuperscript{th} and 8\textsuperscript{th} grade data, respectively) nested in 624 schools (mean school student population=559.91, SD=713.34).

3.2.2. Primary Exposure
The primary exposure was a categorical variable representing age- and gender-specific percent change in a composite fitness score based on mean performance on aerobic capacity (Progressive Aerobic Cardiovascular Endurance Run (PACER)), muscle strength and endurance (curl-up and push-up) tests relative to peers (categorized as: >20% decrease, 10-20% decrease, <10% change, 10-20% increase, and >20% increase (reference category) from the year prior, consistent with prior longitudinal research on fitness and academic outcomes drawing the NYC Fitnessgram dataset\textsuperscript{52}). The PACER comprises an aerobic capacity test for which individuals must run back and forth across a 20-meter space at a specified pace which increases incrementally. The pushup and curlup (i.e. sit-up) also are set to a specified pace. For all three assessments, students are asked to complete as many repetitions as possible.\textsuperscript{28,92} Students pass the tests based on whether they achieve a score within age- and gender-specific Healthy Fitness Zones (Appendix C).

3.2.3. Primary outcome

The primary outcome variable for this analysis was student days absent per year. Students were excluded (n=6,225) if they were enrolled for less than n-5 days to ensure a consistent period of observation across school years with different total instructional days per year, where n is the maximum number of days enrolled across all students each year (range: 292-297 days). The process for selecting study variables from the NYC Fitnessgram dataset that pertain to attendance and period of observation is detailed in prior work (See Chapter 2).

3.2.4. Individual-level variables
Demographic variables included gender, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian or Pacific Islander, Native American, and other (multiracial or parent refused), and place of birth (NYC, United States (not NYC), or foreign born). This information was based on DOE demographic surveys administered annually to parents and linked to Fitnessgram data by unique student identifiers.

Change in obesity status (obese to not obese, consistently not obese, consistently obese, not obese to obese) was also included (see Chapter 2 for detailed background on variable specification). Height and weight are collected on NYC students during routine physical education classes as part of the Fitnessgram assessment. Age- and gender-specific BMI for children are computed using the following formula: \( \text{BMI} = \frac{\text{weight (kg)}}{\text{height (m)}^2} \). Obesity status was defined according to CDC growth chart-derived norms for gender and age in months based on a historical reference population, and used to compute the BMI percentile for each child.\(^{28}\)

Obesity was defined as having a BMI \( \geq 95^{\text{th}} \) percentile for youth in the same gender and age in months group.\(^{85,86}\)

### 3.2.5. School-level variables

A categorical school-area poverty variable was based on percentage of households in the school zip code living below the federal poverty threshold (low (<10%), medium (10%-20%), high (20%-30%), and very high (>30%) area poverty) drawing from the American Community Survey (ACS) 2007-2013. Area poverty data was linked to individual student Fitnessgram records based on school zip code.
3.2.6. Statistical methods

Descriptive statistics were computed to summarize sample characteristics. Next, trends in attendance (days absent) by fitness, grade and gender were examined.

A 3-level longitudinal Generalized Linear Mixed Model (GLMM) was fit to assess the causal fitness-attendance relationship while accounting for school clustering. Individual attendance observations were nested within students, nested within schools. A random intercept was used to model the mean response by child, and a separate random intercept was used to model the mean response by school. Mixed models simultaneously examined the effects of individual- and school-level factors on student-specific attendance. Ignoring within-school clustering introduces violations of assumptions of independence of standard errors that may result in an underestimate of standard errors and increased Type I (false positive) errors.\textsuperscript{52,53}

First, an unconditional longitudinal 3-level random-intercepts model (model 1) was fit to determine the extent of variation in attendance at the student- and school-levels. The student-level (within) Intraclass Correlation (ICC) estimate represented the extent of within-student clustering (the variation in attendance due to clustering of observations within the same student). The student-level (between) ICC represented the effect of between-student differences. The school-level ICC estimate represented the effect of school clustering. The student- and school-level ICC's were calculated as the ratio of the variance for the student (within or between) or school, respectively, divided by the sum of the 3 variance parameter estimates, represented as:
\[ \frac{\sigma^2_{\epsilon}}{(\sigma^2_{\text{student}} + \sigma^2_{\text{school}} + \sigma^2_{\epsilon})}, \quad \frac{\sigma^2_{\text{student}}}{(\sigma^2_{\text{student}} + \sigma^2_{\text{school}} + \sigma^2_{\epsilon})} \quad \text{and} \quad \frac{\sigma^2_{\text{school}}}{(\sigma^2_{\text{student}} + \sigma^2_{\text{school}} + \sigma^2_{\epsilon})}, \]

respectively.

Next, the predictor, student-specific change in fitness from the year prior was added to the model (model 2) to assess the causal longitudinal effect of change in fitness on subsequent attendance. Change in ICC estimates represented the proportion of variance explained by including student-level change in fitness in the model. Beta coefficients represented the effects of the exposure, change in fitness (categorized as: >20% increase, 10-20% increase, <10% change, 10-20% decrease, and >20% decrease (reference category) from the year prior) on outcome, 1-year lagged attendance (days absent per year), among students enrolled in 6-8th grades.

The above model was next adjusted for potential individual- and group-level confounders (model 3). Confounding variables included time (grade-level, level-one time-varying covariate), calendar year (level-one time-varying covariate to control for potential cohort effects), individual socio-demographic factors (gender, race/ethnicity, place of birth (NYC, US (not NYC) and foreign; level-two covariates), change in obesity status from the year prior (level-one time-varying covariate), starting fitness (grand-mean centered; level-one covariate), and school-level school-area poverty (percentage of households in the zip code living below the federal poverty level based on the ACS (2008-2012) categorized as <10%, 10%-20%, 20%-30%, and >30%; level-three covariate). The final, 3-level model was represented as:

\[
\begin{align*}
\text{Absent}_{ij} &= \beta_{000} + \beta_{100}(\text{GRADE})_{ij} + \beta_{200}(\text{FITNESS})_{ij} + \beta_{300}(\text{CHANGE IN OBESITY STATUS})_{ij} + \beta_{400}(\text{YEAR})_{ij} + \beta_{510}(\text{RACE/ETHNICITY})_{ij} + \beta_{520}(\text{GENDER})_{ij} + \beta_{530}(\text{PLACE of BIRTH})_{ij} + \beta_{540}(\text{STARTING FITNESS})_{ij} + \beta_{501}(\text{SCHOOL AREA POVERTY})_{ij} + \\
&\quad \beta_{510}(\text{GRADE})_{ij}^{*}(\text{RACE/ETHNICITY})_{ij} + \beta_{520}(\text{GRADE})_{ij}^{*}(\text{GENDER})_{ij} + \beta_{530}(\text{GRADE})_{ij}^{*}
\end{align*}
\]
(PLACE of BIRTH)ij + β140(GRADE)ij*(STARTING FITNESS)ij + β011(SCHOOL AREA POVERTY)ij*(RACE/ETHNICITY)ij + εij + r0ij + μ00j

Lastly, gender effect measure modification in fitness on lagged attendance was examined.

GLMM models were stratified by gender (models 4 and 5), where individual student change in days absent (outcome variable) was predicted by change in fitness from the year prior (independent variable, categorized as above). As above, time was represented by each student’s grade and potential confounders were included in final models (individual level race/ethnicity, place of birth, change in obesity status, year, starting fitness, and group-level school-area poverty).

In these analyses, students contributed 1-3 years of fitness-change data. In this sense, students may have contributed fitness-change data for 5-6th, 6th-7th, and/or 7th-8th grades (n=349,381 unique students; 675,318 observations). All analyses were performed using PROC MIXED in SAS 9.4 software (Cary, NC).

3.2.7. Regression diagnostics and sensitivity analyses

A full residual analysis was performed (see Appendix F). Univariate distributions for the days absent variable demonstrated a long right-tailed Poisson distribution, therefore models also were run using PROC GLIMMIX (see Appendix G For Poisson model estimates) to determine the extent to which model misspecification impacted findings. Also, sensitivity analyses were performed on whether enrollment exclusions and BMI variable specification impacted findings.
3.3. Results

3.3.1. Sample characteristics

As shown in Table 3.1, the analytic population included approximately 350,000 students and 624 schools, and was comprised of slightly more females (51%), Hispanic (38%), and English speaking students (57%). The majority of students were born in NYC (77%) and were consistently not obese across all years (73%). Also, just under 40% of students had <10% change in fitness from the year prior, followed by >20% increase (20%), >20% decrease (19%), 10-20% increase (12%), and 10-20% decrease (12%).

<table>
<thead>
<tr>
<th>Gender</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>177355</td>
<td>51</td>
</tr>
<tr>
<td>Male</td>
<td>172026</td>
<td>49</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>1222</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Asian and/or Pacific Islander</td>
<td>58295</td>
<td>17</td>
</tr>
<tr>
<td>Hispanic</td>
<td>134453</td>
<td>38</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>99363</td>
<td>28</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>55857</td>
<td>16</td>
</tr>
<tr>
<td>Language Spoken at Home</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>197727</td>
<td>57</td>
</tr>
<tr>
<td>Spanish</td>
<td>86052</td>
<td>25</td>
</tr>
<tr>
<td>Other language</td>
<td>65602</td>
<td>19</td>
</tr>
<tr>
<td>Place of Birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYC</td>
<td>269251</td>
<td>77</td>
</tr>
<tr>
<td>US (Not NYC)</td>
<td>19909</td>
<td>6</td>
</tr>
<tr>
<td>Foreign</td>
<td>60149</td>
<td>17</td>
</tr>
<tr>
<td>Change in Fitness (all years)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
3.3.2. Attendance by fitness, grade and gender

As shown in Table 3.2 and Figures 3.1 and 3.2, girls showed more days absent than boys across all grades, and boys had higher levels of fitness as well as greater variability in fitness over time compared with girls. Also, attendance rates by grade, fitness and gender showed steeper declines in days absent from grades 6-8 across increasing fitness categories for girls compared with boys (Table 3.2). For example, mean days absent for girls with increasing fitness (>20% decrease, 10-20% decrease, <10% change, 10-20% increase, and >20% increase in fitness composite percentiles from the year prior) were 11.06 (SD=11.76), 10.60 (SD=11.20), 10.43 (SD=11.07), 10.08 (SD=10.65) and 9.97 (SD=10.39) vs. 14.27 (SD=14.93), 12.76 (SD=13.36), 12.27 (SD=13.00), 11.71 (SD=12.32) and 11.97 (SD=12.65), for 6 vs. 8th grades, respectively. In contrast, mean days absent for boys with decreasing fitness (>20% increase, 10-20%, increase, <10% change, 10-20% decrease, and >20% decrease) were 10.23 (SD=10.78), 9.75
(SD=10.57), 9.55 (SD=10.66), 9.32 (SD=10.06) and 9.14 (SD=9.86) vs. 13.51 (SD=13.77),
12.33 (SD=13.19), 11.60 (SD=12.28), 11.27 (SD=12.15) and 11.80 (SD=12.85), for 6 vs. 8<sup>th</sup>
grades, respectively.

**Figure 3.1. Mean days absent by grade across fitness-change<sup>a</sup> categories in girls<sup>b</sup>**

![Graph showing mean days absent by grade across fitness-change categories in girls](image)

<sup>a</sup>Based on change in fitness composite percentile scores from the year prior. <sup>b</sup>Tabulated mean estimates.

