# EARLY CHILDHOOD INVESTMENTS SUBSTANTIALLY BOOST ADULT HEALTH

**Frances Campbell**[*], **Gabriella Conti**[†], **James J. Heckman**[‡], **Seong Hyeok Moon**[§], **Rodrigo Pinto**[**], **Elizabeth Pungello**[††], and **Yi Pan**[‡‡]

[*]Frances Campbell is a Senior Scientist at the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill.

[†]Gabriella Conti is a Senior Lecturer (Assistant Professor) in the Department of Applied Health Research, University College London.

[‡]James J. Heckman is the Henry Schultz Distinguished Service Professor of Economics at the University of Chicago; the Professor of Science and Society, University College Dublin; and a Senior Fellow at the American Bar Foundation.

[§]Seong Hyeok Moon is a Research Fellow at the Economics Research Center at the University of Chicago.

[**]Rodrigo Pinto is a Research Fellow at the Economics Research Center at the University of Chicago.

[††]Elizabeth Pungello is a Scientist at the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill.

[‡‡]Yi Pan is a Statistician Investigator at the Frank Porter Graham Child Development Institute at the University of North Carolina at Chapel Hill.

## Abstract

High-quality early childhood programs have been shown to have substantial benefits in reducing crime, raising earnings, and promoting education. Much less is known about their benefits for adult health. We report the long-term health impacts of one of the oldest and most heavily cited early childhood interventions with long-term follow-up evaluated by the method of randomization: the Carolina Abecedarian Project (ABC). Using recently collected biomedical data, we find that disadvantaged children randomly assigned to treatment have significantly lower prevalence of risk factors for cardiovascular and metabolic diseases in their mid-30s. The evidence is especially strong for males. The mean systolic blood pressure among the control males is 143, while only 126 among the treated. One in four males in the control group is affected by metabolic syndrome, while none in the treatment group is. To reach these conclusions, we address several statistical challenges. We use exact permutation tests to account for small sample sizes and conduct a parallel bootstrap confidence interval analysis to confirm the permutation analysis. We adjust inference to account for the multiple hypotheses tested and for nonrandom attrition. Our evidence shows the potential of early life interventions for preventing disease and promoting health.

## 1 Introduction

Noncommunicable diseases are responsible for roughly two-thirds of worldwide deaths (1). Most policies that combat disease currently focus on treatment after disease occurs and on reducing risk factors in adult life. Recent discussions of effective ways of controlling the soaring costs of the US health care system emphasize tertiary prevention; i.e., reducing the worsening of the conditions of those already ill (see, e.g., (2)) and "bending the cost curve" for such treatments (2–5).

A complementary approach is to prevent disease, or to delay its onset. A large body of evidence shows that adult illnesses are more prevalent and problematic among those who have experienced adverse early life conditions (6, 7). The exact mechanisms through which early life experiences translate into later life health are being actively investigated (8, 9).

This paper shows that high-quality, intensive interventions in the early years can be effective in preventing, or at least delaying, the onset of adult disease. The recent literature establishes that interventions that enrich the environments of disadvantaged children have substantial impacts on a variety of outcomes throughout their lives (see, e.g., (10–12)). However, little is known about their benefits on health (see, e.g. , (13, 14)).

We study the long-term health effects of one of the oldest and most cited early childhood programs: the Carolina Abecedarian Project (ABC). ABC was designed as a social experiment to test if a stimulating early childhood environment could prevent the development of mild mental retardation in disadvantaged children. The study was conducted on four cohorts of disadvantaged children born between 1972 and 1977 who were living in or near Chapel Hill, North Carolina. The intervention consisted of a two-stage treatment targeted to different segments of child life cycles: an early childhood stage (from birth to age 5) and a subsequent school-age stage (from age 6 to 8). The first stage of the intervention involved periods of cognitive and social stimulation interspersed with caregiving and supervised play throughout a full eight-hour day for the first 5 years. The stimulation component was based on a curriculum that emphasized development of language, emotional regulation, and cognitive skills (15, 16). The second stage of the intervention focused on improving early math and reading skills through having "home-school resource teachers" customize learning activities based on materials being covered at school and then deliver these materials to the parents to use at home. The treatment and control groups from the first stage were randomly assigned to treatment and control groups in the second stage. We analyze data on treatment and control groups created by the first-stage randomization. We find no evidence of any treatment effect on adult health from the second-stage randomization. The treatment effects are much smaller in magnitude than those estimated for the first-stage treatment, and fail to achieve statistical significance at conventional levels. See Supplementary Material, Section F, for evidence on this issue. References (17–19) show that for most outcomes the early educational intervention had much stronger effects than the school-age treatment. Additionally, a previous work has also shown no health effects from a school-age (as compared to a preschool) educational intervention (20). The available evidence on interventions to prevent obesity points to the years 0–5 as a critical period (as compared to after 5, see e.g. (21–23)).

The Abecedarian intervention had also a nutritional and health care component. Treated children had two meals and a snack at the childcare center during the first stage. They were offered primary pediatric care (both well- and ill-child care), with periodic check-ups and daily screening, in both stages. More details on the intervention are given in the Supplementary Material, Section A.

## Data

Data were collected on both treated and control cases from the beginning of their participation in the study, using surveys administered to children, parents, and teachers, as well as direct assessments. Before the intervention started, baseline information was gathered on parental characteristics, family structure, socioeconomic status, and birth circumstances. For both treated and control cases, data on cognition, personality, health, achievement, and behavior were then collected at multiple stages from birth until the end of school-age treatment. At the end of the second stage of treatment, children were followed up at ages 12, 15, 21, 30, and in the mid-30s. Details on the outcomes and covariates used in this analysis are provided in the Supplementary Material, Section B.

A biomedical survey of cardiovascular and metabolic risk factors was conducted when participants were in their mid-30s. Information on biomeasures was collected from two sources. The first was a physical exam carried out by a local physician in the Chapel Hill Internal Medicine practice, in which the same doctor (blind to treatment status) examined all subjects. In this exam, measurements were collected on weight (lbs), height (inches), waist (inches), hips (inches), and systolic and diastolic blood pressure. The physician also checked the status of several body systems. The physician carried out a complete physical exam and checked whether there was abnormality in relation to the following systems: skin, HEENT (Head, Ear, Eye, Nose, Throat), neck, chest, lung, breast, cardiovascular, abdomen, neurologic, muscle strength and tone, musculoskeletal, lymphatic. The second source was lab tests, based on non-fasting venous blood collected from the subjects during the medical visit (the phlebotomist was blind to treatment status and the blood samples were sent out to another facility for analysis and report preparation).

Several issues arise in evaluating the health effects of the ABC intervention. First, the sample size is small. Conventional testing approaches that rely on large-sample properties of test statistics may be inappropriate. To surmount this problem, we use exact (small-sample) block permutation tests. However, as we show in Tables S25—S26 in Supplementary Material, Section I, when we use bootstrap methods that have a large sample justification, we obtain the same inference about treatment effects. Bootstrapping has the additional benefit of producing confidence intervals to gauge the uncertainty inherent in our estimates.

Second, numerous treatment effects are analyzed. This creates an opportunity for "cherry picking"—finding spurious treatment effects merely by chance if conventional one-hypothesis-at-a-time approaches to testing are used. We account for the multiplicity of the hypotheses being tested using recently developed stepdown procedures (24).

Third, information is missing due to non-random attrition from the survey, potentially undermining the validity of inference. We investigate the causes of missing information and

correct for potential bias using inverse probability weighting (IPW) (25, 26). More information on the methodology and a detailed analysis of the attrition patterns is presented in Supplementary Material, Sections C, D, and H.

## 2 Results

### Physical Health

Estimated treatment effects and associated test statistics are given in Tables 1 (males) and 2 (females). Throughout the paper, we report one-sided single hypothesis block permutation *p*-values associated with the IPW treatment effect estimates; multiple hypothesis stepdown *p*-values are reported in Tables 1 and 2. We first report the experimental results on the biomarkers of cardiovascular functioning. On average, treated males have lower values of both systolic and diastolic blood pressure. This difference amounts to 13.5 mmHg for diastolic blood pressure ( *p* =0.024) and to 17.5 mmHg for systolic blood pressure ( *p* =0.018). Treated females are less likely to be pre-hypertensive. The prevalence of pre-hypertension (systolic bp ≥120 or diastolic bp ≥80; see (27)) is 0.909 in the control group and 0.667 in the treatment group, and the difference is statistically significant (*p* =0.042). Using two different definitions of hypertension (Systolic bp ≥140 and diastolic bp ≥90, (28), and systolic bp ≥140 or diastolic bp ≥90, (27)), treated males are less likely to fall into the stage I hypertension category (a prevalence of only 0.105 or 0.211) as compared to a much higher prevalence observed in the control group (0.444 and 0.556). Both treatment effects are statistically significant (*p* =0.010 and *p* =0.038) (29).

