BMI Trajectories from Adolescence to Adulthood: Evidence from Five Waves of Add Health

Introduction

In 2015 – 2016, about 19% of youth and 40% of adults were obese in the United States (U.S.).[1] There is emerging evidence that upward body mass index (BMI) trajectories begin to form early in life and may track over the long-term,[2,3] and that obese children tend to remain obese in adolescence and even adulthood.[4-9] Studies of recent U.S. cohorts with nationally representative data have shown that obesity tends to persist from birth through early childhood,[10,11] from early through middle childhood,[12-14] and from adolescence to early adulthood.[15,16] Thus, there is evidence of tracking over periods, albeit in a piecewise manner and using different datasets.

Differences in obesity prevalence, incidence, and tracking across sex and race have been previously reported in the literature, with women and people of color being the more vulnerable groups, regardless of the life stage being studied.[12-16] Furthermore, obesity differences between urban and rural residents have also become increasingly prominent in the U.S.[17,18] Obesity in youth is associated with poorer physical health, mental health, and psycho-social well-being in the long run,[19-21] so the combination of high levels of obesity, strong tracking of obesity, and diverging obesity outcomes by demographic characteristics could make already-existing health disparities across population segments even more pronounced throughout the adult years.

In this study, we used Waves I (1994 – 1995, mean age 15) through IV (2008, mean age 28) and an additional sub-sample of Wave V (2016 – 2017, mean age 37) of the National

Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health is one of the few nationally representative prospective studies that can depict obesity dynamics over the developmentally critical period from adolescence to adulthood. Comprising contemporary U.S. data over one of the longest time horizons, it allows for a thorough exploration of obesity dynamics over more than 20 years. Obesity is a major public health problem in the U.S. that has important implications for adult health. Enhancing our understanding of its overall patterns, identifying particularly critical windows from adolescence to adulthood, and identifying especially susceptible demographic subgroups are crucial for prevention and treatment.

Data and Methods

We used data from Add Health, a longitudinal study of a nationally representative sample of adolescents in grades seven through twelve in 1994. Four follow-up waves have been conducted through 2018, though only a sub-sample of the fifth wave is currently available.[1] The mean ages at the five waves were about 15, 16, 22, 28, and 37. These waves allowed for an analysis of more than two decades of data.

At each wave of Add Health, height and weight were self-reported. At Waves II, III, and IV, height and weight were also measured by interviewers. We calculated BMI in kg/m^2, and measured anthropometrics were used where possible. Heights of seven feet or taller and weights of 700 pounds or heavier were coded as missing, due to their biological implausibility. BMI values less than 10 kg/m^2 or greater than 75 kg/m^2 were also coded as missing.

[1] Wave II, by design, was a sub-sample of Wave I and excluded those already in twelfth grade at Wave I. The Wave V sub-sample, consisting of respondents interviewed in 2016 and 2017, was an early release of Wave V. Wave V in its entirety will be a full wave, but the currently available sub-sample should be representative of all the Wave V respondents.

To categorize BMI, BMI z-scores (calculated using the 2000 Centers for Disease Control and Prevention growth reference curves, and adjusted by age and sex) were used for adolescents, and BMI values were used for adults. Overweight was defined as a BMI z-score between the 85^{th} and 95^{th} percentiles for those under 18 years of age, and as a BMI between 25 and 30 for those 18 years of age or older. Obesity was defined as a BMI z-score $\geq 95^{th}$ percentile for those under 18 years of age, and as a BMI ≥ 30 for those 18 years of age or older. All other adolescents/adults were considered to have a normal BMI.

Transition probabilities between consecutive waves were calculated to ascertain the likelihood of changing BMI categories. People could have normal, overweight, or obese BMI status in any given wave, resulting in nine possibilities between consecutive waves. These were grouped into three main transitions of interest – staying in the same BMI category, moving to a higher BMI category, and moving to a lower BMI category. Changes in continuous BMI were also calculated to determine how many BMI points, on average, people gained or lost between waves. Because of the differences in duration between waves, these changes in continuous BMI were annualized to facilitate comparison. Looking at shifts in categorical and continuous BMI between consecutive waves allowed for us to zero in on shorter windows within the period from adolescence to adulthood.