**Figure 3.2. Mean days absent by grade across fitness-change<sup>a</sup> categories in boys<sup>b</sup>**

![Graph showing mean days absent by grade across fitness-change categories in boys](image)

<sup>a</sup>Based on change in fitness composite percentile scores from the year prior. <sup>b</sup>Tabulated mean estimates.
Table 3.2. Mean in attendance\(^a\) for New York City public school students in grades 6-8, by level of fitness-change from the previous year across gender (N=349,381; 675,318 observations\(^{bc}\))

<table>
<thead>
<tr>
<th>Girls (n=177355)</th>
<th>Grade 6</th>
<th>Grade 7</th>
<th>Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Fitness(^d)</td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
</tr>
<tr>
<td>&gt;20% decrease</td>
<td>21471</td>
<td>11.06(11.76)</td>
<td>23035</td>
</tr>
<tr>
<td>10-20% decrease</td>
<td>11017</td>
<td>10.60(11.20)</td>
<td>16023</td>
</tr>
<tr>
<td>&lt;10% change</td>
<td>32325</td>
<td>10.43(11.07)</td>
<td>51272</td>
</tr>
<tr>
<td>10-20% increase</td>
<td>12174</td>
<td>10.08(10.65)</td>
<td>16401</td>
</tr>
<tr>
<td>&gt;20% increase</td>
<td>26588</td>
<td>9.97(10.39)</td>
<td>22949</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Boys (n=172026)</th>
<th>Grade 6</th>
<th>Grade 7</th>
<th>Grade 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Fitness</td>
<td>n</td>
<td>Mean (SD)</td>
<td>n</td>
</tr>
<tr>
<td>&gt;20% decrease</td>
<td>22053</td>
<td>10.23(10.78)</td>
<td>23640</td>
</tr>
<tr>
<td>10-20% decrease</td>
<td>10446</td>
<td>9.75(10.57)</td>
<td>15492</td>
</tr>
<tr>
<td>&lt;10% change</td>
<td>29804</td>
<td>9.55(10.66)</td>
<td>50242</td>
</tr>
<tr>
<td>10-20% increase</td>
<td>11509</td>
<td>9.32(10.06)</td>
<td>15536</td>
</tr>
<tr>
<td>&gt;20% increase</td>
<td>27202</td>
<td>9.14(9.86)</td>
<td>22613</td>
</tr>
</tbody>
</table>

\(^{a}\)Days absent per year. \(^{b}\)Observations account for 1-3 years of fitness-change per student. \(^{c}\)Tabulated mean estimates. \(^{d}\)Change in fitness composite percentile scores from the year prior.
Figure 3.3. Between- and within school-level variances in attendance in empty models and with fitness as the predictor

<table>
<thead>
<tr>
<th>Attendance^c</th>
<th>All students 3-level Empty Model^a (Model 1)</th>
<th>All students 3-level Fitness Added Model^b (Model 2)</th>
<th>Girls 3-level Empty Model^a (Model 3)</th>
<th>Girls 3-level Fitness Added Model^b (Model 4)</th>
<th>Boys 3-level Empty Model^a (Model 5)</th>
<th>Boys 3-level Fitness Added Model^b (Model 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between (schools)</td>
<td>13.24</td>
<td>12.89</td>
<td>15.76</td>
<td>15.33</td>
<td>14.41</td>
<td>14.08</td>
</tr>
<tr>
<td>Between (students)</td>
<td>89.96</td>
<td>90.47</td>
<td>95.17</td>
<td>95.57</td>
<td>85.44</td>
<td>86.05</td>
</tr>
<tr>
<td>Within (residual)</td>
<td>39.93</td>
<td>38.54</td>
<td>41.42</td>
<td>40.22</td>
<td>38.77</td>
<td>37.16</td>
</tr>
<tr>
<td>ICC^d (proportion of variance at the school level)</td>
<td>9.25%</td>
<td>9.08%</td>
<td>10.34%</td>
<td>10.14%</td>
<td>10.39%</td>
<td>10.26%</td>
</tr>
<tr>
<td>ICC (proportion of variance between students)</td>
<td>62.85%</td>
<td>63.76%</td>
<td>62.47%</td>
<td>76.35%</td>
<td>61.64%</td>
<td>62.68%</td>
</tr>
<tr>
<td>ICC (proportion of variance within students)</td>
<td>27.90%</td>
<td>27.16%</td>
<td>27.19%</td>
<td>26.61%</td>
<td>27.97%</td>
<td>27.07%</td>
</tr>
<tr>
<td>Intercept</td>
<td>11.93</td>
<td>14.07</td>
<td>12.20</td>
<td>14.27</td>
<td>11.63</td>
<td>13.80</td>
</tr>
</tbody>
</table>

^a Random intercept models with no predictors. ^b Change in fitness composite percentile scores from the year prior. ^c Days absent per year. ^d p<.0001 for all estimates.
3.3.3. Longitudinal individual-level and school-level clustering of attendance

The overall mean attendance rates across all schools (n=624) was 11.93 days per year (model 1 intercept). Table 3.3 shows variance in attendance at the student and school levels in model 1 (random intercept model with no predictors) and model 2 (including the predictor, change in health-related fitness from the year prior). Between-student variation in attendance was found to be much higher compared with both within-student and between-school variation in attendance for both models. ICC estimates, however, demonstrated a large degree of clustering at the school level (9% for both models). The percent of between-school, between-student, and within-student explainable variance in attendance accounted for by fitness was 3% ([13.24-12.89]/13.24), 1% ([89.96-90.47]/89.96), and 3% ([39.93-38.54]/39.93), respectively. Also, the percent of variability in attendance attributable to fitness across schools was 2.7% and 2.3% for girls and boys, respectively, 0.42% and 0.71% across students within schools for girls and boys, respectively, 2.9% and 4.2% across observations within student for girls and boys, respectively (Table 3.3).

3.3.4. Longitudinal causal effects of fitness-change on attendance

Results from model 2 showed all levels of change in fitness were significantly associated with subsequent attendance (p<.001). Compared to the reference category (decrease >20%), beta estimates were -.64 (95% CI: -0.70, -0.58), -.53 (95% CI: -0.61, -0.46), -0.34 (95% CI: -0.40, -0.28) and -0.22 (95% CI: -0.30, -0.149) for those who had a >20% increase, 10-20% increase, <10% change, and 10-20% decrease in fitness composite percentile scores from the year prior, respectively (Table 3.4).
After adjusting for covariates (gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), starting fitness, and school-area poverty, and including interactions (Grade*Ethnicity, Grade*Place of Birth, Grade*Starting Fitness, and School-Area Poverty*Ethnicity), beta estimates for the effects of fitness-change on days absent diminished but remained significant (p<.005). Relative to the reference category (decrease >20%), beta estimates were -.64 (95%CI: -.70,-.57), -.54 (95%CI: -.61,-.46), -.34 (95% CI: -.40,-.28), and -.23 (95%CI: -.30,-.15) days absent for students who had a >20% increase, 10-20% increase, <10% change, and 10-20% decrease in fitness composite percentile scores from the year prior (model 3, Table 3.4).

3.3.5. **Longitudinal causal effects of fitness-change on attendance by gender**

Results from mixed models stratified by gender showed slightly larger improvements in attendance with increased fitness in girls compared with boys. Girls with a large increase in fitness (>20%) demonstrated 0.66 fewer days absent per year (95%CI: 0.56, 0.75) compared with boys who demonstrated 0.64 fewer days absent per year (95%CI: 0.55, 0.72) relative to the

### Table 3.4. Longitudinal causal effects of fitness-change and attendance in New York City public school students in grades 6-8 (n=349 381 students; 624 schools)

<table>
<thead>
<tr>
<th>Change in Fitness</th>
<th>Unadjusted (Model 2)</th>
<th>Adjusted (Model 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β&lt;sup&gt;a&lt;/sup&gt;</td>
<td>95%CI lower</td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>-1.145</td>
<td>-1.262</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>-1.124</td>
<td>-1.250</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>-0.777</td>
<td>-0.875</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>-0.455</td>
<td>-0.581</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

<sup>a</sup> All beta estimates p<.0001;  
<sup>b</sup> Change in fitness composite percentile scores based on PACER (Progressive Aerobic Cardiovascular Endurance Run) Push-up and Curl-up Fitnessgram tests from the year prior.  
<sup>c</sup> Adjusted for gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), starting fitness, and school-area poverty, and including interactions (Grade*Ethnicity, Grade*Place of Birth, Grade*Starting Fitness, and School-Area Poverty*Ethnicity.)
reference group (>20% decrease in fitness composite percentile scores from the year prior; Table 3.5, models 4 and 5). School-area poverty was the only covariate shown not to be significant in males (p=.346), but was significant in females (p<.001). Moreover, the difference in days absent for the most compared with least improved fitness groups in girls was greatest for those attending schools in high- and very high- compared with mid- and low-poverty areas (0.567 vs. 0.249, respectively).

3.3.6 Regression diagnostics and sensitivity analysis results
Normal probability (Q-Q) plots demonstrated the outcome variable (days absent) was slightly right tail-distributed. Cook’s distance plots revealed extreme values for the main predictor variable with very low estimates (range: .000003199-.00042393), therefore these data were retained in analyses. In addition, in the plot of level-1 residuals by predicted values, there was relatively good agreement between the observed and predicted values with the exception of some outliers (retained in the models).

| Table 3.5. Longitudinal causal effects of fitness-change on attendance by gender in New York City public school students in grades 6-8^d |
|---|---|---|---|---|---|---|---|
| | Females (Model 4) | | Males (Model 5) | |
| | n=177355 students; 559 schools | | n=172026 students; 552 schools | |
| Predictors | β^a | 95%CI | β | 95%CI |
| Change in Fitness^b | lower | upper | lower | upper |
| >20% Increase | -0.655 | -0.747 | -0.564 | -0.636 | -0.724 | -0.547 |
| 10-20% Increase | -0.559 | -0.665 | -0.452 | -0.521 | -0.625 | -0.417 |
| <10% Change | -0.345 | -0.430 | -0.259 | -0.344 | -0.425 | -0.262 |
| 10-20% Decrease | -0.193 | -0.303 | -0.084 | -0.262 | -0.368 | -0.156 |
To determine whether enrollment exclusions impacted findings, sensitivity analyses were performed. When inclusion criteria were widened to include all students who were excluded from the analysis due to having fewer than 287 days enrolled, estimates remained significant,
though more conservative as predicted (p<.001; B=-.60, -.54, -.35, and -.19 days absent for >20% increase, 10-20% increase, <10% change, and 10-20% decrease) relative to the reference group (>20% decrease in fitness, respectively). Further, a sensitivity analysis was run excluding individuals with extreme observations for days absent from the final models. These models also showed slightly more conservative estimates for the magnitude of effects of fitness on attendance, although the dose response relationship remained consistent (p<.001; B=-.52, -.42, -.26, and -.16 days absent for >20% increase, 10-20% increase, <10% change, and 10-20% decrease) relative to the reference group (>20% decrease in fitness, respectively).

Sensitivity analyses were also performed to determine the impact of the specification of the BMI variable on findings. This paper subscribed to the literature which suggests that BMI in children should be used to track changes in weight status (underweight, normal weight, overweight, and obese), and also as a screening tool to identify those youth with potential weight problems. Changes across weight categories are found to be indicative of risk for obesity-related health problems in children and were therefore thought to be the best way to control for changes in BMI in this analysis. However, the distribution of the change in obesity status variable showed most students clustering in one value (consistently not obese). Moreover, the influence of changes in BMI percentile on the fitness-attendance association may be better captured by categorizing BMI percentiles into deciles or quartiles. To ensure that the BMI variable specification used here did not influence findings, sensitivity analyses involving re-fitting models with change in BMI percentiles categorized into quartiles and deciles showed that findings remained significant. Moreover, the magnitude of effects of the fitness-attendance relationship were almost identical across models using different BMI variable specifications.
(p<.001 all models; $B=-.64, -.53, -.33,$ and $-.23$ for most to least improved fitness, beta estimates identical to the second decimal for quartiles and deciles, vs. $B=-.64, -.53, -.34,$ and $-.22$ for change in obesity status model; p<.001 for all models).

3.4. Discussion

This study presents the first paper to the author’s knowledge that examines the causal effects of longitudinal change in fitness on attendance, overall and stratified by gender, and drawing from multiple years of multilevel data. Strengths of this study also include a large and diverse study sample of almost 350,000 students and comprising 6 cohorts followed for 4 consecutive years over a seven-year study period (2006/07-2012/2013).