Biomarkers of metabolic activity from blood tests (lipid panel) show that treated individuals have higher levels of High-Density Lipoprotein Cholesterol (HDL-C)—"good" cholesterol. The magnitude of the difference between treated and control groups is larger for males. The control males have a level of HDL cholesterol of 42 mg/dL which is just above the lower recommended limit of 40 mg/dL (30), while the level for the treated males is 11 mg/dL higher. The treatment effect is marginally significant ( *p* =0.066). This is reflected in the prevalence of dyslipidemia (elevated lipid levels). The difference in the prevalence of this condition between treatment and control groups is 0.311 for males (HDL-C < 40 mg/dL; *p* =0.179) and 0.177 for females (HDL-C < 50 mg/dL; *p* =0.099). The healthier metabolic status experienced by the male treatment group is confirmed by the lower prevalence of pre-diabetes indicators (Glycosylated Hemoglobin ≥5.7%, (31) 0.473 vs. 0.583), although the difference does not attain statistical significance ( *p* =0.426). Control males are also twice as likely to be affected by vitamin D deficiency (Total Vitamin D < 20 ng/mL, (32); 0.368 vs. 0.750; *p* =0.021).

The prevalence of both severe and abdominal obesity is lower among treatment group males but the differences are not statistically significant at the 10% level. Treated females are less likely than controls to be affected by abdominal obesity, both when considering the waist-hip ratio (WHR) and when analyzing a dichotomous measure of WHR> 0.85 ((33); 0.563 vs. 0.762); both treatment effects are marginally significant (*p* =0.063 and *p* =0.080, respectively).

The health effects of the Abecedarian intervention translate into lower prevalence of *multiple* risk factors that are particularly striking for males. Those in the treatment group are less likely to experience *both* obesity and hypertension (diff.=0.389; *p* =0.016), severe obesity and hypertension (diff.=0.375; *p* =0.005), and dyslipidemia and hypertension (diff.=0.333; *p* =0.006). *None* of the treated males have the cluster of conditions known as Metabolic Syndrome (defined as waist circumference > 102 cm or 40 inches (34); HDL-C < 40 mg/dL; blood pressure ≥130/85 mmHg; see (30)), associated with greater risk of heart disease, stroke, and diabetes, while *one in four* in the control group is affected by it (*p* =0.007). The prevalence of the Metabolic Syndrome for females (defined as waist circumference > 88 cm or 35 inches (34); HDL-C < 50 mg/dL; blood pressure ≥130/85 mmHg; see (30)) is lower in the treatment group but the differences are not statistically significant at the 10% level. Finally, results for the Framingham Risk Score (35) reveal that both treated males and females have a significantly lower risk of experiencing "total" Coronary Heart Disease (CHD), defined as both stable and unstable angina, myocardial infarction, or CHD death, within the next 10 years (diff.=2.154, *p* =0.038; diff.=0.339, *p* =0.070).

In sum, the available evidence from the biomedical survey of ABC shows that the children who attended the child care center in the first five years of their lives enjoy better physical health in their mid-30s with significant markers indicating better future health. The benefits of these health improvements are substantial and wide-ranging. Reference (36) provides a detailed review of the labor market costs of obesity, which range from increased absenteism to lower productivity and wages. There are considerable losses in life expectancy due to obesity. Reference (37) reports estimates that 35-year-old males with hypertension would gain 1.1–5.3 years of expected life (0.9–5.7 years for females) from reducing their diastolic blood pressure to 88 mmHg using the Coronary Heart Disease Policy Model based on data from the Framingham Heart Study. Reference (38), using data from the Framingham Heart Study, finds that 40-year-old male nonsmokers suffer a loss of life expectancy of 3.1 years (3.3 years for females) of life because of overweight, and of 5.8 years (7.1 years for females) because of obesity. Reference (39), using data from the National Longitudinal Study of Adolescent Health, shows that diabetics are less likely to be employed (by 8–11 percentage points), more likely to participate in social programs (by 8–13 percentage points), and earn on average lower wages (by $1,500–6,000). Reference (40) provides further evidence from the National Longitudinal Survey of Youth 1979 that the duration of diabetes is negatively associated with employment and wages. Reference (41) reports a hazard ratio of 1.47 (95% CI 1.13–1.92) for all-cause mortality and of 2.53 (95% CI 1.74–3.67) for cardiovascular mortality caused by metabolic syndrome (NCEP definition) in the San Antonio Heart Study.

### Health Care

Availability of health care is a necessary condition for enjoying better health, although not a sufficient one (42). The upper panel of Table 3 reveals that treated males were more likely to be covered by health insurance at age 30 (0.704 vs. 0.476; *p* =0.039) and to be cared for in a hospital or by a doctor when sick (0.815 vs. 0.524; *p* =0.037).

### Physical Development

We analyze the effects of the intervention on early physical development, assessed using anthropometric measurements (height and weight) taken at the times the children had their routine assessments at multiple times in childhood. We transform the BMI measures into standard normal variates (*z*-scores) using the LMS method developed in (43, 44) and (45). The results are reported in the bottom panel of Table 3. Treated males were less likely than controls to be overweight throughout their preschool years, with almost no treated child having a weight-for-length above the 85th percentile (the age-specific measure for being "at-risk overweight", see (46)) in the first two years of life. Control males had a greater weight-for-length *z*-score change between birth and 24 months of age. More rapid increases in weight-for-length in the first 6 months of life have been associated with increased risk of obesity at age 3 (47). Looking at the full BMI distribution by treatment status shown in Figure 1, it is evident that the distribution is both less spread out and shifted to the left for treated males relative to controls. These results are consistent with the obesity-reducing effects found in Head Start (13, 14) and are consistent with evidence in the literature of the important role played by early life nutrition (48). Further evidence on the importance of these early growth patterns is shown in Figure 2. The two upper figures show the evolution of BMI-for-age during childhood for males by treatment status. It is noticeable that, while the BMI-for-age of the treatment group is always centered around the median for the reference population, the control group experiences a surge in the first year, which peaks at 18 months, becomes partially attenuated, and then exhibits diverging growth patterns after 5 years of age. And it is striking that, when we consider the early growth trajectory by obesity status in adulthood (panels c–f of Figure 2), those who are obese or severely obese in their mid-30s are already on a trajectory of above-normal BMI in the first five years of their lives.

## Conclusions

This paper analyzes recently collected biomedical data for the Abecedarian intervention. Children randomly assigned to the treatment group for ages 0–5 have a significantly lower prevalence of risk factors for cardiovascular and metabolic diseases in their mid-30s. Treated males have a healthier body mass in their childhood years. These early benefits persist into adulthood.

The precise mechanisms through which these effects are obtained remain to be determined. It may be improved health due to access to pediatric care and proper nutrition in the early years, improved noncognitive skills as in the Perry study (49), improved cognitive skills, or some combination of all three factors. Uniformity of the treatment does not provide the necessary independent variation in the components of the intervention that would allow us to examine the sources of treatment effects. A simple mediation analysis (presented in Appendix G, Tables S19 and S20), suggests that half of the effect of the treatment on hypertension and obesity in the mid-30s may be mediated by the body mass index of the child around one year of age, while no significant role seems to be played by the availability of health insurance, or improved socioeconomic status at age 30. However, the estimated mediation effects are not precisely determined, so these findings are necessarily speculative at this point. Whatever the channel, our evidence supports the importance of intervening in

the first years of life, and suggests that early childhood programs can make a substantial contribution to improving the health of adult Americans and reducing the burden of health care costs. An intervention that lasted five years and cost $67,000 (in 2002 dollars, see (50)) produced sustained and substantial health benefits. Early childhood interventions are an unexplored and promising new avenue of health policy.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

## References

1. Alwan, A.; Armstrong, T.; Bettcher, D.; Branca, F.; Chisholm, D.; Ezzati, M.; Garfield, R.; MacLean, D.; Mathers, C.; Mendis, S.; Poznyak, V.; Riley, L.; Tang, KC.; Wild, C. Global status report on noncommunicable diseases 2010: Description of the global burden of NCDs, their risk factors and determinants. Geneva, Switzerland: World Health Organization; 2011.

2. Emanuel E. Prevention and Cost Control. Science. 2012; 337:1433–1433. [PubMed: 22997291]

3. Cutler DM, Nikhil SR. If Slow Rate Of Health Care Spending Growth Persists, Projections May Be Off By $770 Billion. Health Affairs. 2013; 32

4. Antos JR, Pauly MV, Wilensky GR. Bending the Cost Curve through Market-Based Incentives. New England Journal of Medicine. 2012; 367:954–958. [PubMed: 22852881]

5. Ginsburg PB, Ichiseki H, Punwani N. Correspondence: Bending the Health Care Cost Curve. New England Journal of Medicine. 2012; 367:2454–2456. [PubMed: 23252543]

6. Danese A, Pariante CM, Caspi A, Taylor A, Poulton R. Childhood Maltreatment Predicts Adult Inflammation in a Life-Course Study. Proceedings of the National Academy of Sciences of the United States of America. 2007; 104:1319–1324. [PubMed: 17229839]

7. Galobardes B, Lynch JW, Davey-Smith G. Is the association between childhood socioeconomic circumstances and cause-specific mortality established? Update of a systematic review. Journal of epidemiology and community health. 2008; 62:387–390. [PubMed: 18413449]

8. Hertzman C. The Biological Embedding of Early Experience and its Effects on Health in Adulthood. Annals of the New York Academy of Sciences. 1999; 896:85–95. [PubMed: 10681890]

9. Entringer S, Buss C, Wadhwa PD. Prenatal Stress, Telomere Biology, and Fetal Programming of Health and Disease Risk. Science Signaling. 2012; 5:pt12. [PubMed: 23112344]

10. Heckman JJ. Schools, Skills and Synapses. Economic Inquiry. 2008; 46:289–324. [PubMed: 20119503]

11. Hines PJ, Mervis J, McCartney M, Wible B. Investing Early in Education. Science. 2011 Aug 19.333:909–1056.

12. Nores M, Barnett WS. Benefits of Early Childhood Interventions across the World: (Under)Investing in the Very Young. Economics of Education Review. 2010; 29

13. Frisvold DE, Lumeng JC. Expanding Exposure: Can Increasing the Daily Duration of Head Start Reduce Childhood Obesity? Journal of Human Resources. 2011; 46:373–402.