To more broadly examine overall changes from adolescence to adulthood, we used generalized estimating equations (GEE), which consider the longitudinal nature of the data and allow for correlation between observations of the same subject.[22-24] These population-average models were used to explore differences in long-term trajectories by various demographic subgroups. The demographic characteristics studied here were wave (I = reference, II, III, IV, V), sex (male = reference, female), race (non-Hispanic white = reference, non-Hispanic black, Hispanic, Asian, other), place of residence at Wave I (rural = reference, suburban, urban), and parents' education at Wave I (neither parent graduated from college = reference, one parent graduated from college, both parents graduated from college).[2] While place of residence and parents' education could be time-varying, what is of interest here was how an adolescent's contextual characteristics could impact future BMI outcomes. An indicator for obesity was used as the dependent variable to study the likelihood of obesity of various population subgroups from adolescence to adulthood.

We then broke the population into six different groups based on their BMI trajectories across the first four waves of Add Health.[3] BMI was condensed into two categories – non-obese (normal and overweight) and obese. The first trajectory group consisted of those who were non-obese at all four waves. The second group consisted of those who were non-obese at Wave I, who made exactly one change in classification and were obese by Wave IV. The third group consisted of those who were non-obese at Wave I, and fluctuated at least twice between non-obese and obese across waves. Trajectory groups four, five, and six were defined similarly – those who were obese at all four waves, those who started as obese and became non-obese, and those who started as obese and fluctuated between obese and non-obese, respectively.

Among the three trajectory groups that started as non-obese at Wave I, there was a natural ordering of groups – always non-obese, non-obese and fluctuating, and non-obese to obese. An ordinal logistic regression was run using this ordinal categorical variable as the dependent variable, and with the same covariates as in the GEE model. A corresponding ordinal

---

[2] In Wave I, the respondent's parent was asked "How far did you go in school?" and "How far did your current (spouse/partner) go in school?" Even though the current spouse/partner might not have been the respondent's biological parent, this person was currently in the household with the respondent.
[3] The Wave V sub-sample was excluded because there were too few people with available data.

logistic regression was run for those who started off as obese at Wave I, with the ordering obese to non-obese, obese and fluctuating, and always obese.

For these analyses, the complex survey design of Add Health was taken into account and survey weights were used where possible.[4] For the descriptive statistics by wave, survey weights for the corresponding wave were used. For the models, longitudinal survey weights were used. Listwise deletion was used for observations with missing data. All analyses were performed in Stata 15.1.

Results

Table 1 shows the BMI transitions between waves of Add Health. The sub-tables display the proportion of people who transitioned between BMI categories in consecutive waves or from the beginning to the end of the survey period. The numbers along the main diagonal represent staying in the same BMI category, those shaded in red and above the main diagonal represent gaining weight to a higher BMI category, and those shaded in blue and below the main diagonal represent losing weight to a lower BMI category.

Table 1: Percentage of people who experienced each BMI transition between waves

(n = 14,133)		Wave II * (mean age 16) BMI category			
		Normal	Overweight	Obese	Row totals
Wave I (mean age 15) BMI category	Normal	68.75	5.55	0.79	75.09
	Overweight	3.98	7.49	3.00	14.47
	Obese	0.59	1.71	8.14	10.44
	Column totals	73.32	14.75	11.93	100.00
(n = 10,862)		Wave III (mean age 22) BMI category			
		Normal	Overweight	Obese	Row totals
Wave II *	Normal	47.92	19.82	5.39	73.13

[4] Survey weights were not available for the Wave V sub-sample.