Individual- and school-level clustering in attendance rates were found to be sizeable and statistically significant. These findings are not surprising given children are naturally clustered in schools. Youth typically spend many hours per day interacting with same-school peers and within the same classroom contexts. Students attending the same schools also share common teachers and administrations, and are furthermore exposed to common physical spaces both within and surrounding their schools. Similar to findings here, Bezold et al. detected large school clustering effects (ICC=25%) in their paper on fitness and academics.$^{52}$ Given ignoring clustering may produce underestimates of standard errors, and can lead to inaccurate beta coefficient estimates and increased Type I (false positive) errors,$^{52,53}$ this paper provides additional support for using multilevel methods to account for student nesting when studying the association of fitness and school-based outcomes.
This paper also found that 3% of the variability in attendance across schools (2.7% and 2.3% for girls and boys, respectively), 1% of the variability in attendance across students within schools (0.42% and 0.71% for girls and boys, respectively), and 3% of the variability in attendance across observations within students (2.9% and 4.2% for girls and boys, respectively), were attributable to fitness. While the estimates in variability in attendance across schools and students due to fitness are small, fitness is a modifiable factor and therefore may be targeted in school outcomes interventions. Given almost 1 million students attend NYC schools, targeting fitness in school settings holds great potential to change attendance at a population level.

It was also found that all levels of 1-year change in fitness were significantly associated with attendance (p<.001) in both crude and adjusted mixed models. Furthermore, consistent levels of fitness improvement each year at the >20% level (vs. >20% decrease) were found to have the potential to reduce a child’s number of days absent almost 2 days per year over the middle school period (i.e. an individual with mean days absent (10) would shift to having 8 days absent per year), and >4 days between 7-12th grades (i.e. from 10 to <6 days absent per year—a shift to regular attendance). In light of these findings and given the NYC DOE reports an average student attendance rate of 12 days absent per year, fitness interventions should be investigated as an effective approach to increase attendance at the population level. For example, perhaps children’s fitness interventions should target the school built environment, such as improving accessibility of building stairs, and expanding school recreation facilities. Quality physical education can also be integrated into programs before, during, and after school, including recess, extracurricular sports clubs, walk- and bike-to-school programs, and movement clubs. Physical activity can also be increased through establishing
activity breaks in the classroom, and programs to foster family engagement in regular physical activity.\textsuperscript{114-116} Lastly, communities can be engaged to hold intramural sports programs in parks and at school facilities outside of school hours.\textsuperscript{117} Importantly, research on the effects of fitness interventions would also inform population strategies for improving overall learning given attendance is shown to promote academic performance.\textsuperscript{55,99-102} Specifically, future research should examine the effectiveness of fitness interventions on both attendance and academic outcomes, as well as potential mediation by attendance in the fitness-academic performance association.\textsuperscript{10,34,76}

Results from mixed model analyses stratified by gender showed slightly larger increases in attendance with increased fitness in girls compared with boys. These findings are consistent with the fitness-attendance literature, which shows youth fitness may increase attendance,\textsuperscript{32-34} and also that low self-esteem in adolescent girls is significantly associated with both lower physical activity levels\textsuperscript{21} and attendance.\textsuperscript{25,26} While no prior papers were found which examined gender as an effect measure modifier in the fitness-attendance association, findings here are consistent with work documenting a stronger magnitude of effects in the fitness-educational outcomes literature in girls compared with boys.\textsuperscript{52} Based on findings here, if should also be examined whether youth fitness interventions should specifically target increased physical activity programming in adolescent girls, particularly in light of high absenteeism rates in girls shown in the descriptive analyses.

Lastly, gender differences across school-area poverty were found. More specifically, school-area poverty was a significant predictor in girls but not boys (p<.001; p=.346, respectively). In
fact, female gender and area poverty may interact to foster a combined negative effect on self-perception of athletic ability, which may be related to feelings of self-worth. Adolescent girls’ self-perception of athletic ability and fitness are shown to be associated with geographic poverty.\textsuperscript{103,118} For example, girls have lower self-perception of athletic competence compared with boys in more versus less impoverished neighborhoods.\textsuperscript{118} Similarly, low school-area recreational resource abundance has been shown to predict reduced physical activity patterns in middle-school girls vs. boys.\textsuperscript{103} Given the literature shows self-esteem predicts lower attendance,\textsuperscript{25,26} nuanced research on the combined effects of gender and area poverty on the fitness-attendance association are warranted.

3.4.1. \textit{Limitations}

Findings from this study may not be generalized to the entire NYC middle school population given not all students are not required to take the Fitnessgram. For example, students with <75% attendance rate or a long term absence, and those with particular medical conditions are exempt from taking the Fitnessgram. This analysis also may be limited in that a large number of students were not included due missing Fitnessgram tests for ≥2 consecutive years, insufficient enrollment period, or moving schools. It is likely that students who do not meet inclusion criteria, however, would be more likely to have lower attendance given psychosocial and family factors associated with moving and long-term absences. These effects potentially would move the association further from the null. In this sense, this study’s estimates would be conservative.

There also is potential for systematic bias and differential measurement error given the Fitnessgram dataset includes data not collected for research purposes. Data was not available on many student and school-level factors, including psychosocial measures, drug and alcohol
use, family structure, and individual household poverty. Potential sources of random measurement error and systematic bias include variation across Fitnessgram testing sites where school staff may vary in their testing protocol, may have permitted students to practice tests to a varied extent, or may have allowed students to perform tests in attire that impacted performance. However, prior research has not detected a learning effect when fitness tests are repeated.\textsuperscript{105} Moreover, despite variability across administrators at different schools, testing protocols are designed to promote consistency across administrators, including manuals, video-based training and site-visits, as well as use of calibrated scales.\textsuperscript{92} Further, Morrow et al.\textsuperscript{77} demonstrated reliability and validity of the Fitnessgram across testing sites.

Also of note, in this analysis, the effect of fitness change on lagged days absent was examined across gender strata by looking at whether beta estimates for fitness on attendance were different in models ran separately for boys compared with girls. However, it was not tested whether differences in beta estimates across gender strata were statistically significant. A formal effect measure modification by gender analysis may be more advantageous to formally assess whether the relationship between fitness change and lagged days absent is stronger for boys compared with girls. Formal effect measure modification could have been tested by including a cross-level interaction term in the model. This term ($\beta_{220}(\text{FITNESS})_{ij} \times (\text{GENDER})_{ij}$) would have assessed whether the strength of the relationship between two level-one variables (the main predictor, fitness-change, and the outcome, days absent) changed as a function of a level-two variable (gender). A significant interaction term in the fixed effects of the model output would indicate that the difference in the magnitude of effects of fitness-change on lagged days absent across gender is statistically significant.\textsuperscript{121} While this approach may be preferable,
including cross-level interaction terms in a 3-level longitudinal model adds a large degree of complexity. Beta estimates corresponding to the association of the interaction terms between the level-two variable gender and each level of the level-one time varying predictor (fitness-change) and the outcome (lagged days absent) are outputted and must be interpreted appropriately. Additional complexity is added with the possibility of including 3-level interactions in a 3-level longitudinal model \((\beta_{220}(\text{FITNESS})_{ij} \times (\text{GENDER})_{ij} \times (\text{SCHOOL-AREA POVERTY})_{ij})\), which would assess whether the strength of the relationship between two level-one variables (the main predictor, fitness-change, and the outcome, days absent) changed as a function of a level-two variable (gender) and a level-three variable (school-area poverty). In this sense, model interpretation using a stratified analysis is simpler and more intuitive, although a formal interaction analyses would be informative to better understand differences in the magnitude of effects of the fitness-attendance relationship across gender. Lastly, while this paper offers evidence in support of causal effects of fitness on attendance, the bi-directional causal relationship between exposure and outcome should be explored in more detail in future studies.

3.4.2. Conclusion

This paper presents evidence for inverse dose-response effects of fitness on attendance in both genders. A slightly stronger fitness-attendance effect was observed in girls, and youth attending schools in high-poverty areas. Given only stratified mixed models analyses are presented here, future analyses should examine fitness-change*gender interactions. Moreover, given preliminary findings from this study on gender differences in the effects of school-area poverty on the fitness-attendance association, further nuanced research is recommended on the potential
interactive effects of gender and area poverty on the fitness-attendance association. This research would inform school youth physical activity programming, particularly with respect to policy targeting adolescent girls in impoverished areas.
Chapter 4: Examining the causal longitudinal effects of fitness and chronic absenteeism in New York City 6th-8th grade youth

4.1. Background

Youth chronic absenteeism rates remain high. Nationally, 10-15% of (5-7.5 million) students are chronically absent, meaning they miss ≥10% of the school year (or ≥20 days of school per year). In NYC, approximately 20% are students are chronically absent (roughly 200,000 students). Chronic absenteeism rates increase with increasing student age, and are strongly associated with student race/ethnicity, and socioeconomic status. Moreover, chronic absenteeism is shown to reduce academic performance, and has long-term effects on graduation rates. Reducing chronic absenteeism may diminish racial/ethnic disparities in academic achievement. For example, Musser et al. found that moving from chronic absenteeism to average attendance was associated with a 17% and 26% decrease in the achievement gap between non-Hispanic white and minority 4th grade students on English and math standardized tests, respectively.

A growing body of literature in education over the last 5 years highlights efforts to reduce chronic absenteeism with school-based interventions. For example, in New York City (NYC) the Mayor’s Interagency Task Force on Truancy, Chronic Absenteeism and School Engagement has worked steadily with schools with above-average rates of chronic absenteeism to increase attendance through 1) prevention and intervention programs through “early warning” flags to identify students at risk of chronic absenteeism; 2) monitoring of student and school-level progress through accountability strategies and incentives; 3) “success mentors” to provide personalized support to students and families; 4) principal-led “student success meetings” to
foster data-driven planning by school-wide partners; 5) expanding strategies for collaborating with community partners; and 6) promoting awareness around chronic absenteeism among schools, families, and the public. Despite these and similar efforts, however, chronic absenteeism rates are not sufficiently reduced.

4.1.1. Fitness and attendance in youth

Recent literature has suggested health-related fitness (fitness) may increase attendance in youth, similar to findings on the association of fitness and reduced workplace attendance in adult populations. For example, cardiorespiratory fitness and physical activity in adults are shown to be positively associated with work attendance. Moreover, interventions targeting improvements in adult fitness have demonstrated an increase in work attendance. Given these findings, it is plausible that improvements in youth fitness may reduce chronic absenteeism. Hypothesized causal mechanisms include the contribution of fitness to global health improvements that advance adults along a wellness continuum. For example, it is shown that individuals with higher fitness have lower risk of cardiovascular disease, insulin sensitivity, hypertension, and metabolic syndrome, perhaps contributing to reduced absences related to illness. Research on the association of global health and school attendance in child populations has shown similar findings. Improvements in diet and physical activity may similarly reduce negative health effects and psychosocial problems associated with overweight and obesity.

To the author’s knowledge, only 1 longitudinal study and 4 cross-sectional studies have addressed the specific association between fitness and attendance in youth. These
studies draw predominantly from cross-sectional data, and also do not account for a wide range of potential confounders, address the influences of contextual effects, or investigate potential moderation by gender, despite findings suggesting that it may be important. No studies were identified which examined the longitudinal effects of fitness on chronic absenteeism in youth.

This paper examined the longitudinal causal effects of change in fitness on chronic absenteeism by drawing from 5 prospective cohorts of approximately 350,000 NYC middle-school students followed 4 years each (grades 5-8) over a seven-year study period (2006/7-2012/13). It was hypothesized that higher positive change in fitness (cardiorespiratory and muscular endurance, and muscular strength fitness composite percentile scores) would predict lower probability of 1-year lagged chronic absenteeism after accounting for potential individual- and school-level confounders, as well as for clustering by individual and school, and time-dependent interactions.