14. Carneiro, P.; Ginja, R. Institute for the Study of Labor (IZA); 2012. Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start.

15. Sparling, J.; Lewis, I. LearningGames for the First Three Years : A Guide to Parent/Child Play. New York, NY: Berkley Books; 1979.

16. Sparling, J.; Lewis, I. LearningGames for Threes and Fours. New York, NY: Walker and Company; 1984.

17. Campbell FA, Ramey CT. Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. Child Development. 1994; 65:684–698. [PubMed: 8013248]

18. Campbell FA, Ramey CT. Cognitive and School Outcomes for High-Risk African-American Students at Middle Adolescence: Positive Effects of Early Intervention. American Educational Research Journal. 1995; 32:743–772.

19. F.A.C. Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. Early Childhood Research Quarterly. 2008; 23:452–466. e. al.

20. Reynolds AJ, Temple Judy A, Ou Suh-Ruu, Arteaga Irma A, Barry AB. White, School-based early childhood education and age-28 well-being: Effects by timing, dosage, and subgroups. Science. 2011; 333:360–364. [PubMed: 21659565]

21. Davis K, Kaufer Christoffel Katherine. Obesity in preschool and school-age children: treatment early and often may be best. Archives of pediatrics & adolescent medicine. 1994; 148:1257. [PubMed: 7951803]

22. Flynn MAT, McNeil DA, Maloff B, Mutasingwa D, Wu M, Ford C, Tough SC. Reducing obesity and related chronic disease risk in children and youth: a synthesis of evidence with 'best practice' recommendations. Obesity reviews. 2006; 7:7–66. [PubMed: 16371076]

23. Waters E, de Silva Sanigorski A, Hall BJ, Brown T, Campbell KJ, Gao Y, Armstrong R, Prosser L, Summerbell CD. Interventions for preventing obesity in children (review). Cochrane collaboration. 2012; 12:1–212.

24. Romano JP, Wolf M. Exact and Approximate Stepdown Methods for Multiple Hypothesis Testing. Journal of the American Statistical Association. 2005; 100:94–108.

25. Johnston, J.; DiNardo, JE. Econometric Methods. ed. 4. New York, NY: McGraw-Hill; 1997.

26. Horvitz DG, Thompson DJ. A Generalization of Sampling Without Replacement from a Finite Universe. Journal of the American Statistical Association. 1952; 47:663–685.

27. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL Jr, Jones DW, Materson BJ, Oparil S, Wright JT Jr. Seventh report of the joint national committee on prevention, detection, evaluation, and treatment of high blood pressure. Hypertension. 2003; 42:1206–1252. [PubMed: 14656957]

28. Strong, K.; Bonita, R. World Health Organization; 2003. Surveillance of Risk Factors related to noncommunicable diseases: Current Status of Global Data.

29. Treated males also report a lower blood pressure in the age 30 interview (0.190 vs. 0.408) and in the history section of the mid-30s medical visit (0.158 vs. 0.444); in neither of these cases are the differences statistically significant, although the rates are remarkably similar to those obtained for hypertension in the physical exam.

30. National Cholesterol Education Program, Third Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) Final Report. Circulation. 2002; 106:3143–3421. [PubMed: 12485966]

31. American Diabetes Association, Standards of Medical Care in Diabetes--2013. Diabetes Care. 2013; 36:S11–S66. [PubMed: 23264422]

32. Holick MF, Chen TC. Vitamin D deficiency: A worldwide problem with health consequences. American Journal of Clinical Nutrition. 2008; 87:1080S–1086S. [PubMed: 18400738]

33. WHO Consultation. World Health Organization; 1999. Definition, diagnosis and classification of diabetes mellitus and its complications.

34. Grundy SM, Cleeman JI, Daniels SR, Donato KA, Eckel RH, Franklin BA, Gordon DJ, Krauss RM, Savage PJ, Smith SC, Spertus JA, Costa F. Diagnosis and management of the metabolic syndrome. Circulation. 2005; 112:2735–2752. [PubMed: 16157765]

35. Wilson PWF, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of Coronary Heart Disease Using Risk Factor Categories. Circulation. 1998; 97:1837–1847. [PubMed: 9603539]

36. Cawley, J. The economics of obesity. Oxford: Oxford University Press; 2011.

37. Tsevat J, Weinstein Milton C, Williams Lawrence W, Tosteson Anna N, Goldman Lee. Expected gains in life expectancy from various coronary heart disease risk factor modifications. Circulation. 1991; 83:1194–1201. [PubMed: 2013141]

38. Peeters A, Barendregt Jan J, Willekens Frans, Mackenbach Johan P, Al Mamun Abdullah, Bonneux Luc. Obesity in adulthood and its consequences for life expectancy: a life-table analysis. Annals of internal medicine. 2003; 138:24–32. [PubMed: 12513041]

39. Fletcher JM, Richards Michael R. Diabetes's 'Health Shock' To Schooling And Earnings: Increased Dropout Rates And Lower Wages And Employment In Young Adults. Health Affairs. 2012; 31:27–34. [PubMed: 22232091]

40. Minor T. An Investigation into the Effect of Type I and Type II Diabetes Duration on Employment and Wages. Economics & Human Biology. 2013; 11:534–544. [PubMed: 23659822]

41. Hunt KJ, Resendez Roy G, Williams Ken, Haffner Steve M, Stern Michael P. National Cholesterol Education Program versus World Health Organization metabolic syndrome in relation to all-cause and cardiovascular mortality in the San Antonio Heart Study. Circulation. 2004; 110:1251–1257. [PubMed: 15326061]

42. Levy, H.; Meltzer, D. Health Policy and the Uninsured. MacLaughlin, CG., editor. Washington, DC: Urban Institute Press; 2004. p. 179-204.

43. Cole TJ. The LMS method for constructing normalized growth standards. European Journal of Clinical Nutrition. 1990; 44:45–60. [PubMed: 2354692]

44. Cole TJ, Green PJ. Smoothing reference centile curves: The LMS method and penalized likelihood. Statistics in Medicine. 1992; 11:1305–1319. [PubMed: 1518992]

45. The LMS method is an age- and gender-dependent transformation that translates BMI measures into z-scores. The transformation is applied to two widely used population-based reference data: (1) the 2000 Centers for Disease Control and Prevention growth standards (CDC); and (2) the 2006 WHO child growth standards scale (WHO). The LMS method transforms the information on the median (M), coefficient of variation (S), and skewness (L) of the BMI distribution of these population-based reference data into a Box-Cox power function. This information is then used to transform BMI measures into z-scores.

46. Barlow SE. the Expert Committee, Expert Committee Recommendations Regarding the Prevention, Assessment, and Treatment of Child and Adolescent Overweight and Obesity: Summary Report. Pediatrics. 2007; 120:S164–S192. [PubMed: 18055651]

47. Taveras EM, Rifas-Shiman SL, Belfort MB, Kleinman KP, Oken E, Gillman MW. Weight Status in the First 6 Months of Life and Obesity at 3 Years of Age. Pediatrics. 2009; 123:1177–1183. [PubMed: 19336378]

48. A reduction in BMI can be caused by either better nutrition, and/or more physical exercise. Physical environment is certainly important: the FPG center included a large open space, and outside play was a part of the daily routine (11, 30am and 3pm); however, this is likely not to have played a significant role only after the toddlers had started walking. Of course, better nutrition does not necessarily come only in the daycare center: the treated children could have enjoyed more nutritious food at home as well, both because their preferences for food might have been affected, and because their parents were counselled by the pediatricians on site during the physical exams.

Finally, Frisvold and Lumeng, 2001, provide evidence - based on the What We Eat in America 2003–2004, combined with the National Health and Nutrition Examination Survey (NHANES) 2003–2004 - that Head Start participants consume similar level of calories as non-Head Start participants at evenings and in weekends, but fewer calories in the morning and in the afternoon during the week.

49. Heckman JJ, Pinto R, Savelyev PA. Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes. American Economic Review. 2013; 103:2052–2086. [PubMed: 24634518]

50. Barnett SW, Masse LN. Comparative benefit–cost analysis of the Abecedarian program and its policy implications. Economics of Education Review. 2007; 26:113–125.

51. Ramey CT, Campbell FA, Skinner ML, Burchinal M, Gardner DM, Ramey SL. Persistent Effects of Early Childhood Education on High-Risk Children and Their Mothers. Applied Developmental Science. 2000; 4:2–14.

52. Breitmayer BJ, Ramey CT. Biological Nonoptimality and Quality of Postnatal Environment as Codeterminants of Intellectual Development. Child Development. 1986; 57:1151–1165. [PubMed: 3769604]

53. Ramey CT, Campbell FA. Preventive Education for High-Risk Children: Cognitive Consequences of the Carolina Abecedarian Project. American Journal of Mental Deficiency. 1984; 88:515–523. [PubMed: 6731489]

54. Ramey CT, Smith BJ. Assessing the Intellectual Consequences of Early Intervention with High-Risk Infants. American Journal of Mental Deficiency. 1977; 81:318–324. [PubMed: 836631]

55. Ramey C, Yeates KO, Short EJ. The Plasticity of Intellectual Development: Insights from Preventive Intervention. Child Development. 1984; 55:1913–1925. [PubMed: 6510061]

56. Here we refer to the first randomization, which took place before the preschool intervention; a second randomization took place before the school-age intervention (5–8).

57. (19) explain the higher rate of study rejection in the treatment group by noting "mothers wanting to care for their infants at home."

58. The two interventions each targeted at-risk children, employed similar teacher-child ratios in their preschool-aged classrooms, and shared a common perspective on child development and learning, such that: First, each drew their conceptual frameworks from theorists Piaget and Vygotsky, wherein the role of the teacher is to design the environment in order to create learning experiences at the appropriate developmental level of the child, and the role of each individual child is to direct their own learning within this structured environment. Further, teachers work with each child at their unique level until new skills are internalized and the child can extend their learning at higher levels of complexity. Second, both emphasized early language development, literacy skills, and healthy social-emotional development by integrating caregiving experiences with structured verbal interactions including adult narration of child's activities, adult repetition of child's sounds and words, frequent joint reading of shared books, and the adult use of open-ended questions to extend a child's conversation. Third, both emphasized individualized instruction and the use of formative assessment, with teachers maintaining detailed records of their observations of each individual child's activities and growth and using these assessments to guide subsequent teaching experiences.

59. Campbell FA, Ramey CT, Pungello E, Sparling J, Miller-Johnson S. Early Childhood Education: Young Adult Outcomes from the Abecedarian Project. Applied Developmental Science. 2002; 6:42–57.

60. Campbell FA, Pungello EP, Miller-Johnson S, Burchinal M, Ramey CT. The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. Developmental Psychology. 2001; 37:231–242. [PubMed: 11269391]

61. Barnett, WS.; Masse, LN. New Brunswick, NJ: National Institute for Early Education Research (NIEER); 2002. A Benefit-Cost Analysis of the Abecedarian Early Childhood Intervention.

62. The medical care was no-cost to the families, but they were responsible for buying medicines. The original plan was to also provide medical care at the FPG center to the control families, but after the first year that proved to be impractical.

63. Ramey, CT.; McGinness, GD.; Cross, L.; Collier, AM.; Barrie-Blackley, S. The Social Life of Children in a Changing Society. Borman, KM., editor. Hillsdale, NJ: Lawrence Erlbaum Associates; 1982. p. 145-174.

64. Roberts JE, Sanyal MA, Burchinal MR, Collier AM, Ramey CT, Henderson FW. Otitis Media in Early Childhood and Its Relationship to Later Verbal and Academic Performance. Pediatrics. 1986; 78:423–430. [PubMed: 3748675]

65. Sanyal MA, Henderson FW, Stempel EC, Collier AM, Denny FW. Effect of upper respiratory tract infection on eustachian tube ventilatory function in the preschool child. Journal of Pediatrics. 1980; 97:11–15. [PubMed: 6966684]

66. The licensed practical nurse visited the classroom daily to review the health status of the children and receive reports from the parents (64).

67. HDL-C=High-Density Lipoprotein Cholesterol.

68. WHR=Waist-Hip Ratio.

69. Individuals who reported to have smoked in the past 12 months were coded as smokers. Diabetes status was defined as having HbA1C 6.5 or use of diabetes medication reported by the doctor.

70. Responses A="I have never smoked cigarettes regularly" were coded as missing, while nobody replied B="Less than 9 years old", or C="9 or 10 years old" at age 21.

71. The alcohol section is preceded by the following prompt: "The next questions ask about drinking alcohol. This includes drinking beer, wine, wine coolers, and liquor such as rum, gin, vodka, or whiskey. For these questions, drinking alcohol does not include drinking a few sips of wine for religious purposes".

72. Nobody replied B="Less than 9 years old".

73. The alcohol section is preceded by the following prompt: "The next questions ask about drinking alcohol. This includes drinking beer, wine, wine coolers, and liquor such as rum, gin, vodka, or whiskey. For these questions, drinking alcohol does not include drinking a few sips of wine for religious purposes".

74. The alcohol section is preceded by the following prompt: "Questions about drinking alcohol. This includes beer, wine, liquor. Sips of wine for religious purposes excluded."

75. "Other than a few sips" is in parentheses at the age 30 survey.

76. Responses of A="I have never had a drink of alcohol other than a few sips" were coded as missing, while nobody replied B="Less than 9 years old" at age 21.

77. Responses of A="I have never tried marijuana" were coded as missing, while nobody replied B="Less than 9 years old" at age 21.

78. Miller, A. Subset selection in regression. CRC Press; 2002.

79. Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. Journal of the American Statistical Association. 1995; 90:106–121.

80. John F, Gottschalk P, Moffitt R. An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. Journal of Human Resources. 1998:251–299.

81. Akaike H. A new look at the statistical model identification. IEEE Transactions on Automatic Control. 1974; 19:716–723.

82. One child was diagnosed to be biologically retarded, two died at very early ages (3 and 4 months), and other two children died later (12 and 50 months). One child was withdrawn from the sample before 6 months of age. In addition, other 10 children dropped out of the sample before 54 months.

83. For males, we compare attriters vs. non-attriters both in the lab tests and the physical exam sample, since participation varies between the two, as seen in Table S3.

84. Although more likely to have other special circumstances.

85. Control females who took part in the biomedical sweep were also more likely to report having any illness or headache at age 30 (Table S5).

86. The only exception is the males in the control group who took part in the blood tests, for whom no significant difference emerges between attrited and non-attrited (Table S9).

87. The variables in Tables S4–S9 are not normalized in the"favorable" direction. In these tables, the difference in means reported is simply the difference between the mean of the non-attrited and the

mean of the attrited. The test is based on a two-sided p-value. Namely, we do not test for the health effects of the intervention, but for statistically significant differences between attrited and non-attrited, within each treatment group (and gender), with respect to both baseline and age 30 characteristics. This is in sharp contrast with the tables that test for the health effects of the intervention: Tables 1–2 in the paper and Tables S10–S18 in the appendix. For those we use one-sided p-values as we test the hypothesis that the preschool-age intervention had either positive or no effects on the health outcomes.

88. Burchinal MR, Campbell FA, Bryant DM, Wasik BH, Ramey CT. Early intervention and mediating processes in cognitive performance of children of low-income African American families. Child Development. 1997; 68:935–954.

89. Campbell FA, Pungello Elizabeth P, Burchinal Margaret, Kainz Kirsten, Pan Yi, Barbara H Wasik, Oscar A Barbarin, Sparling Joseph J, Ramey Craig T. Adult outcomes as a function of an early childhood educational program: An Abecedarian Project follow-up. Developmental Psychology. 2012; 48:1033. [PubMed: 22250997]

90. Lloyd LJ, Langley-Evans SC, McMullen S. Childhood obesity and risk of the adult metabolic syndrome: a systematic review. International Journal of Obesity. 2011; 36.1:1–11. [PubMed: 22041985]

91. Lobstein T, Baur Louise, Uauy Ricardo. Obesity in children and young people: a crisis in public health. Obesity reviews. 2004; 5:4–85. [PubMed: 15096099]

92. Johnston FE, Mack Robert W. Obesity in urban black adolescents of high and low relative weight at 1 year of age. American journal of diseases of children. 1978; 132:862–864. [PubMed: 685902]

93. Yu ZB, Han SP, Cao XG, Guo XR. Intelligence in relation to obesity: a systematic review and meta-analysis. Obesity reviews. 2010; 11:656–670. [PubMed: 19780990]

94. Starr JM, Taylor Michelle D, Hart Carole L, Smith George Davey, Whalley Lawrence J, Hole David J, Wilson Valerie, Deary Ian J. Childhood mental ability and blood pressure at midlife: linking the Scottish Mental Survey 1932 and the Midspan studies. Journal of hypertension. 2004; 22:893–897. [PubMed: 15097227]

95. Richards M, Black Stephanie, Mishra Gita, Gale Catharine R, Deary Ian J, Batty David G. IQ in childhood and the metabolic syndrome in middle age: Extended follow-up of the 1946 British Birth Cohort Study. Intelligence. 2009; 37:567–572. [PubMed: 23226908]

96. Pulkki-Råback L, Elovainio Marko, Kivimäki Mika, Raitakari Olli T, Keltikangas-Järvinen Liisa. Temperament in childhood predicts body mass in adulthood: the Cardiovascular Risk in Young Finns Study. Health Psychology. 2005; 24:307. [PubMed: 15898867]

97. Ravaja N, Keltikangasjarvinen Liisa. Temperament and metabolic syndrome precursors in children: a three-year follow-up. Preventive medicine. 1995; 24:518–527. [PubMed: 8524728]

98. Lochner L. Nonproduction Benefits of Education: Crime, Health, and Good Citizenship. Handbook of the Economics of Education. 2011; 4:183.

99. Smith JP. Health bodies and thick wallets: the dual relation between health and economic status. The Journal of Economic Perspectives. 1999; 13:145–166.

100. Marmot M. The influence of income on health: views of an epidemiologist. Health Affairs. 2002; 21:31–46. [PubMed: 11900185]

101. Levy H, Meltzer D. The impact of health insurance on health. Annual Review of Public Health. 2008; 29:399–409.

102. Currie J, Phongsack Manivong Mark Stabile, Roos Leslie L. Child health and young adult outcomes. Journal of Human Resources. 2010; 45:517–548.