4.2. Methods

4.2.1. Data source, collection, management and study population

Data were drawn from the NYC Fitnessgram dataset jointly managed by NYC Department of Education (DOE) and Department of Health and Mental Hygiene (DOHMH), and comprised of annual fitness assessments collected by DOE for approximately 870,000 NYC public school students per year (grades K-12) starting in 2006-07. The Fitnessgram is demonstrated to have both strong reliability and validity.28,29 Individual student Fitnessgram data from multiple years are linked by a unique identifier. Detailed descriptions of the dataset, including sample selection criteria may be found in prior work (see Chapter 1).
Inclusion criteria for this study included active enrollment status in a NYC public school for the entire academic year (number of days enrolled for ≥ 287 days) for ≥2 consecutive years while in grades 6-8 during the study period (2006/7-2012/13) while attending a school in districts 1-32 (i.e. the schools that have Fitnessgram measurements; n=451,172). Students who attended schools with poor quality fitness data were excluded from the analysis (n=350). Students also were excluded from the analysis if they changed schools (to account for school clustering); n=44,977). In addition, students were excluded if they did not take the Fitnessgram for ≥2 consecutive years (n=56,464). After the above exclusions, the final sample of 6-8th graders was comprised of 349,381 unique students (51% female, 77% NYC born, 38% Hispanic, 28% Non-Hispanic black, and 17% non-Hispanic white; 177,281, 220,769, and 186,135 student-years contributed 6th, 7th and 8th grade data, respectively), and 624 schools (mean student population=541 students; SD=632).

4.2.2. Primary exposure

The primary exposure was a categorical variable representing age- and gender-specific percent change in a composite fitness score based on mean performance on aerobic capacity (Progressive Aerobic Cardiovascular Endurance Run (PACER)), muscle strength and endurance (curl-up and push-up) tests relative to peers (categorized as: >20% decrease, 10-20% decrease, <10% change, 10-20% increase, and >20% increase (reference category) from the year prior, consistent with prior longitudinal research on fitness and academic outcomes drawing from the NYC Fitnessgram52). The PACER comprises an aerobic capacity test for which individuals must run back and forth across a 20-meter space at a specified pace which increases
incrementally. The pushup and curlup (i.e. sit-up) also are set to a specified pace. For all three assessments, students are asked to complete as many repetitions as possible. Students pass the tests based on whether they achieve a score within age- and gender-specific Healthy Fitness Zones (Appendix C). Change in composite fitness percentiles was calculated based on fitness composite percentile scores collected the year prior to attendance data collection. Individual student fitness data from multiple years are linked in the dataset by a unique identifier.

4.2.3. Primary outcome

The outcome was chronic absenteeism (binary variable, missed ≥ 20 days of school per year) measured 1 year after the individual’s corresponding Fitnessgram assessments (e.g. change in fitness grades 6-7 and chronic absenteeism status at 7th grade year-end). Attendance data was also drawn from the NYC Fitnessgram dataset. Attendance information is collected at year-end and linked to Fitnessgram data by unique student identifiers. Student admission and discharge dates are included in the dataset if they occurred prior to or post-end of the school year.

4.2.4. Individual-level variables

Demographic variables included gender, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, Asian or Pacific Islander, Native American, and other (multiracial or parent refused), and place of birth (NYC, United States (not NYC), or foreign born). This information was based on DOE demographic surveys administered annually to parents and linked to Fitnessgram data by unique student identifiers.
Change in obesity status (obese to not obese, consistently not obese, consistently obese, not obese to obese) was also included (see Chapter 2 for detailed background on variable specification). Height and weight are collected on NYC students during routine physical education classes as part of the Fitnessgram assessment. Age- and gender-specific BMI for children are computed using the following formula:

\[ BMI = \frac{weight (kg)}{(height (m))^2} \].

Obesity status was defined according to CDC growth chart-derived norms for gender and age based on a historical reference population, and used to compute the BMI percentile for each child.\textsuperscript{28} Obesity was defined as having a BMI greater than or equal to the 95\textsuperscript{th} percentile for youth in the same gender and age in months group.\textsuperscript{85,86}

4.2.5. School-level variables

A categorical school-area poverty variable was based on percentage of households in the school zip code living below the federal poverty threshold (low (<10%), medium (10%-20%), high (20%-30%), and very high (>30%) area poverty) drawing from the American Community Survey (ACS) 2007-2013.

4.2.6. Statistical methods

Descriptive statistics were computed to summarize sample characteristics. Trends in chronic absenteeism (≥20 absent per year) by demographic and fitness-change characteristics were examined.

Next, cross-sectional, 2-level logistic mixed models with random intercepts were run by grade (6-8; models 1, 2, and 3, respectively). These models assessed the effects of change in fitness
composite percentile (categorized into >20% decrease, 10-20% decrease, and <10% change, 10-20% increase, and >20% increase (reference category) from the year prior) on 1-year lagged predicted probability of chronic absenteeism (≥20 days absent per year). These models were represented as: Logit(Chronic absenteeism)_{ij} = \beta_0 + \beta_1(FITNESS)_{ij} + \varepsilon_{ij} and were run for grades 6, 7 and 8 separately.

An unconditional longitudinal 3-level (observations nested in students, nested in schools) logistic Generalized Linear Mixed Model with a random intercept (model 4) was next run. This model was represented as: Logit(Chronic absenteeism)_{ij} = \beta_{00j} + \beta_{01}(Grade)_{ij}, where \beta_{00j} = \gamma_{00} + \mu_{0j}, and \mu_{0j} \sim N(0, \tau^2), where \tau^2 is the estimated between-school variance.

The Intraclass Correlation Coefficient (ICC) was calculated to determine the total variability in chronic absenteeism attributable to the school-level. The formula: ICC= \tau^2 / (\tau^2 + 3.29) was applied based on the assumption in logistic Generalized Linear Mixed Models that there is no error at level-1, therefore a modification of the calculation for the ICC is necessary. For this calculation, a dichotomous outcome is assumed to come from an unknown latent variable with a level-1 residual that follows a logistic distribution with a mean of 0 and a variance of 3.29. The estimate for the intercept in the unconditional model was used to determine the change in predicted probability (Pp) of chronic absenteeism: Pp= \frac{e^{\eta_{ij}}}{1 + e^{\eta_{ij}}}, where \eta_{ij}= Logit(Chronic absenteeism)_{ij}).

Next, the main exposure, categorical change in fitness was added to the longitudinal 3-level (observations nested in students, nested in schools) logistic Generalized Linear Mixed Model.
with a random intercept (model 5). This model was represented as: Logit(chronic absenteeism)\(_{ij}\) = \(\beta_{000} + \beta_{200}(FITNESS)_{ij} + \beta_{01}(Grade)_{ij}\). The ICC for this model was calculated as the difference between the unconditional between-school variance and conditional between-school variance, divided by the unconditional between-school variance. Change in ICC estimates represented the proportion of variance explained by including student-level change in fitness in the model. Change in within-student, between-student, and between-school variance estimates represented the proportion of explainable variation in attendance scores that was accounted for by fitness.

The final model (model 6) was adjusted for covariates (grade, time, gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), and school-area poverty (level 3), and including interactions (school-area poverty*ethnicity, and grade interactions with gender, place of birth, and race/ethnicity), represented as:

\[
\text{Logit(chronic absenteeism)}_{ij} = \beta_{000} + \beta_{100}(\text{GRADE})_{ij} + \beta_{200}(\text{FITNESS})_{ij} + \beta_{300}(\text{CHANGE IN OBESITY STATUS})_{ij} + \beta_{400}(\text{YEAR})_{ij} + \beta_{010}(\text{RACE/ETHNICITY})_{ij} + \beta_{020}(\text{GENDER})_{ij} + \beta_{030}(\text{PLACE OF BIRTH})_{ij} + \beta_{110}(\text{GRADE})_{ij} \ast (\text{RACE/ETHNICITY})_{ij} + \beta_{120}(\text{GRADE})_{ij} \ast (\text{GENDER})_{ij} + \beta_{130}(\text{GRADE})_{ij} \ast (\text{PLACE OF BIRTH})_{ij} + \beta_{011}(\text{SCHOOL-AREA POVERTY})_{j} \ast (\text{RACE/ETHNICITY})_{ij} + \beta_{001}(\text{SCHOOL-AREA POVERTY})_{j}
\]

All analyses were performed using \textit{PROC GLIMMIX} in SAS 9.4 software (Cary, NC).

4.3. Results

4.3.1. Sample characteristics

Table 4.1 presents demographic, fitness, and school characteristics of the study population. The population was comprised of slightly more females (n=177,355; 51%) and mostly Hispanic and
Non-Hispanic Black (n=134,453; 38% and n=99,363; 28%, respectively). Most students were English-speaking (n=197,727; 57%) and most were born in NYC (n=269,251; 77%).

Across all years and grades, 37% of students (n= 253,161) had <10% change in fitness scores, 20% (n=134,753) had a >20% increase in fitness scores, and 12% had a 10-20% increase (n=82,117; 12%) in fitness scores from the year prior. Nineteen percent of students (n=126,115) had a >20% decrease in fitness scores, and 12% had a 10-20% decrease (n=79,172; 12%) in fitness scores from the year prior. Also, most students were consistently not obese across all years (n=504,762; 73%), followed by consistently obese (n=119,235; 17%), obese to not obese (n= 36,029; 5%), and not obese to obese (n=27,273; 4%).

There were 624 schools included in the analysis. The average school size was 560 students (SD=713.34), including 365 (58%) small schools (<400 students per school). Twenty six percent (n=89,407) and 22% (n=78,510) of students attended a school in a high- or very high-poverty area.

<table>
<thead>
<tr>
<th>Table 4.1. Demographic and fitness-change characteristics of the study population (N=349,381)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
</tr>
<tr>
<td>American Indian</td>
</tr>
<tr>
<td>Asian and/or Pacific Islander</td>
</tr>
<tr>
<td>Race/Ethnicity</td>
</tr>
<tr>
<td>-------------------------------</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Language Spoken at Home</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>197727</td>
<td>57</td>
</tr>
<tr>
<td>Spanish</td>
<td>86052</td>
<td>25</td>
</tr>
<tr>
<td>Other language</td>
<td>65602</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Student Socioeconomic Status</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Area Poverty</td>
<td>62238</td>
<td>18</td>
</tr>
<tr>
<td>Medium Area Poverty</td>
<td>119219</td>
<td>34</td>
</tr>
<tr>
<td>High Area Poverty</td>
<td>89407</td>
<td>26</td>
</tr>
<tr>
<td>Very High Area Poverty</td>
<td>78510</td>
<td>22</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Place of Birth</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYC</td>
<td>269251</td>
<td>77</td>
</tr>
<tr>
<td>US (Not NYC)</td>
<td>19909</td>
<td>6</td>
</tr>
<tr>
<td>Foreign</td>
<td>60149</td>
<td>17</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Fitness (all years)</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;20% Increase</td>
<td>134753</td>
<td>20</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>82117</td>
<td>12</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>253161</td>
<td>37</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>79172</td>
<td>12</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>126115</td>
<td>19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Obesity Status (all years)</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obese to not obese</td>
<td>36029</td>
<td>5</td>
</tr>
<tr>
<td>Consistently not obese</td>
<td>504762</td>
<td>73</td>
</tr>
<tr>
<td>Consistently obese</td>
<td>119235</td>
<td>17</td>
</tr>
<tr>
<td>Not obese to obese</td>
<td>27273</td>
<td>4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Attendance</th>
<th>M</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 6</td>
<td>10.23</td>
<td>11.03</td>
</tr>
<tr>
<td>Grade 7</td>
<td>10.92</td>
<td>12.45</td>
</tr>
<tr>
<td>Grade 8</td>
<td>13.13</td>
<td>14.47</td>
</tr>
</tbody>
</table>

*a* Based on change in fitness composite percentile scores from the year prior. *b* Days absent per year. N_missing Place of Birth=72; N_missing Area Poverty=7; N_missing or > 1 race/ethnicity=177.
4.3.3. Chronic absenteeism prevalence rates by demographics and fitness

Overall, 15%, 16% and 20% of 6th, 7th and 8th grade students, respectively, were chronically absent (≥ 20 days absent per year; n=52,393, 55,115, and 55,857, respectively; Table 4.2). Chronic absenteeism prevalence rates were highest in students with a >20% decrease and 10-20% decrease in fitness from the year prior (18%; n=23,098 and 16%; n=12965, respectively). Chronic absenteeism also was highest in girls (18%; n=86,918), American Indian, Hispanic and Non-Hispanic Black students (21%; n=251, 19%; n=25,819 and 19%; n=18,507, respectively), and students whose primary home language was English (19%; n=36,574). Students who attended schools in very high poverty areas, and those who were born in NYC or the US (not NYC) had the highest chronic absenteeism rates (22%; n=17,283, 17%; n=44,668, and 17%; n=3,413, respectively).