103. Kleinert S, Chertok IR, Zimmerman DR, Mathai M, Gulmezoglu AM, Hill S, Madamombe I, et al. Adolescent health: an opportunity not to be missed. Lancet. 2007; 369:1057–1058. [PubMed: 17398287]

104. Di Cesare M, Khang Young-Ho, Asaria Perviz, Blakely Tony, Cowan Melanie J, Farzadfar Farshad, Guerrero Ramiro, et al. Inequalities in non-communicable diseases and effective responses. The Lancet. 2013; 381:585–597.

105. Blinder AS. Wage discrimination: reduced form and structural estimates. Journal of Human Resources. 1973:436–455.

106. Oaxaca R. Male-female wage differentials in urban labor markets. International Economic Review. 1973; 14:693–709.

107. Fortin NM. The Gender Wage Gap among Young Adults in the United States: The Importance of Money versus People. Journal of Human Resources. 2008; 43:884–918.

108. Reimers CW. Labor market discrimination against Hispanic and black men. The Review of Economics and Statistics. 1983; 65:570–579.

109. Westfall, PH.; Young, SS. Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment. New York, NY: John Wiley and Sons; 1993.

110. Lehmann, EL.; Romano, JP. Testing Statistical Hypotheses. ed. 3. New York, NY: Springer Science and Business Media; 2005.

111. Romano JP. Bootstrap and Randomization Tests of Some Nonparametric Hypotheses. Annals of Statistics. 1989; 17:141–159.
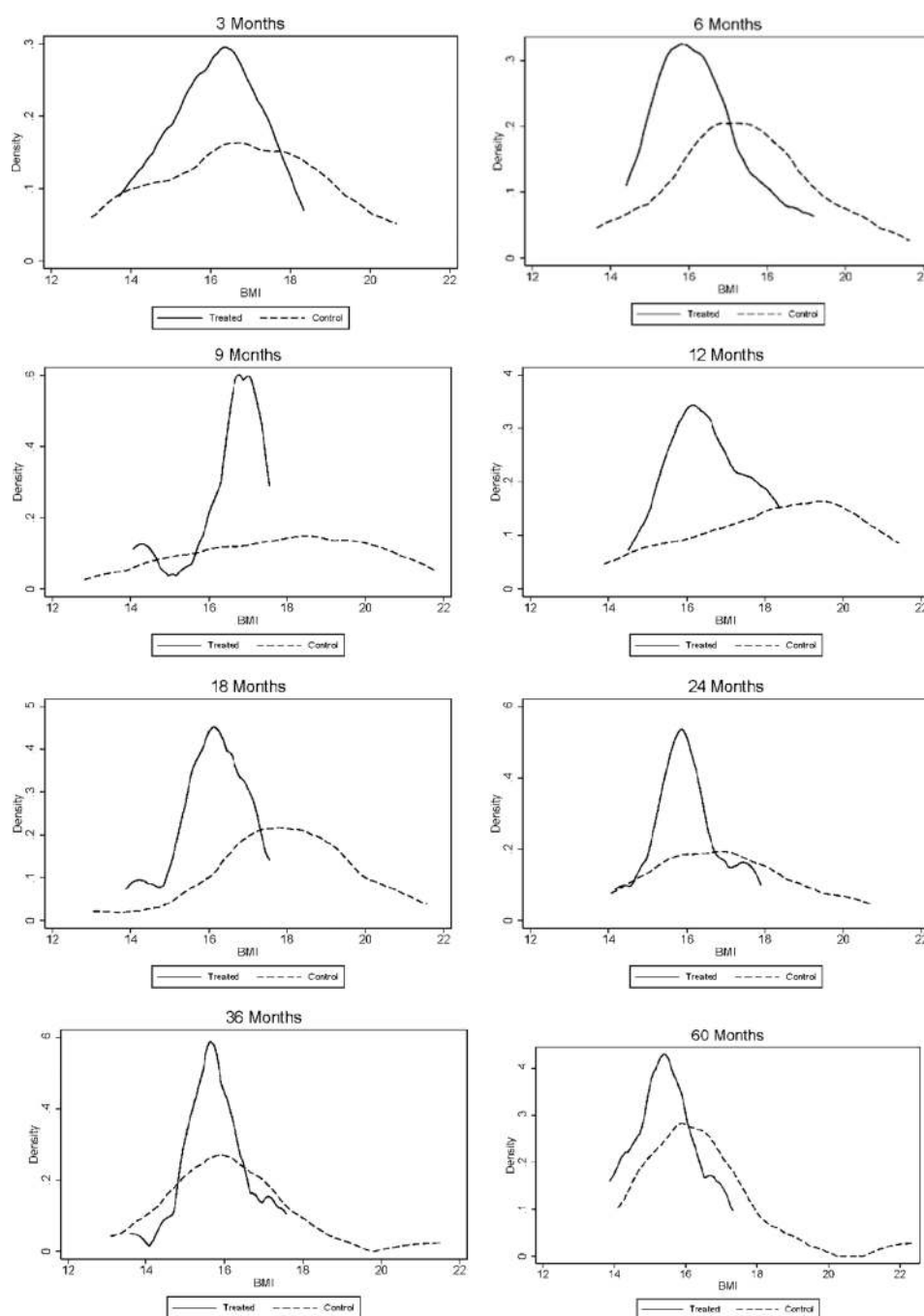
**Fig. 1.**
Body Mass Index (BMI) ages 0–5 by Treatment Status, Males. The black solid line depicts the density for treated males; the black dashed line depicts the density for control males. The graphs display non-parametric kernel estimates of the probability density function based on the Epanechnikov kernel. The kernel $K$ is $K(u) = \frac{3}{4}(1-u^2)\mathbf{1}[|u| \leq 1]$, where $\mathbf{1}[\cdot]$ is an indicator function.
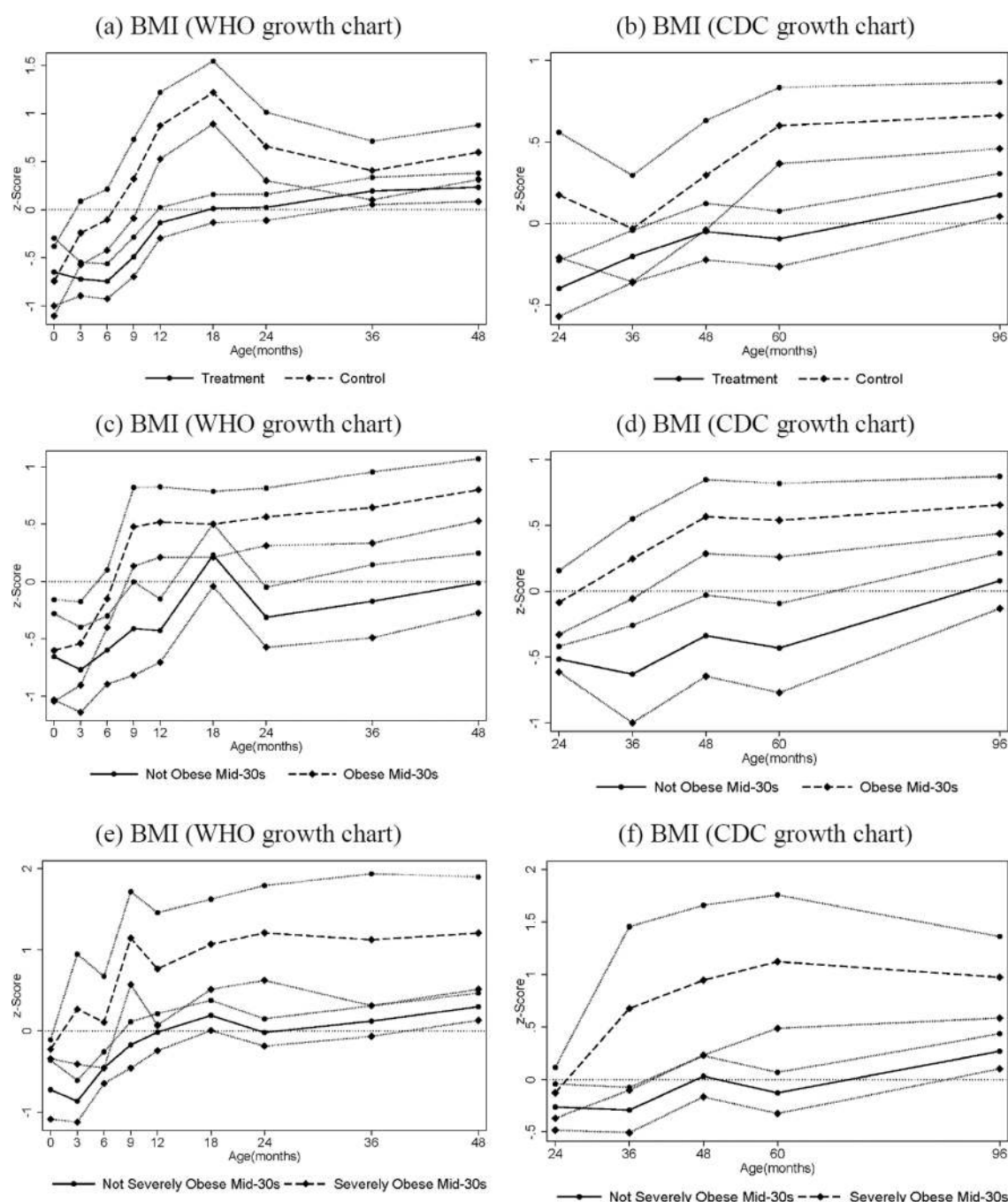
**Fig. 2.**
Body Mass Index ages 0–4 (figures a, c, e) and 2–8 (figures b, d, f), by Treatment and Obesity Status at Mid-30s, Males. The graphs show BMI *z*-scores at different points in childhood (0, 3, 6, 9, 12, 18, 24, 36, 48, 60, and 96 months) by treatment and control status (panels a–b), by obesity status (BMI ≥30) in adulthood (panels c–d), and by severe obesity status (BMI ≥35) in adulthood (panels e–f). Solid and dashed lines represent mean BMI by age for different groups while the bands around each line represent standard errors for the corresponding means (one standard error above and below). Figures (a), (c) and (e) use the