<table>
<thead>
<tr>
<th>Table 4.2. Chronic absenteeism overall and across demographic and fitness-change characteristics (N=349381)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
</tr>
<tr>
<td><strong>Chronic Absenteeism (all years)</strong></td>
</tr>
<tr>
<td>Grade 6</td>
</tr>
<tr>
<td>Grade 7</td>
</tr>
<tr>
<td>Grade 8</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
</tr>
<tr>
<td>American Indian</td>
</tr>
<tr>
<td>Asian and/or Pacific Islander</td>
</tr>
<tr>
<td>Hispanic</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
</tr>
<tr>
<td>Language Spoken at Home</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td>English</td>
</tr>
<tr>
<td>Spanish</td>
</tr>
<tr>
<td>Other Language</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Place of Birth</th>
</tr>
</thead>
<tbody>
<tr>
<td>NYC</td>
</tr>
<tr>
<td>US (Not NYC)</td>
</tr>
<tr>
<td>Foreign</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Change in Fitness (all years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;20 Increase</td>
</tr>
<tr>
<td>10-20% Increase</td>
</tr>
<tr>
<td>&lt;10% Change</td>
</tr>
<tr>
<td>10-20 Decrease</td>
</tr>
<tr>
<td>&gt;20 Decrease</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School-Area Poverty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Area Poverty</td>
</tr>
<tr>
<td>Medium Area Poverty</td>
</tr>
<tr>
<td>High Area Poverty</td>
</tr>
<tr>
<td>Very High Area Poverty</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥20 days absent per year.</td>
</tr>
</tbody>
</table>
| *Based on change in fitness composite percentile scores from the year prior. N\text{missing} Place of Birth=69; N\text{missing Area Poverty}=7; N\text{missing or} >1 race/ethnicity=177.

4.3.4. *Cross-sectional fitness-change-chronic absenteeism association by grade*

Cross-sectional logistic models on the association of change in fitness composite percentile scores and 1-year lagged chronic absenteeism run by grade are shown in Table 4.3 and Figure 4.1 (models 1-3). Relative to the reference category (increase of >20% in fitness), all categories of fitness-change were significantly associated with chronic absenteeism (p<.001) across all grades. Also, the magnitude of effects of change in fitness on subsequent chronic absenteeism increased with increasing grade: predicted probability of chronic absenteeism was 14.65%
(95% CI: 11.68-17.43%), 20.98% (95% CI: 18.59-23.25%), and 23.89% (95% CI: 21.57-26.07%), for 6, 7 and 8th grades, respectively, for students who had a decrease >20% in fitness scores from the year prior relative to the reference group (increase >20% in fitness scores). Effects were also substantial for students with 10-20% decrease in fitness scores: predicted probability of chronic absenteeism was 11.22% (95% CI: 7.26-14.86%), 13.01% (95% CI: 9.71-16.09%), and 14.67% (95% CI: 11.43-17.68%), for 6, 7 and 8th grades, respectively, relative to the reference group (increase >20% in fitness scores).

Figure 4.1. Predicted probability of chronic absenteeism by fitness-change and grade

*All fitness-change categories relative to the reference group, >20% increase in fitness scores from the year prior; all categories p<.001 except increase 10-20%.
### Table 4.3. Predicted probability of chronic absenteeism in New York City public school students grades 6-8, by level of fitness-change from the previous year (N=349381)

<table>
<thead>
<tr>
<th>Change in Fitness</th>
<th>Grade 6 (Model 1)</th>
<th>Grade 7 (Model 2)</th>
<th>Grade 8 (Model 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pp(^b)</td>
<td>95% CI</td>
<td>Pp</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>14.65%(^*)</td>
<td>11.68</td>
<td>17.43%</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>11.22%(^*)</td>
<td>7.26</td>
<td>14.86%</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>10.88%(^*)</td>
<td>7.94</td>
<td>13.64%</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>4.27%</td>
<td>-0.24</td>
<td>8.39%</td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

*ICC 21.33% 18.35% 14.57%

\(^a\)Totals for grades reflect some student missing fitness data for some grades (1-year of change in fitness composite percentile scores from the year prior was required to meet inclusion criteria). \(^b\)P<.0001. \(^p\)=predicted probability.

### Table 4.4. Overall effects of fitness-change on predicted probability of chronic absenteeism in New York City public school students in grades 6-8 (n=349381 students; 624 schools)

<table>
<thead>
<tr>
<th>Fitness-change(^a)</th>
<th>Unadjusted (Model 5)</th>
<th>Adjusted (Model 6)(^b)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pp(^b)</td>
<td>95% CI</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td>20.26%(^*)</td>
<td>18.86%</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>14.70%(^*)</td>
<td>12.84%</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>11.67%(^*)</td>
<td>10.14%</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>3.26%(^A)</td>
<td>0.82%</td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>.</td>
<td>.</td>
</tr>
</tbody>
</table>

* ICC 17.20% 12.26%

\(^a\)Change in fitness composite percentile scores based on PACER (Progressive Aerobic Cardiovascular Endurance Run) Push-up and Curl-up Fitnessgram tests from the year prior. \(^b\)Pp=predicted probability.
4.3.5. *Longitudinal causal effects of fitness-change on chronic absenteeism*

The overall predicted probability of chronic absenteeism across all schools was 16.50% (n_{schools}=624, model 4 intercept (not shown)). ICC estimates for the unconditional (model 4, no predictors) and crude models (model 5; Table 4.4) showed a large degree of clustering at the school-level (18% and 12%, respectively). Including student-level change in fitness in the model reduced the proportion of variance explained by 36%, as indicated by the change in the school-level variance estimates for models 4 and 5.

Greater decreases in fitness predicted higher probability of chronic absenteeism the following year (Pp=20.26% (95%CI: 18.86, 22.16), Pp=14.70% (95%CI: 12.84, 16.47), Pp=11.67% (95%CI: 10.14, 13.15), and Pp=3.26% (95%CI: 0.82, 5.59) for >20% decrease, 10-20% decrease, <10% change, and 10-20% increase, respectively) relative to the reference group (>20% increase in fitness, Table 4.4, model 5).

After adjusting for covariates (grade, time, gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), and school-area poverty (level 3)), and including interactions (school-area poverty*ethnicity, and grade interactions with gender, place of birth, and race/ethnicity; model 6), all estimates for the association of change in fitness composite percentile and 1-year lagged chronic absenteeism diminished but remained significant (p<.001). Higher positive change in fitness still predicted lower probability of chronic absenteeism the following year. The predicted probability of chronic absenteeism for >20 decrease in fitness was 19.49% (95%CI: 18.05, 20.88) relative to the reference group (>20% increase in fitness scores). The predicted probability for 10-20% decrease, <10%
change, and 10-20% increase in fitness scores was 12.58% (95%CI: 10.60, 14.46), 8.07% (95%CI: 6.38, 9.71), and -1.04% (95%CI: -1.54, -3.49), respectively, relative to the reference group (>20% increase in fitness; Table 4.4).

4.4. Discussion

This study presents the first paper to the author’s knowledge that prospectively examines the association of change in fitness and chronic absenteeism over multiple years and taking into account both student- and school-level factors. Chronic absenteeism prevalence rates increased with increasing grade. Also, predicted probability of chronic absenteeism by grade was highest in students with the least to most improved fitness from the year prior based on cross-sectional mixed models. Three-level logistic mixed models demonstrated an inverse dose-response relationship between fitness-change and time-lagged chronic absenteeism. Indeed, greater decreases in fitness predicted increased probability of chronic absenteeism the following year. Although no papers were identified by the author which addressed the specific association of fitness and chronic absenteeism, findings here are consistent with the literature on fitness and attendance, and a more general literature on fitness and academic outcomes demonstrating a positive effects of fitness on academics.3,10-12

A large body of work has in fact demonstrated a positive association between children’s fitness and academic outcomes.3,10-12 The Center for Disease Control’s (CDC) review from 2010, for example, on 50 studies examining the association between school-based physical activity and several indicators of academic performance (course grades, academic behavior, cognitive skills and attitudes), found positive associations in more than half of the studies reviewed, and
nonsignificant and negative associations in 48% and 1.5%, respectively (effect sizes not shown). Similarly, a meta-analysis of 59 studies from 1947-2009 concluded that school-based physical activity programs in children have a significant positive association with cognitive performance based on school grades or standardized test performance, with a greater effect size for experimental and quasi-experimental in comparison to observational designs (mean effect size=0.35, 95%CI: 0.27-0.43 vs. 0.32, 95%CI: 0.26-0.37, for experimental and observational studies, respectively). 

In addition to the above review and meta-analysis, six studies on children’s physical activity and academic performance show a significant positive association. For example, Mohar demonstrated a significant inverse association between MVPA (based on accelerometry) and grade point average (M=2.7(SD=0.03), M=3.1(SD=0.04) and M=3.1 (SD=0.4), for the lowest, middle and highest physical activity tertiles, respectively) in a cross-sectional study of primary and middle school children (n=324), and adjusting for student age, gender, grade level, BMI, and time of day of physical activity. Moreover, Welk et al. using cross-sectional data (n=38,992) found a difference of 10% in fitness attainment was associated with a 0.36% higher value in standardized test performance after controlling for school-level SES, minority status, school size, gender. Similarly, London et al. in a prospective study of primary and middle school children (n=2,735) followed over four years demonstrated that fitness composite scores (aerobic capacity, abdominal strength and endurance, trunk extensor strength and endurance, upper body strength and endurance, and flexibility) significantly predicted standardized test performance. More specifically, middle school students who were consistently fit had higher
math and English test scores compared to those who were consistently unfit 
($\beta \pm SE = 0.218 \pm 0.058$ and $0.154 \pm 0.049$, respectively).  

Given this study’s findings, it is plausible that children’s physical activity may predict academic 
performance through a pathway involving improved fitness, followed by improved attendance, 
and resulting in higher academic performance. Attendance in fact is well documented in the 
literature to predict academic performance. For example, a one standard deviation 
increase in the days a student is absent is shown to be associated with a statistically significant 
0.45 and 0.39 standard deviation change in school grade point average for elementary and 
middle school students, respectively ($p < .01$). More work must therefore draw from methods 
which address the causal pathway linking fitness to academic performance. Indeed, only three 
papers have postulated (but did not test for) the role of attendance as a mediator linking 
children’s fitness to academic performance, although this framework is plausible given 
research into adult fitness and work absenteeism. 

If further research shows that fitness is causally related to attendance, public health 
interventions should promote opportunities for youth physical activity, including ≥60 minutes of 
physical activity per day for 6-17 year olds. Opportunities for expanding youth engagement in 
physical activity include offering quality physical education before, during, and after school, 
including recess, extracurricular sports clubs, walk- and bike-to-school programs, and physical 
activity breaks in the classroom. Additional ideas to promote youth physical activity include 
the creation of community programs to increase family engagement in regular physical 
activity.
Strengths of this study include drawing from repeated measures and multilevel data, and comprising a large and diverse study sample of almost 350,000 individuals comprising 6 cohorts of students followed for 4 consecutive years during a seven-year study period (2006/07-2012/2013). This is also the first paper to the author’s knowledge that examined the specific fitness-lagged chronic absenteeism association. Given temporality of exposure and outcome are known, this paper lends support in favor of a causal relationship between fitness-change and chronic absenteeism.

4.5.1. Limitations

Limitations for this analysis include a potential for sample selection bias due to a large number of students given student inclusion requirements included for ≥2 consecutive years of Fitnessgram data measured at the same school to be included in the analytic population. However, it is likely that students who are missing Fitnessgram data (such as due to moving schools frequently, or long-term absences, leading to school absences on the days Fitnessgram tests were administered) would be more likely to have lower attendance given mental health and familial factors associated with moving schools and inconsistent school enrollment potentially leading to effect sizes further from the null. It is suspected therefore that exclusions based on enrollment or missing Fitnessgram data likely led to conservative effect estimates.

Given the Fitnessgram dataset is not collected for research purposes, there is additional potential for systematic bias in that participation in the Fitnessgram is not randomized. For example, schools are required to collect Fitnessgram assessments on ≥ 85% of eligible students. Not all
students, all years, receive all components of the Fitnessgram. Given this study aimed to assess the longitudinal effects of fitness on chronic absenteeism, additional sensitivity analyses were conducted on a smaller cohort of students for whom data was available for the entire study period \( n_{\text{sensitivity}} = 97,476 \) students and 493 schools; i.e. 4 consecutive years of fitness data and consistent enrollment at the same school). Although estimates based on this small cohort weakened in magnitude, they remained substantial and significant \( p < .001 \).

Lastly, while this paper offers evidence in support of a causal association between fitness-change and chronic absenteeism, there is the potential for a bi-directional relationship between exposure and outcome. For example, it is possible that children who have higher chronic absenteeism are more sedentary, particularly if they are home resting, ill, or occupied in nonactive ways (e.g. video-game playing, watching television, etc.). In this sense, although chronic absenteeism was lagged to fitness in this analysis, the temporality of exposure and outcome could be reversed. In other words, chronic absenteeism may precede fitness level. Future research should explore the directionality of fitness and attendance in more detail.

### 4.5.2. Conclusion

This study demonstrated an inverse dose-response relationship between fitness and 1-year lagged probability of chronic absenteeism in a large and diverse study sample, and drawing from multi-level, repeated measures logistic mixed models. Based on this study’s findings, school-based fitness programs should be examined as a population health strategy for reducing youth chronic absenteeism rates. For example, NYC DOE, the largest school district in the US, serves approximately 1.1 million students. Given children spend over 6 hours a day in school,
systematic efforts to foster positive attitudes towards fitness, and promote lifelong physical activity and exercise skills should be considered standard elements of school-based curriculums. Indeed, the CDC presents schools as an ideal population-wide setting for promoting physical activity among US youth. This study’s findings may therefore inform policy targeting reduced chronic absenteeism with fitness interventions, including school curricula, after school programs, and the school built environment. Future research should examine the potential for school-based physical activity interventions to reduce chronic absenteeism rates.
Chapter 5: Discussion

5.1. Overview of the Dissertation

This dissertation aimed to describe differences in student- and school-level attendance, examine the causal association of health-related fitness (fitness) and attendance (days absent and chronic absenteeism), and assess gender effect measure modification in the fitness-attendance association in a large and diverse sample of NYC middle school students. Chapter 2 described cross-sectional trends in attendance and fitness at the individual student- and school- levels, by grade and school-area poverty. Chapter 2 also drew from mixed modeling techniques to present the variation in the attendance accounted for at the student and school levels. Chapter 3 used 3-level repeated measures linear models to examine the causal longitudinal association of change in fitness composite percentile scores and 1-year lagged days absent from school per year, while accounting for important covariates at the student- and school-levels, and time-dependent interactions. Chapter 3 also investigated effect measure modification by gender on the causal longitudinal fitness-attendance association. Chapter 4 explored the casual longitudinal association of change in fitness and 1-year lagged chronic absenteeism drawing from 3-level repeated measures logistic models.

5.2. Summary of Findings

5.2.1. Chapter 2

Chapter 2 aimed to characterize individual-level and between-school variation in fitness and attendance in middle school students using the NYC Fitnessgram dataset (2006/7-2012/13). It was hypothesized that attendance would increase with increasing fitness levels, decreasing
grade levels, and decreasing school-area poverty levels. It was also hypothesized that clustering students in schools would account for a small but significant proportion of total variability in attendance levels.

Attendance was found to decrease (increasing days absent) with decreasing student fitness levels based on student-specific data and increasing grade level. For example, mean days absent for students with the greatest increase in fitness (>20% positive change in composite percentile fitness scores) was 9.56 (SD=10.11), 9.85 (SD=10.81) and 11.87 (SD=12.73), for students in 6, 7 and 8th grades, respectively. In contrast, mean days absent for students with the greatest decline in fitness (>20% negative change in fitness) was 10.62 (SD=11.27), 11.57 (SD=12.62), and 13.87 (SD=14.32), for students in 6, 7 and 8th grades, respectively.

Attendance also was found to decrease with decreasing fitness and increasing school-area poverty. For example, students with the greatest decrease (>20%) in fitness who attended schools in the lowest compared with highest area poverty had a mean attendance of 10.11 (SD=10.52) vs. 14.04 (SD=14.74) days absent per year, respectively. In contrast, students with the greatest increase (>20%) in fitness who attended schools in the lowest compared with highest area poverty had a mean attendance of 8.63 (SD=9.07) vs. 12.44(SD=13.07) days absent per year, respectively.

Based on the unconditional 2-level mixed model (students nested in schools), the overall mean days absent across all schools (n_schoo=624) was 11.85 days per year. ICC estimates for the empty model demonstrated a large degree of clustering at the school level (11%). Given
attendance reporting is typically aggregated at the school-level,\textsuperscript{32,73,74,89-91} these rates may not fully capture the extent of school absenteeism, nor accurately present nuanced patterns of attendance across students compared with figures based on student-level data. Moreover, findings here suggest attendance is shaped not only by student-specific factors, but also the characteristics of the schools they attend.

5.2.2. Chapter 3

The purpose of Chapter 3 was to analyze the causal effects of change in health-related fitness on subsequent attendance in 6 cohorts of NYC Department of Education (DOE) middle school students followed consecutively over 4 years during a seven-year study period (2006/7-2012/13). It was hypothesized that change in individual-level fitness (cardiorespiratory, muscular endurance, and muscular strength fitness composite percentile scores) from the year prior would positively predict change in individual-level days absent per year after accounting for potential individual- and school-level confounders, as well as accounting for clustering by individual and school, and time-dependent interactions. It was also hypothesized that gender would modify the relationship between change in fitness and 1-year lagged change in attendance, and further that fitness would be a stronger predictor of attendance in females compared with males.

Based on 3-level repeated measures linear models, the between-student variation in attendance was found to be much higher than the between-school variation in attendance for both models 1 and 2. ICC estimates, however, demonstrated a large degree of clustering at the school level (9% for both models). This analysis also found that 3% of the variability in attendance across
schools and within-students, and 1% of the variability in attendance across students were attributable to fitness. Also, the percent of variability in attendance attributable to fitness across schools was 2.7% and 2.3% for girls and boys, respectively, 0.42% and 0.71% across students within schools for girls and boys, respectively, 2.9% and 4.2% across observations within student for girls and boys, respectively. While the estimates in variability in attendance across schools and students due to fitness are small, fitness is a modifiable factor and may hold great potential for population-level interventions targeting school settings.

All levels of 1-year change in fitness were significantly associated with attendance (p<.001) in both crude and adjusted mixed models. Furthermore, after adjusting for covariates (gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), starting fitness, and school-area poverty, and including interactions (Grade*Ethnicity, Grade*Place of Birth, Grade*Starting Fitness, and School-Area Poverty*Ethnicity), beta estimates for the effects of fitness-change on days absent diminished but remained significant (p<.001). Relative to the reference category (decrease >20%), beta estimates were -0.64 (95%CI: -0.70, -0.57), -0.54 (95%CI: -0.61, -0.46), -0.34 (95% CI: -0.40, -0.28), and -0.23 (95%CI: -0.30, -0.15) days absent for those who had a >20% increase, 10-20% increase, <10% increase or decrease, and 10-20% decrease in fitness composite percentile scores from the year prior. These findings indicate that consistent levels of fitness improvement each year at the >20% level (vs. >20% decrease) were found to have the potential to reduce a child’s number of days absent almost 2 days per year over the middle school period (i.e. an individual with mean days absent (10) would shift to having 8 days absent per year), and >4 days between 7-12th grades (i.e. from 10 to <6 days absent per year—a shift to regular attendance). In light of these
results and given the NYC DOE reports an average student attendance rate of 12 days absent per year\textsuperscript{89,93,94} it should be investigated whether fitness interventions may be an effective approach to increase attendance at the population level.

Mixed models stratified by gender showed slightly larger improvements in attendance with increased fitness in girls compared with boys. Girls with a large increase in fitness scores (>20\% increase) from the year prior demonstrated 0.66 fewer days absent per year (95\%CI: 0.56, 0.75) compared with boys who demonstrated 0.64 fewer days absent per year (95\%CI: 0.55, 0.72) relative to the reference group (>20\% decrease in fitness from the year prior).

School-area poverty was the only covariate shown not to be significant in males (p=.346), but was significant in females (p<.001) Moreover, the difference in days absent for the most compared with least improved fitness groups in girls was greatest for those attending schools in high- and very high- compared with mid- and low-poverty areas (0.567 vs. 0.249, respectively). These findings suggest that physical activity interventions targeting student attendance may be most effective for girls who attend schools in high-poverty areas.

5.2.3. Chapter 4

Chapter 4 examined the causal association of change in fitness and chronic absenteeism by drawing from 5 prospective cohorts of approximately 350,000 NYC middle-school students followed 4 years each (grades 5-8) over a seven-year study period (2006/7-2012/13). It was hypothesized that higher positive change in fitness (cardiorespiratory and muscular endurance, and muscular strength fitness composite percentile scores) would predict lower probability of 1-year lagged chronic absenteeism after accounting for potential individual- and school-level
confounders, as well as accounting for clustering by individual and school, and time-dependent interactions.

Cross-sectional logistic models on the association of change in fitness composite percentile scores and 1-year lagged chronic absenteeism run by grade (>20% decrease, 10-20% decrease, <10% change, 10-20% increase, and >20% increase (reference category)) showed all categories of fitness-change were significantly associated with chronic absenteeism (p<.001) across all grades. Also, the magnitude of effects of change in fitness on subsequent chronic absenteeism increased with increasing grade: predicted probability of chronic absenteeism was 14.65% (95%CI: 11.68-17.43%), 20.98% (95%CI: 18.59-23.25%), and 23.89% (95%CI: 2157-26.07%), for 6, 7 and 8th grades, respectively, for students who had a decrease >20% in fitness composite percentile from the year prior relative to the reference group (increase >20% in fitness).

The overall predicted probability of chronic absenteeism across all schools was 16.50% based on the empty 3-level model intercept (not shown). ICC estimates for the unconditional and crude models showed a large degree of clustering at the school level (18% and 12%, respectively).