WHO (World Health Organization) growth charts to construct the $z$-scores; figures (b), (d) and (f) use the CDC (Center for Disease Control) growth charts. The CDC recommends the use of the WHO growth charts for less than 2 years of age (see www.cdc.gov/growthcharts/who_charts.htm).
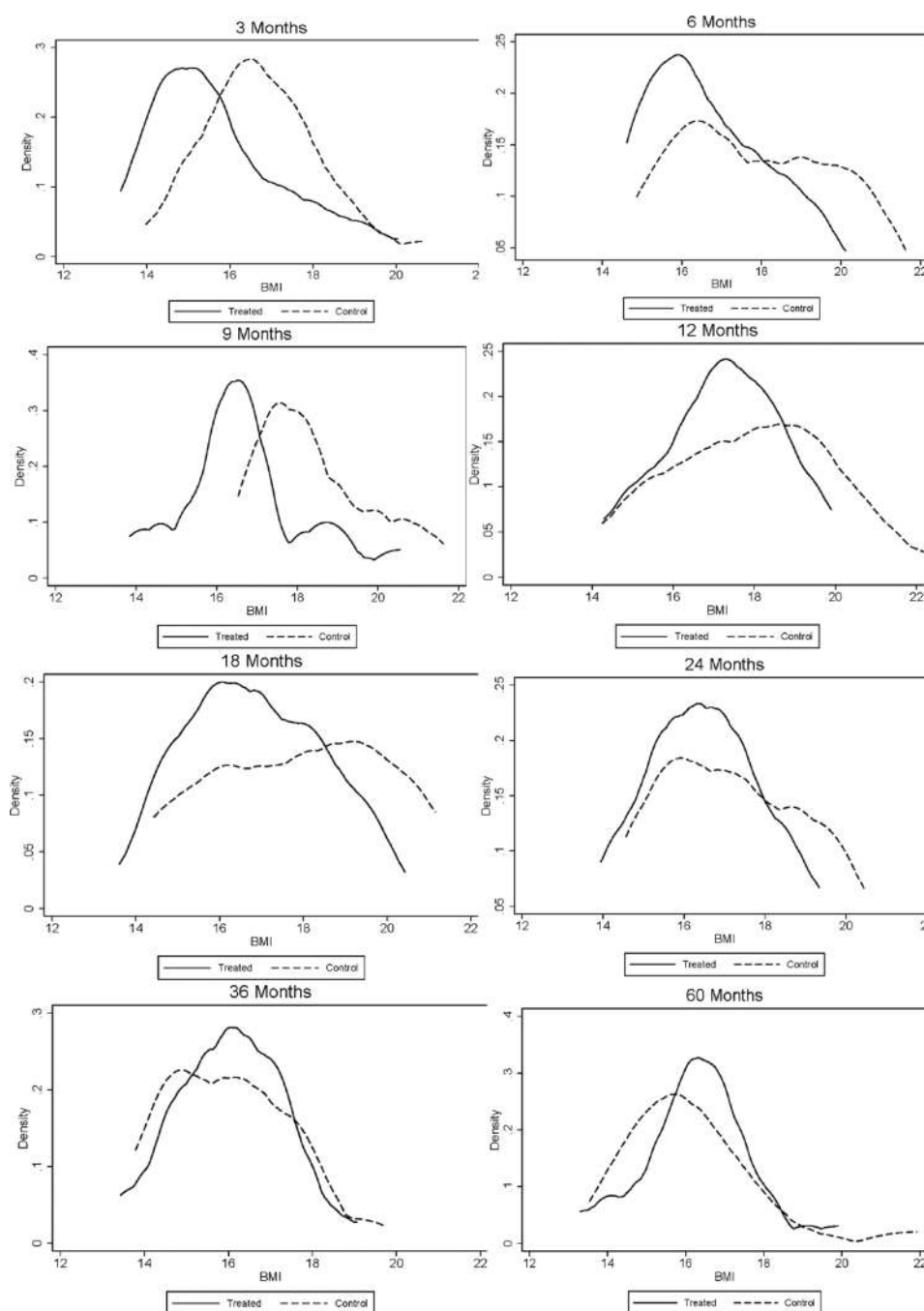
**Fig. 3.**
Body Mass Index (BMI) ages 0–5 by Treatment Status, Females. The black solid line depicts the density for treated males; the black dashed line depicts the density for control males. The graphs display non-parametric kernel estimates of the probability density function based on the Epanechnikov kernel. The kernel $K$ is $K(u) = \frac{3}{4}(1-u^2)\mathbf{1}[|u| \leq 1]$, where $\mathbf{1}[\cdot]$ is an indicator function.
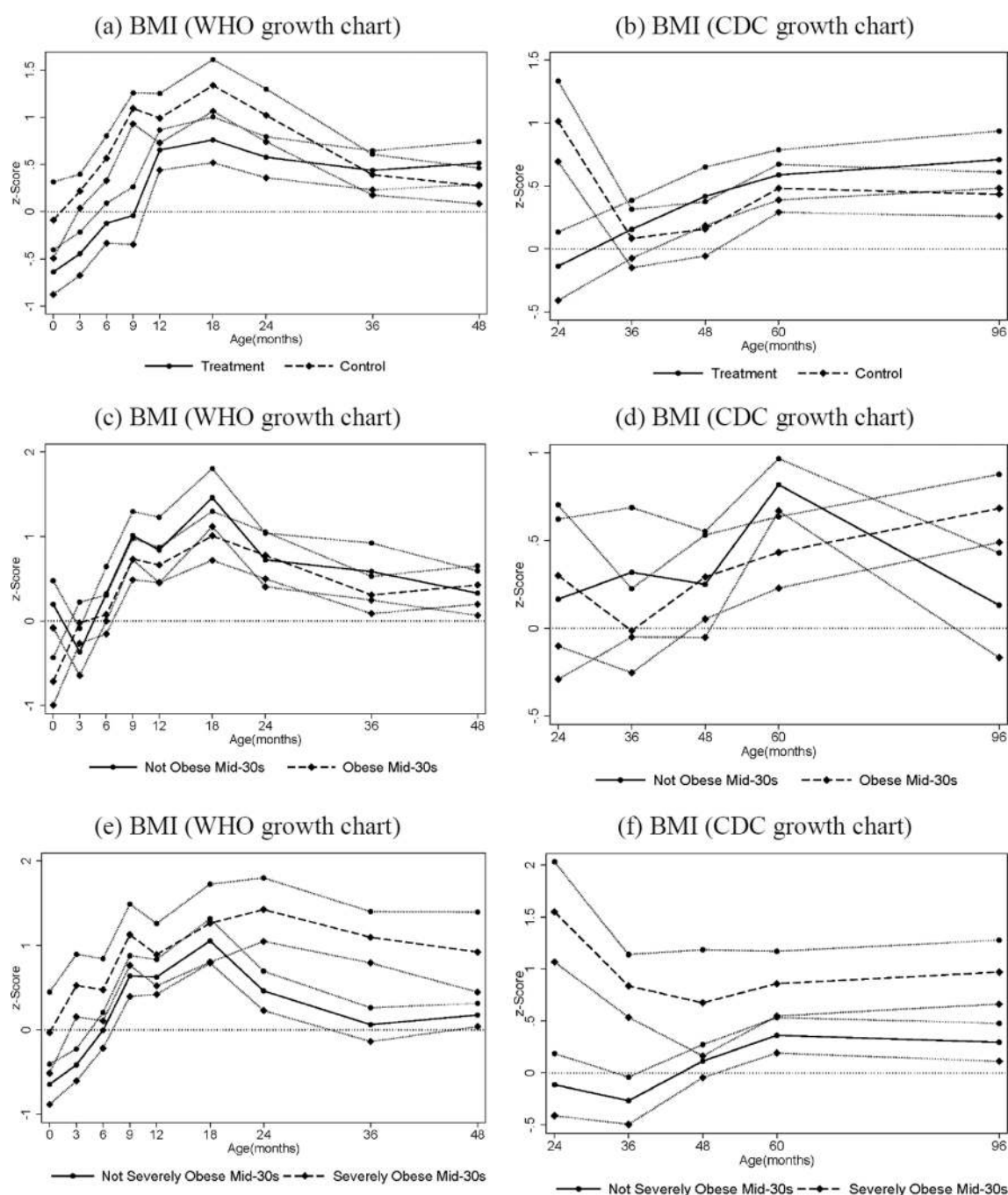
**Fig. 4.**
Body Mass Index ages 0–4 (figures a, c, e) and 2–8 (figures b, d, f), by Treatment and Obesity Status at Mid-30s, Females. The graphs show BMI *z* -scores at different points in childhood (0, 3, 6, 9, 12, 18, 24, 36, 48, 60, and 96 months) by treatment and control status (panels a–b), by obesity status (BMI ≥30) in adulthood (panels c–d), and by severe obesity status (BMI ≥35) in adulthood (panels e –f). Solid and dashed lines represent mean BMI by age for different groups while the bands around each line represent standard errors for the corresponding means (one standard error above and below). Figures (a), (c) and (e) use the

WHO (World Health Organization) growth charts to construct the $z$-scores; figures (b), (d) and (f) use the CDC (Center for Disease Control) growth charts. The CDC recommends the use of the WHO growth charts for less than 2 years of age (see www.cdc.gov/growthcharts/who_charts.htm).