When fitness was added to the unconditional model (crude analysis), greater decreases in student-level fitness predicted higher probability of chronic absenteeism the following year (predicted probability=20.26% (95%CI: 18.86, 22.16), 14.70% (95%CI: 12.84, 16.47), 11.67% (95%CI:10.14, 13.15), and 3.26% (95%CI: 0.82, 5.59) for >20% decrease, 10-20% decrease, <10% increase or decrease, and 10-20% increase, respectively) relative to the reference group.
 (>20% increase in fitness). After adjusting for covariates (grade, time, gender, race/ethnicity, change in obesity status from the year prior, place of birth (US (not NYC), NYC, or foreign), and school-area poverty (level 3)), and including interactions (school-area poverty*ethnicity, and grade interactions with gender, place of birth, and race/ethnicity), all estimates for the association of change in student fitness composite percentile and 1-year lagged chronic absenteeism diminished but remained significant (p<.001). Higher positive change in student-level fitness still predicted lower probability of chronic absenteeism the following year (19.49% (95%CI: 18.05, 20.88), 12.58% (95%CI: 10.60, 14.46), 8.07% (95%CI: 6.38, 9.71), and -1.04% (95%CI: -1.54, -3.49), for >20% decrease, 10-20% decrease, <10% increase or decrease, and 10-20% increase, respectively) relative to the reference group (>20% increase in fitness). In sum, this study demonstrated an inverse dose-response relationship between fitness and 1-year lagged probability of chronic absenteeism in a large and diverse study sample, and drawing from multi-level, repeated measures logistic mixed models. Based on this study’s findings, school-based fitness programs should be examined as a population health strategy for reducing youth chronic absenteeism rates.

5.2.4. Overall findings

To summarize, school-level clustering in attendance was found to be sizable and statistically significant across all analyses. School-area poverty accounted for a large proportion of the variance in student attendance at the school-level in Chapter 2, and effects of fitness on attendance were more pronounced in schools situated in very high- and high- compared with mid- and low-poverty areas. Also, there was in inverse dose-response relationship found between fitness-change and lagged days absent for both sexes, with slightly stronger effects in
girls compared with boys. In addition, an inverse dose-response relationship was found between fitness-change and lagged chronic absenteeism. Collectively, these findings contribute to a literature which suggests that schools may be a unique setting shape youth attitudes and behaviors towards routine physical activity. US schools serve approximately 56 million youth, and children spend at minimum 6 hours per day in school settings. However, only 10% of middle schools in 2006 provided daily physical education to its students in all grades, and only 57% of school districts required regularly scheduled recess for elementary school students. In this sense, findings from this dissertation support CDC’s call for policies which promote and sustain school-based physical activity programs for youth. Future work should address the mechanistic pathways in the fitness-attendance association, and focus on gender and area-poverty interactions with fitness to better inform school-based fitness programs and policy targeting all youth, and adolescent girls attending schools in high-poverty areas in particular.

5.3. Limitations

These studies had several limitations. To begin, a large number of students were not included due to insufficient period of school enrollment, moving schools or missing Fitnessgram tests over for ≥2 consecutive years. To determine whether these exclusions impacted findings, sensitivity analyses were performed. When inclusion criteria were widened to include students who were excluded from the analyses due to insufficient period of enrollment, estimates remained significant, though slightly more conservative (p<.001; B=-.606, -.543, -.352, and -.199 days absent for >20% increase, 10-20% increase, <10% increase or decrease, and 10-20% decrease, respectively) relative to the reference group ( >20% decrease in fitness).
There is also a potential for sample selection bias due to a large number of students being excluded from analyses based on missing $\geq 2$ consecutive years of Fitnessgram data measured at the same school. However, it is likely that students who are missing Fitnessgram data (such as due to moving schools frequently, or long-term absences, leading to school absences on the days Fitnessgram tests were administered) would be more likely to have lower attendance given mental health and familial factors associated with moving schools and inconsistent school enrollment potentially leading to effect sizes further from the null. It is therefore suspected that study exclusions based on missing Fitnessgram data likely led to conservative effect estimates.

Given the Fitnessgram dataset is not collected for research purposes, there is additional potential for systematic bias in that participation in the Fitnessgram is not randomized. For example, schools are required to collect Fitnessgram assessments on $\geq 85\%$ of eligible students. Not all students, all years, receive all components of the Fitnessgram. Given this study aimed to assess the causal longitudinal association of fitness and attendance, additional sensitivity analyses were conducted on a smaller cohort of students for whom data was available for the entire study period ($n_{\text{sensitivity}}=97,476$ students and 493 schools; i.e. 4 consecutive years of fitness data and consistent enrollment at the same school). Although estimates based on this small cohort weakened in magnitude, they remained substantial and significant.

This study also may be limited in its application of school-based poverty as a proxy for individual household poverty. Following recent guidelines from the NYC DOHMH, analyses drew from an area-based poverty measure, school-area poverty, in lieu of individual student-
level meal code status. Area-based poverty may better capture health disparities resulting from socioeconomic differences across individuals, particularly in NYC where great disparities exist in health resources and opportunities across different neighborhoods. Given some students may attend schools in different neighborhoods from where they reside, area-based socioeconomic factors may differ between school and home areas potentially having different influences on the fitness-attendance association. While a more accurate poverty measure may have been to use home-area poverty, home address or zip code was not available in the NYC Fitnessgram dataset for the majority of the analytic population. Future research should be devoted to examining the relationship between school and home zip code to better address the impact of employing different poverty measures in fitness and attendance research.

5.4. Strengths and public health significance

To the author’s knowledge, this is the first multi-year study to examine the causal association between fitness and attendance in youth by drawing from a large and diverse sample of almost 350,000 students followed over 4 years (grades 5-8 per cohort, 2006/7-2012/13), and including a rich body of information pertaining to both individual-student and contextual factors. The analysis also drew from data collected using the Fitnessgram test, which is demonstrated to have both strong reliability and validity.28,29

5.4.1. Public health significance: Prevention paradox and school attendance

When considered at a population level, promoting attendance is a universal goal that can improve overall learning. For example, Gottfried et al. found that a 1 standard deviation increase in the days a student is present in school was associated with a statistically significant
0.28-0.29 standard deviation change in GPA. Some programs, such as the Early Warning Indicator system are used in multiple states to “flag” students who have a high number of absences and provided them with additional outreach and support. However, based on the significance and magnitude of the causal longitudinal fitness-attendance association found here, perhaps effective fitness interventions should target all students, not just those perceived as “high risk”. Indeed, reducing the average risk of poor attendance across the population may be a more effective approach. As per Rose, most of the risk attributable to a disease (or in this scenario, poor attendance), occurs in the left side of the days absent distribution, rather than the tail (where chronically absent students would lie). Put alternatively, population strategies to promote attendance aim to shift the distribution to the left in order to have the largest impact on population attendance. Rose further posits that population-wide approaches to health promotion have the potential to shift entire disease distribution curves to the left. As per Rose, “A large number of people at a small risk may give rise to more cases of disease than the small number who are at a high risk. This situation seems common, and it limits the utility of the ‘high-risk’ approach to prevention.”134(p.431) Rather than focusing on screening programs or other strategies for identifying or treating high-risk individuals, findings here similarly suggest population-wide strategies to improve youth fitness have the potential to more effectively reduce the incidence of poor attendance overall.

5.5. Policy recommendations and future research directions

This study’s findings demonstrate an inverse dose-response relationship between fitness and days absent in both genders, with slightly stronger effects in girls, and youth attending schools in high-poverty areas. Also, findings here suggest that consistent levels of fitness improvement each year at the >20% level may have potential to reduce a child’s days absent almost 2 days
per year over the middle school period, and > 4 days between 7-12th grades compared with students who have >20% decrease in fitness each year. Given over 200,000 NYC students are chronically absent (≥ 20 days absent) each year,89 this work also suggests that fitness programs may be effective in increasing attendance at the population level.

This work supports CDC’s call for policies which promote and sustain physical activity programs in schools.3 Faced with increasing emphases on high-stakes testing, many schools have replaced physical education and other opportunities for physical activity with instructional time. For example, <10% of US middle schools in 2006 provided daily physical education to its students in all grades, and only 57% of school districts required regularly scheduled recess for elementary school students.3 These findings are particularly surprising given the literature which demonstrates no negative impact on children’s cognition when instructional time is reassigned to aerobic physical activity programs.135 Moreover, schools present a unique setting in which to shape the attitudes and behaviors of young individuals towards routine physical activity.64-70 Consistent with the WHO Health Promoting School Framework,10 policy should target increased student physical activity in schools, and further establishing the close association between health and education. Schools in the U.S. serve approximately 56 million youth, and in fact represent an ideal setting in which to promote physical activity in children.3,64,136-139 In this sense, school administrators would likely serve as important stakeholders in supporting school-based physical activity programs and policy.

Current recommendations by the National Association for Sport and Physical Education (NASPE) and the World Health Organization (WHO) state youth aged 6-17 should obtain ≥60
minutes of physical activity per day.\textsuperscript{1,2} The literature suggests this may be achieved through quality physical education before, during, and after school, including recess, extracurricular sports clubs, walk- and bike-to-school programs, movement activities and physical activity breaks in the classroom, and programs to foster family engagement in regular physical activity.\textsuperscript{114-116} Based on this study’s findings, additional work should examine mechanistic pathways in the fitness-attendance association to inform school youth physical activity programming, particularly with respect to policy targeting adolescent girls in impoverished areas to maximize attendance and potentially academic performance benefits from increases in fitness. For example, mediation analyses should be conducted exploring the role of mental and physical health in the fitness--attendance pathway. Moreover, further research should explore the particular effects of gender and area-poverty on the fitness-attendance mechanism.

5.5.1 Conclusion

In sum, this study’s findings demonstrate a dose-response relationship between change in fitness and days absent in middle school youth, with stronger effects in girls and students attending schools in very high- and high- compared with mid- and low-poverty areas. School-level factors accounted for a sizeable and significant proportion of variability in student-specific attendance, both for change in days absent and also predicted probability of chronic absenteeism the year subsequent to change in fitness composite percentile scores. Further research should examine the mechanistic pathway linking change in fitness to student attendance to better inform school-based fitness interventions targeting attendance in youth.
Appendix

Appendix A. Data Cohorts

<table>
<thead>
<tr>
<th>Academic Year of Data Collection</th>
<th>Grades included for Fitnessgram data in analysis:</th>
<th>Grades included for Attendance data in analysis:</th>
<th>Cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006/7</td>
<td>5th(^\wedge) 6th, 7th</td>
<td>n/a*</td>
<td>1</td>
</tr>
<tr>
<td>2007/8</td>
<td>5th, 6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>1,2</td>
</tr>
<tr>
<td>2008/9</td>
<td>5th, 6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>1,2,3</td>
</tr>
<tr>
<td>2009/10</td>
<td>5th, 6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>1,2,3,4</td>
</tr>
<tr>
<td>2010/11</td>
<td>5th, 6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>2,3,4,5</td>
</tr>
<tr>
<td>2011/12</td>
<td>5th, 6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>3,4,5,6</td>
</tr>
<tr>
<td>2012/13</td>
<td>6th, 7th, 8th</td>
<td>6th, 7th, 8th</td>
<td>4,5,6</td>
</tr>
</tbody>
</table>

\(^\wedge\) Grade 5 Fitnessgram data not available for all students.

*Fitness-change from grades 6-7 paired with grade 7 attendance (days absent). Students were required to have at minimum 2 consecutive years of fitness (one year of fitness change) to be included in analytic sample. Attendance data for year 2006/7 does not contribute to analysis given FG initiated 2006/7.
Appendix B. Demographic profile of students excluded from the analysis.

The proportion of males and females in the excluded and analytic populations was identical (49% and 51%, respectively). The excluded population compared with the analytic population was comprised of fewer Asian/Pacific Islanders and Non-Hispanic whites (12% vs. 17%, and 14% vs. 16%, respectively) and slightly more Non-Hispanic blacks and Hispanics (41% vs. 38%, and 32% vs. 28%, respectively. The excluded population also was comprised of fewer foreign born students, and more students born in NYC (vs. foreign and US (not NYC) compared with the analytic population (12% vs. 17%, and 82% vs. 77%, respectively; proportion of students born in the US (not NYC) was identical for the excluded and analytic populations (6%)).