## Table 1

Abecedarian Intervention, Males: Main Health Results, Biomedical Sweep. This table presents the inference and descriptive statistics of selected outcomes of the Abecedarian Intervention. The first column describes the outcome analyzed. The remaining six columns present the statistical analysis. The columns present the following information: (1) control mean; (2) treatment mean; (3) unconditional difference in means across treatment and control groups. We multiply the difference in means by (−1) when a higher value of the variable in the raw data represents a worse outcome so that all outcomes are normalized in a favorable direction (but are not restricted to be positive). (4) conditional treatment effect controlling for cohort, number of siblings, mother's IQ and high-risk index at birth, and accounting for attrition using Inverse Probability Weighting (IPW). Probabilities of IPW are estimated using the following variables: prematurity (gestational age <37 weeks), a dichotomous indicator for not having an exam for illness or injury in the past two years at age 30, Achenbach DSM Attention-Deficit/Hyperactivity (AD/H) problems scale at age 30, and Achenbach substance abuse scale at age 30. The selection of covariates for IPW is based on the lowest Akaike Information Criteria (AIC) among models examining all combinations of covariates that present statistically significant imbalance between attriters and non-attriters. See Supplementary Material Section C and Table S1 for details. (5) one-sided single hypothesis block permutation *p*-value associated with the IPW treatment effect estimate. By block permutation we mean that permutations are done within strata defined by the pre-program variables used in the randomization protocol: cohort, gender, number of siblings, mother's IQ, and high-risk index. (6) Multiple Hypothesis stepdown *p*-values associated with (5). The multiple hypothesis testing is applied to blocks of outcomes. Blocks of variables that are tested jointly using the stepdown algorithm are delineated by horizontal lines. *p* -values ≤0.10 are in bold type. BMI: Body Mass Index; HbA1C: Glycosylated Hemoglobin; NCEP: National Cholesterol Education Program. See Table S11 for complete estimation results.

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| **Blood Pressure** | | | | | | |
| Diastolic Blood Pressure (mmHg) | 92.000 | 78.526 | 13.474 | 19.220 | **0.024** | **0.024** |
| Systolic Blood Pressure (mmHg) | 143.333 | 125.789 | 17.544 | 24.828 | **0.018** | **0.029** |
| Pre-Hypertension (systolic bp ≥120 & diastolic bp ≥80) | 0.667 | 0.421 | 0.246 | 0.321 | 0.119 | 0.172 |
| Pre-Hypertension (systolic bp ≥120 or diastolic bp ≥80) | 0.778 | 0.684 | 0.094 | 0.096 | 0.235 | 0.235 |
| Hypertension (systolic bp ≥140 & diastolic bp ≥90) | 0.444 | 0.105 | 0.339 | 0.537 | **0.010** | **0.018** |
| Hypertension (systolic bp ≥140 or diastolic bp ≥90) | 0.556 | 0.211 | 0.345 | 0.404 | **0.038** | **0.038** |
| **Lab Tests** | | | | | | |
| High-Density Lipoprotein (HDL) Cholesterol (mg/dL) | 42.000 | 53.211 | 11.211 | 11.720 | **0.066** | 0.110 |
| Dyslipidemia (HDL < 40 mg/dL) | 0.417 | 0.106 | 0.311 | 0.255 | 0.179 | 0.179 |

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| Pre-Diabetes (HbA1C ≥5.7%) | 0.583 | 0.473 | 0.110 | 0.043 | 0.426 | 0.426 |
| Vitamin D Deficiency ( < 20 ng/mL) | 0.750 | 0.368 | 0.382 | 0.435 | **0.021** | **0.021** |
| **Obesity** | | | | | | |
| Overweight (BMI ≥25) | 0.750 | 0.722 | 0.028 | 0.190 | 0.239 | 0.239 |
| Obese (BMI ≥30) | 0.625 | 0.556 | 0.069 | 0.211 | 0.233 | 0.345 |
| Severely Obese (BMI ≥35) | 0.375 | 0.111 | 0.264 | 0.404 | 0.115 | 0.232 |
| Waist-Hip Ratio (WHR) | 0.962 | 0.937 | 0.025 | 0.045 | 0.293 | 0.293 |
| Abdominal Obesity (WHR > 0.9) | 0.875 | 0.647 | 0.228 | 0.294 | 0.137 | 0.218 |
| **Multiple Risk Factors** | | | | | | |
| Obesity & Hypertension | 0.500 | 0.111 | 0.389 | 0.529 | **0.016** | **0.016** |
| Severe Obesity & Hypertension | 0.375 | 0.000 | 0.375 | 0.502 | **0.005** | **0.012** |
| Hypertension & Dyslipidemia | 0.333 | 0.000 | 0.333 | 0.435 | **0.006** | **0.012** |
| Metabolic Syndrome (NCEP Definition) | 0.250 | 0.000 | 0.250 | 0.465 | **0.007** | **0.014** |
| Framingham Risk Score (35) | 7.043 | 4.889 | 2.154 | 3.253 | **0.038** | **0.038** |

**Table 2**

Abecedarian Intervention, Females: Main Health Results, Biomedical Sweep. This table presents the inference and descriptive statistics of selected outcomes of the Abecedarian Intervention. The first column describes the outcome analyzed. The remaining six columns present the statistical analysis. The columns present the following information: (1) control mean; (2) treatment mean; (3) unconditional difference in means across treatment and control groups. We multiply the difference in means by (−1) when a higher value of the variable in the raw data represents a worse outcome so that all outcomes are normalized in a favorable direction (but are not restricted to be positive). (4) conditional treatment effect controlling for cohort, number of siblings, mother's IQ and high-risk index at birth, and accounting for attrition using Inverse Probability Weighting (IPW). Probabilities of IPW are estimated using the following variables for the biomedical sweep outcomes: prematurity (gestational age <37 weeks), mother WAIS digit symbol score at recruitment, Achenbach rule breaking problem scale at age 30, and Achenbach substance abuse scale at age 30. The selection of covariates for IPW is based on the lowest Akaike Information Criteria (AIC) among models examining all combinations of covariates that present statistically significant imbalance between attriters and non-attriters. See Supplementary Material Section C and Table S2 for details. (5) one-sided single hypothesis block permutation *p*-value associated with the IPW treatment effect estimate. By block permutation we mean that permutations are done within strata defined by the pre-program variables used in the randomization protocol: cohort, gender, number of siblings, mother's IQ, and high-risk index. (6) Multiple Hypothesis stepdown *p*-values associated with (5). The multiple hypothesis testing is applied to blocks of outcomes. Blocks of variables that are tested jointly using the stepdown algorithm are delineated by horizontal lines. *p*-values ≤0.10 are in bold type. BMI: Body Mass Index; HbA1C: Glycosylated Hemoglobin; NCEP: National Cholesterol Education Program. See Table S12 for complete estimation results.

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| **Blood Pressure** | | | | | | |
| Diastolic Blood Pressure (mmHg) | 89.227 | 85.333 | 3.894 | 1.204 | 0.446 | 0.446 |
| Systolic Blood Pressure (mmHg) | 135.636 | 129.666 | 5.970 | 2.185 | 0.300 | 0.380 |
| Pre-Hypertension (systolic bp ≥120 & diastolic bp ≥80) | 0.727 | 0.500 | 0.227 | 0.101 | 0.222 | 0.222 |
| Pre-Hypertension (systolic bp ≥120 or diastolic bp ≥80) | 0.909 | 0.667 | .242 | 0.244 | **0.042** | **0.069** |
| Hypertension (systolic bp ≥140 & diastolic bp ≥90) | 0.318 | 0.222 | 0.096 | −0.003 | 0.375 | 0.499 |
| Hypertension (systolic bp ≥140 or diastolic bp ≥90) | 0.409 | 0.500 | −0.091 | −0.181 | 0.721 | 0.721 |
| **Lab Tests** | | | | | | |
| High-Density Lipoprotein (HDL) Cholesterol (mg/dL) | 55.318 | 60.444 | 5.126 | 6.002 | 0.143 | 0.143 |

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| Dyslipidemia (HDL < 50 mg/dL) | 0.455 | 0.278 | 0.177 | 0.201 | **0.099** | 0.147 |
| Pre-Diabetes (HbA1C ≥5.7%) | 0.364 | 0.353 | 0.011 | 0.070 | 0.580 | 0.580 |
| Vitamin D Deficiency ( < 20 ng/mL) | 0.727 | 0.722 | 0.005 | 0.048 | 0.303 | 0.303 |
| **Obesity** | | | | | | |
| Overweight (BMI ≥25) | 0.955 | 0.889 | 0.066 | 0.054 | 0.482 | 0.690 |
| Obese (BMI ≥30) | 0.727 | 0.666 | 0.061 | −0.112 | 0.790 | 0.790 |
| Severely Obese (BMI ≥35) | 0.364 | 0.223 | 0.141 | 0.143 | 0.354 | 0.653 |
| Waist-Hip Ratio (WHR) | 0.933 | 0.876 | 0.057 | 0.053 | **0.063** | 0.101 |
| Abdominal Obesity (WHR > 0.85) | 0.762 | 0.563 | 0.199 | 0.198 | **0.080** | **0.080** |
| **Multiple Risk Factors** | | | | | | |
| Obesity & Hypertension | 0.364 | 0.278 | 0.086 | −0.028 | 0.501 | 0.641 |
| Severe Obesity & Hypertension | 0.136 | 0.167 | −0.030 | −0.066 | 0.696 | 0.696 |
| Hypertension & Dyslipidemia | 0.182 | 0.167 | 0.015 | −0.043 | 0.486 | 0.725 |
| Metabolic Syndrome (NCEP Definition) | 0.190 | 0.062 | 0.128 | 0.057 | 0.184 | 0.393 |
| Framingham Risk Score (35) | 1.482 | 1.143 | 0.339 | 0.331 | **0.070** | **0.070** |

**Table 3**

Abecedarian Intervention, Males: Health Care at age 30; Physical Development in Childhood. This table presents the inference and descriptive statistics of selected outcomes of the Abecedarian Intervention. The first column describes the outcome analyzed. The remaining six columns present the statistical analysis. The columns present the following information: (1) control mean; (2) treatment mean; (3) unconditional difference in means across treatment and control groups. We multiply the difference in means by (−1) when a higher value of the variable in the raw data represents a worse outcome so that all outcomes are normalized in a favorable direction (but are not restricted to be positive). (4) conditional treatment effect controlling for cohort, number of siblings, mother's IQ and high-risk index at birth, and accounting for attrition using Inverse Probability Weighting (IPW). The selection of covariates for IPW is based on the lowest Akaike Information Criteria (AIC) among models examining all combinations of covariates that present statistically significant imbalance between attriters and non-attriters. See Supplementary Material Section C and Table S1 for details. (5) one-sided single hypothesis block permutation $p$-value associated with the IPW treatment effect estimate. By block permutation we mean that permutations are done within strata defined by the pre-program variables used in the randomization protocol: cohort, gender, number of siblings, mother's IQ, and high-risk index. (6) Multiple Hypothesis stepdown $p$-values associated with (5). The multiple hypothesis testing is applied to blocks of outcomes. Blocks of variables that are tested jointly using the stepdown algorithm are delineated by horizontal lines. $p$-values ≤0.10 are in bold type. CDC: Center for Disease Control. WHO: World Health Organization. We use weight-for-length ≥85$th$ percentile for being "at-risk overweight" under 24 months, and BMI-for-age ≥85$th$ percentile for being overweight for 24 months and older, see (46). See Table S13 for complete estimation results.

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block $p$-value | Step-Down $p$-value |
|---|---|---|---|---|---|---|
| **Health Care at Age 30** | | | | | | |
| Health Insurance Coverage at age 30 | 0.476 | 0.704 | 0.228 | 0.226 | **0.039** | **0.039** |
| Buys Health Insurance at age 30 | 0.333 | 0.630 | 0.296 | 0.248 | **0.035** | **0.080** |
| Hospital or Doctor Office Care When Sick at age 30 | 0.524 | 0.815 | 0.291 | 0.265 | **0.037** | **0.068** |
| **Physical Development in Childhood** | | | | | | |
| At Risk Overweight (CDC) at 3 months | 0.227 | 0.037 | 0.190 | 0.206 | **0.026** | 0.121 |
| At Risk Overweight (CDC) at 6 months | 0.250 | 0.080 | 0.170 | 0.205 | **0.074** | 0.182 |
| At Risk Overweight (CDC) at 9 months | 0.412 | 0.000 | 0.412 | 0.446 | **0.004** | **0.023** |
| At Risk Overweight (CDC) at 12 months | 0.429 | 0.000 | 0.429 | 0.408 | **0.001** | **0.009** |
| At Risk Overweight (CDC) at 18 months | 0.389 | 0.000 | 0.389 | 0.385 | **0.000** | **0.004** |
| Overweight (CDC) at 24 months | 0.333 | 0.000 | 0.333 | 0.343 | **0.001** | **0.011** |
| Overweight (CDC) at 36 months | 0.158 | 0.080 | 0.078 | 0.094 | 0.194 | 0.194 |

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| Overweight (CDC) at 48 months | 0.300 | 0.167 | 0.133 | 0.133 | 0.150 | 0.235 |
| Overweight (CDC) at 60 months | 0.300 | 0.125 | 0.175 | 0.187 | **0.058** | 0.179 |
| Overweight (CDC) at 96 months | 0.421 | 0.120 | 0.301 | 0.286 | **0.030** | 0.117 |
| Weight-for-Length Change 0–24 months (CDC) | 0.858 | −0.105 | 0.963 | 1.176 | **0.058** | **0.058** |
| Weight-for-Length Change 0–24 months (WHO) | 1.265 | 0.166 | 1.100 | 1.397 | **0.049** | **0.057** |

**Table 4**

Abecedarian Intervention, Females: Health Care at age 30; Physical Development in Childhood. This table presents the inference and descriptive statistics of selected outcomes of the Abecedarian Intervention. The first column describes the outcome analyzed. The remaining six columns present the statistical analysis. The columns present the following information for each gender: (1) control mean; (2) treatment mean; (3) unconditional difference in means across treatment and control groups. We multiply the difference in means by (−1) when a higher value of the variable in the raw data represents a worse outcome so that all outcomes are normalized in a favorable direction (but are not restricted to be positive). (4) conditional treatment effect controlling for cohort, number of siblings, mother's IQ and high-risk index at birth, and accounting for attrition using Inverse Probability Weighting (IPW). The selection of covariates for IPW is based on the lowest Akaike Information Criteria (AIC) among models examining all combinations of covariates that present statistically significant imbalance between attriters and non-attriters. See Supplementary Material Section C and Table S2 for details on the selection procedure and the covariates used for the health care and physical development outcomes. (5) one-sided single hypothesis block permutation *p*-value associated with the IPW treatment effect estimate. By block permutation we mean that permutations are done within strata defined by the pre-program variables used in the randomization protocol: cohort, gender, number of siblings, mother's IQ, and high-risk index. (6) Multiple Hypothesis stepdown *p*-values associated with (5). The multiple hypothesis testing is applied to blocks of outcomes. Blocks of variables that are tested jointly using the stepdown algorithm are delineated by horizontal lines. *p*-values ⩽0.10 are in bold type. CDC: Center for Disease Control. WHO: World Health Organization. We use weight-for-length ≥85 *th* percentile for being "at-risk overweight" under 24 months, and BMI-for-age ≥85*th* percentile for being overweight for 24 months and older, see (46). See Table S14 for complete estimation results.

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| **Health Care at Age 30** | | | | | | |
| Health Insurance Coverage at age 30 | 0.857 | 0.760 | −0.097 | −0.159 | 0.943 | 0.943 |
| Buys Health Insurance at age 30 | 0.357 | 0.400 | 0.043 | −0.027 | 0.511 | 0.810 |
| Hospital or Doctor Office Care When Sick at age 30 | 0.929 | 0.800 | −0.129 | −0.131 | 0.875 | 0.964 |
| **Physical Development in Childhood** | | | | | | |
| At Risk Overweight (CDC) at 3 months | 0.192 | 0.190 | 0.002 | −0.036 | 0.418 | 0.757 |
| At Risk Overweight (CDC) at 6 months | 0.423 | 0.167 | 0.256 | 0.212 | **0.040** | 0.237 |
| At Risk Overweight (CDC) at 9 months | 0.360 | 0.143 | 0.217 | 0.181 | 0.169 | 0.548 |
| At Risk Overweight (CDC) at 12 months | 0.478 | 0.208 | 0.270 | 0.141 | **0.055** | 0.276 |
| At Risk Overweight (CDC) at 18 months | 0.440 | 0.318 | 0.122 | 0.118 | 0.311 | 0.669 |
| Overweight (CDC) at 24 months | 0.412 | 0.174 | 0.238 | 0.195 | 0.143 | 0.517 |

| Variable | Control Mean | Treatment Mean | Difference in Means | Conditional Treatment Effect | Block *p*-value | Step-Down *p*-value |
|---|---|---|---|---|---|---|
| Overweight (CDC) at 36 months | 0.261 | 0.143 | 0.118 | −0.020 | 0.202 | 0.556 |
| Overweight (CDC) at 48 months | 0.192 | 0.409 | −0.217 | −0.247 | 0.944 | 0.944 |
| Overweight (CDC) at 60 months | 0.261 | 0.273 | −0.012 | −0.050 | 0.554 | 0.781 |
| Overweight (CDC) at 96 months | 0.174 | 0.350 | −0.176 | −0.230 | 0.943 | 0.985 |
| Weight-for-Length Change 0–24 months (CDC) | 0.857 | 0.918 | −0.062 | −0.052 | 0.658 | 0.688 |
| Weight-for-Length Change 0–24 months (WHO) | 1.129 | 1.215 | −0.085 | −0.006 | 0.660 | 0.660 |