Across area poverty, there was a higher proportion of students attending schools in the very high poverty areas (27% vs. 22%), and a lower proportion of students attending schools in the high, medium and low poverty areas in the excluded population compared with analytic population (27% vs. 26%, 30% vs. 34%, and 16% vs. 18%, respectively). The excluded population had a higher proportion of students who primarily spoke English and Spanish in the home, and a lower proportion of students who primarily spoke a language other than English or Spanish compared with the analytic population (60% vs. 57%, 26% vs. 25%, and 14% vs. 19%, respectively).

Also, the excluded population had a lower proportion of students who were consistently not obese, and a higher proportion who changed weight status from not obese to obese, or who were consistently not obese compared with the analytic population (70% vs. 73%, 21% vs. 17%, and 5% vs. 4%, respectively; percent of students who changed weight status from obese to not obese was identical for the excluded and analytic population (5%)). In addition, mean days absent across students who did not meet inclusion criteria was 15.1 days (SD=17.4) per year (compared with 10.5 days (SD=11.32)).

In sum, the above demographic characteristics of those students excluded from the analysis are not surprising given the two main inclusion criteria for this study were period of enrollment and having at minimum two consecutive years of fitness data. As such, it is not surprising that students who did not meet inclusion criteria had a higher mean number of days absent per year compared with the analytic population. Similarly, it is not surprising that the excluded population had a slightly higher proportion of students who had characteristics shown in the literature to be associated with higher school absenteeism, including slightly higher proportion of Non-Hispanic black and Hispanic students, attending schools in very high poverty areas, born in NYC, and speaking primarily English in the home.19,22,33,34,51
Table 1. Demographic and fitness-change characteristics of excluded population (N=108 552)

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>55 914</td>
<td>51</td>
</tr>
<tr>
<td>Male</td>
<td>52 700</td>
<td>49</td>
</tr>
<tr>
<td><strong>Race/Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Indian</td>
<td>454</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Asian and/or Pacific Islander</td>
<td>13 393</td>
<td>12</td>
</tr>
<tr>
<td>Hispanic</td>
<td>45 005</td>
<td>41</td>
</tr>
<tr>
<td>Non-Hispanic Black</td>
<td>35 003</td>
<td>32</td>
</tr>
<tr>
<td>Non-Hispanic White</td>
<td>14 697</td>
<td>14</td>
</tr>
<tr>
<td><strong>Language Spoken at Home</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>64 680</td>
<td>59</td>
</tr>
<tr>
<td>Spanish</td>
<td>28 628</td>
<td>26</td>
</tr>
<tr>
<td>Other language</td>
<td>15 306</td>
<td>14</td>
</tr>
<tr>
<td><strong>Place of Birth</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NYC</td>
<td>89 140</td>
<td>82</td>
</tr>
<tr>
<td>US (Not NYC)</td>
<td>6 117</td>
<td>6</td>
</tr>
<tr>
<td>Foreign</td>
<td>13 310</td>
<td>12</td>
</tr>
<tr>
<td><strong>Change in Obesity Status (all years)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obese to not obese</td>
<td>270</td>
<td>5</td>
</tr>
<tr>
<td>Consistently not obese</td>
<td>3 789</td>
<td>70</td>
</tr>
<tr>
<td>Consistently obese</td>
<td>1 121</td>
<td>21</td>
</tr>
<tr>
<td>Not obese to obese</td>
<td>257</td>
<td>5</td>
</tr>
<tr>
<td><strong>School-Area Poverty</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low Area Poverty</td>
<td>17 497</td>
<td>16</td>
</tr>
<tr>
<td>Medium Area Poverty</td>
<td>33 033</td>
<td>30</td>
</tr>
<tr>
<td>High Area Poverty</td>
<td>28 898</td>
<td>27</td>
</tr>
<tr>
<td>Very High Area Poverty</td>
<td>29 116</td>
<td>27</td>
</tr>
</tbody>
</table>
Appendix C. NYC Fitnessgram Report

Your BMI

Body Mass Index uses height and weight measurements to find the BMI for age percentile for your age and sex.

Your Previous BMI was 23 placing you in the 72nd percentile for your age.

Your Current BMI is 20, placing you in the 45th percentile for girls aged 17 years. You have a healthy weight.

Height: 5' 5" Weight: 125 lbs BMI: 20
Date Measured: March 2011

NYC FITNESSGRAM

Aerobic Capacity

Aerobic fitness is how well the heart and lungs can perform during physical activity.

The Pacer test measures aerobic capacity. Students run laps between two points in a certain amount of time. The score is the number of laps completed.

Strength, Endurance & Flexibility

Muscle fitness helps to prevent injury and keep the body working properly. Strength, endurance, and flexibility are important for good posture, a healthy lower back, and overall body function.

Simple Steps to A Healthy Weight

Maintaining a healthy weight requires balancing what you eat (calories in) and your physical activity (calories out). Talk with your health care provider to find out what a healthy weight range is for you.

Participating in daily physical activity and eating more fruits and vegetables are two of the best things you can do for your health.
Appendix D. Proc Univariate Descriptive Plot of Days Absent Outcome Variable
Appendix E. Directed Acyclic Graph (DAG) for the causal relationship between health-related fitness and attendance in NYC middle school youth

Figure 1. Illustrative directed acyclic graph showing mediation of the effect of change in health-related fitness on change in school attendance. Green line indicates causal path; red line indicates biasing path; green circle indicates exposure, blue circle indicates outcome, red circle indicates common cause of exposure and outcome.
As shown in the above figure, and consistent with the wellness continuum, it can be hypothesized that health-related fitness improvements may both directly and indirectly promote attendance, working potentially through a pathway involving self-esteem, physical health, mental health, and cognitive processing. Many factors may contribute causally to fitness and attendance, including individual-level sociodemographics, psychosocial factors, BMI, parental factors, and place of birth, as well as area poverty. Moreover, the area poverty-attendance association may be mediated through a pathway involving school and neighborhood resources, psychosocial factors, BMI, and parental attitudes.
Appendix F: Studentized Residual Plots

Conditional Pearson Residuals for absent

Residual Statistics
Observations  675318
Minimum     -2.017
Mean         0.011
Maximum      15.144
Std Dev      0.991

Fit Statistics
Objective    4.93E6
AIC          4.93E6
AICC         4.93E6
BIC          4.93E6
Appendix G. Poisson Model Estimates

<table>
<thead>
<tr>
<th>Change in Fitness(^b)</th>
<th>Unadjusted (Model 2)</th>
<th>Adjusted (Model 3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\beta^a)</td>
<td>95%CI</td>
</tr>
<tr>
<td>&gt;20% Increase</td>
<td>-1.127</td>
<td>-1.119 (\text{lower})</td>
</tr>
<tr>
<td>10-20% Increase</td>
<td>-1.086</td>
<td>-1.077 (\text{lower})</td>
</tr>
<tr>
<td>&lt;10% Change</td>
<td>-1.067</td>
<td>-1.060 (\text{lower})</td>
</tr>
<tr>
<td>10-20% Decrease</td>
<td>-1.020</td>
<td>-1.011 (\text{lower})</td>
</tr>
<tr>
<td>&gt;20% Decrease</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)All beta estimates \(p<0.0001\); \(^b\)Change in fitness composite percentile scores based on PACER (Progressive Aerobic Cardiovascular Endurance Run) Push-up and Curl-up Fitnessgram tests from the year prior. \(^c\)Days absent per year.

The above beta estimates may be interpreted as incident rate ratios. For example, students with higher positive change in fitness had higher attendance per subsequent year. Students with a >20% increase in fitness had 1.12 times the incidence rate for attendance compared with having a >20% decrease in fitness (IRR=1.12). More specifically, a child with consistent improvements in fitness each year at the >20% increase (vs. >20% decrease) level with mean days absent (10) would be expected to have >1 day fewer absences per year; almost 3.5 fewer days absent across the middle-school period and almost 8 fewer days absent per year across the high school period (1.12*7 years=7.84 days).
Appendix H. Model Specifications

Paper 1:
UNCONDITIONAL 2-LEVEL GENERALIZED LINEAR MIXED MODEL:
Attendance_{ij} = \beta_{00} + \mu_{0j} + \epsilon_{ij}, \text{ where } \epsilon_{ij} \sim N(0, \sigma^2) \text{ and } \mu_{0j} \sim N(0, \tau^2)

In this model, i and j correspond to individuals and schools, respectively. \beta_{0j} (\beta_{0j} = \beta_{00} + \mu_{0j}) is the mean days absent across individuals in school j. Also, \beta_{00} represents the overall mean days absent across all schools, \tau^2 is the estimated between-school variance (level 2), and \sigma^2 corresponds to the student-level variance. Lastly, \epsilon_{ij} and u_{0j} (individual-level and school-level effects, respectively) are normally distributed and assumed to be independent of each other.

2-LEVEL AREA POVERTY-ADDED GENERALIZED LINEAR MIXED MODEL:
Attendance_{ij} = \beta_{00} + \beta_{01}(SAP)_j + \mu_{0j} + \epsilon_{ij}

In this model, \beta_{01} is the effect of school-level characteristic (school-area poverty) on mean days absent.

Paper Two:
FINAL 3-LEVEL GENERALIZED LINEAR MIXED MODEL WITH REPEATED MEASURES:

Absent_{ij} = \beta_{000} + \beta_{100}(GRADE)_{ij} + \beta_{200}(FITNESS)_{ij} + \beta_{300}(CHANGE IN OBESITY STATUS)_{ij} + \beta_{400}(YEAR)_{ij} + \beta_{010}(RACE/ETHNICITY)_{ij} + \beta_{020}(GENDER)_{ij} + \beta_{030}(PLACE OF BIRTH)_{ij} + \beta_{040}(STARTING FITNESS)_{ij} + \beta_{050}(SCHOOL AREA POVERTY)_j + \beta_{110}(GRADE)_{ij}*(RACE/ETHNICITY)_{ij} + \beta_{120}(GRADE)_{ij}*(GENDER)_{ij} + \beta_{130}(GRADE)_{ij}*(PLACE OF BIRTH)_{ij} + \beta_{140}(GRADE)_{ij}*(STARTING FITNESS)_{ij} + \beta_{011}(SCHOOL AREA POVERTY)_j*(RACE/ETHNICITY)_{ij} + \epsilon_{ij} + \epsilon_{0ij} + \mu_{00j}

Paper Three:
FINAL 3-LEVEL LOGISTIC MIXED MODEL WITH REPEATED MEASURES:

Logit(chronic absenteeism)_{ij} = \beta_{000} + \beta_{100}(GRADE)_{ij} + \beta_{200}(FITNESS)_{ij} + \beta_{300}(CHANGE IN OBESITY STATUS)_{ij} + \beta_{400}(YEAR)_{ij} + \beta_{010}(RACE/ETHNICITY)_{ij} + \beta_{020}(GENDER)_{ij} + \beta_{030}(PLACE OF BIRTH)_{ij} + \beta_{110}(GRADE)_{ij}*(RACE/ETHNICITY)_{ij} + \beta_{120}(GRADE)_{ij}*(GENDER)_{ij} + \beta_{130}(GRADE)_{ij}*(PLACE OF BIRTH)_{ij} + \beta_{011}(SCHOOL AREA POVERTY)_j*(RACE/ETHNICITY)_{ij} + \beta_{001}(SCHOOL-AREA POVERTY)_j
References


42. van Strien T, Koenders P. How do physical activity, sports, and dietary restraint relate to overweight-associated absenteeism? *J Occup Environ Med*. 2010;52(9):858-64.


72. Balfanz R, Byrnes V. Chronic absenteeism: Summarizing what we know from nationally available data. *Johns Hopkins University Center for Social Organization of Schools*. 2012(May).


84. New York City Department of Health and Mental Hygiene. Childhood obesity is a serious concern in New York City: Higher levels of fitness associated with better academic performance. *NYC Vital Signs*. 2009;1(8).


97. Loeser S, Post J. Mayor Bloomberg launches wake up! NYC campaign to reduce chronic absenteeism and truancy in city schools and releases early data from truancy program. Office of the Mayor. 2011(February 10).