(mean age 16) BMI category	Overweight	1.94	6.10	6.93	14.97
	Obese	0.37	1.66	9.87	11.90
	Column totals	50.23	27.58	22.19	100.00
(n = 13,964)		Wave III (mean age 22) BMI category			
		Normal	Overweight	Obese	Row totals
Wave I (mean age 15) BMI category	Normal	46.93	21.02	6.90	74.85
	Overweight	1.96	5.50	7.52	14.98
	Obese	0.51	1.39	8.27	10.17
	Column totals	49.40	27.91	22.69	100.00
(n = 12,210)		Wave IV (mean age 28) BMI category			
		Normal	Overweight	Obese	Row totals
Wave III (mean age 22) BMI category	Normal	30.25	14.64	3.54	48.43
	Overweight	3.09	13.05	11.65	27.79
	Obese	0.32	2.31	21.16	23.79
	Column totals	33.66	30.00	36.35	100.00
(n = 15,058)		Wave IV (mean age 28) BMI category			
		Normal	Overweight	Obese	Row totals
Wave I (mean age 15) BMI category	Normal	32.23	25.56	16.17	73.96
	Overweight	1.08	3.74	10.43	15.25
	Obese	0.22	1.00	9.57	10.79
	Column totals	33.53	30.30	36.17	100.00
(n = 3,462)		Wave V * (mean age 37) BMI category			
		Normal	Overweight	Obese	Row totals
Wave IV (mean age 28) BMI category	Normal	22.56	10.02	0.90	33.48
	Overweight	4.16	16.12	9.24	29.52
	Obese	0.87	5.49	30.65	37.01
	Column totals	27.59	31.63	40.79	100.00
(n = 3,722)		Wave V * (mean age 37) BMI category			
		Normal	Overweight	Obese	Row totals
Wave I (mean age 15) BMI category	Normal	26.92	27.40	20.28	74.60
	Overweight	1.10	3.68	9.83	14.61
	Obese	0.30	1.02	9.46	10.78
	Column totals	28.32	32.10	39.57	100.00

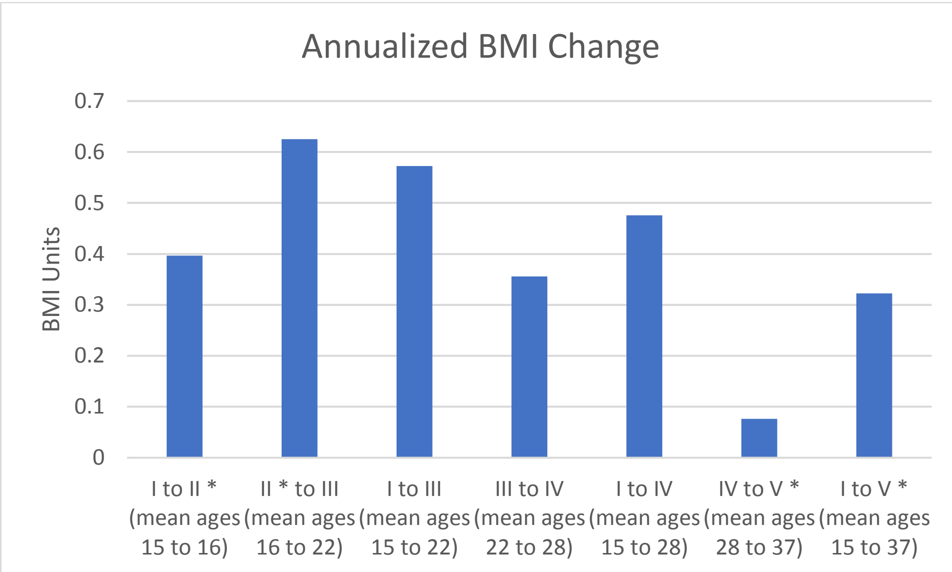
Note: Sub-samples are denoted by a *.

First, we look at the BMI category transitions from Wave I (mean age 15) to Wave V (mean age 37). At Wave I, about 75%, 15%, and 11% of people were classified in the normal, overweight, and obese BMI categories, respectively. By Wave V, the corresponding percentages were about 28%, 32%, and 40%. The percentage of obese people increased by more than four-fold. Within this time frame, almost 60% of people had jumped to a higher BMI category, yet only about 2% of people had fallen to a lower BMI category.

We then look at the transitions between consecutive waves. Between consecutive waves, it was most common to stay in the same category, followed by moving to a higher BMI category, and with moving to a lower BMI category the least common. Despite this, over the course of the entire survey, the majority of people eventually moved upwards to a higher BMI category. Thus, even though the percentage of people who moved upwards to a higher BMI category between any two consecutive waves was never particularly high, with the highest being 35% between Waves I and III (mean ages 15 and 22), the majority moved to a higher BMI category over the course of the two decades.

In Figure 1, we present the annualized BMI change both between consecutive waves and over the entire span of Add Health.

Figure 1: Annualized BMI change between adolescence and adulthood

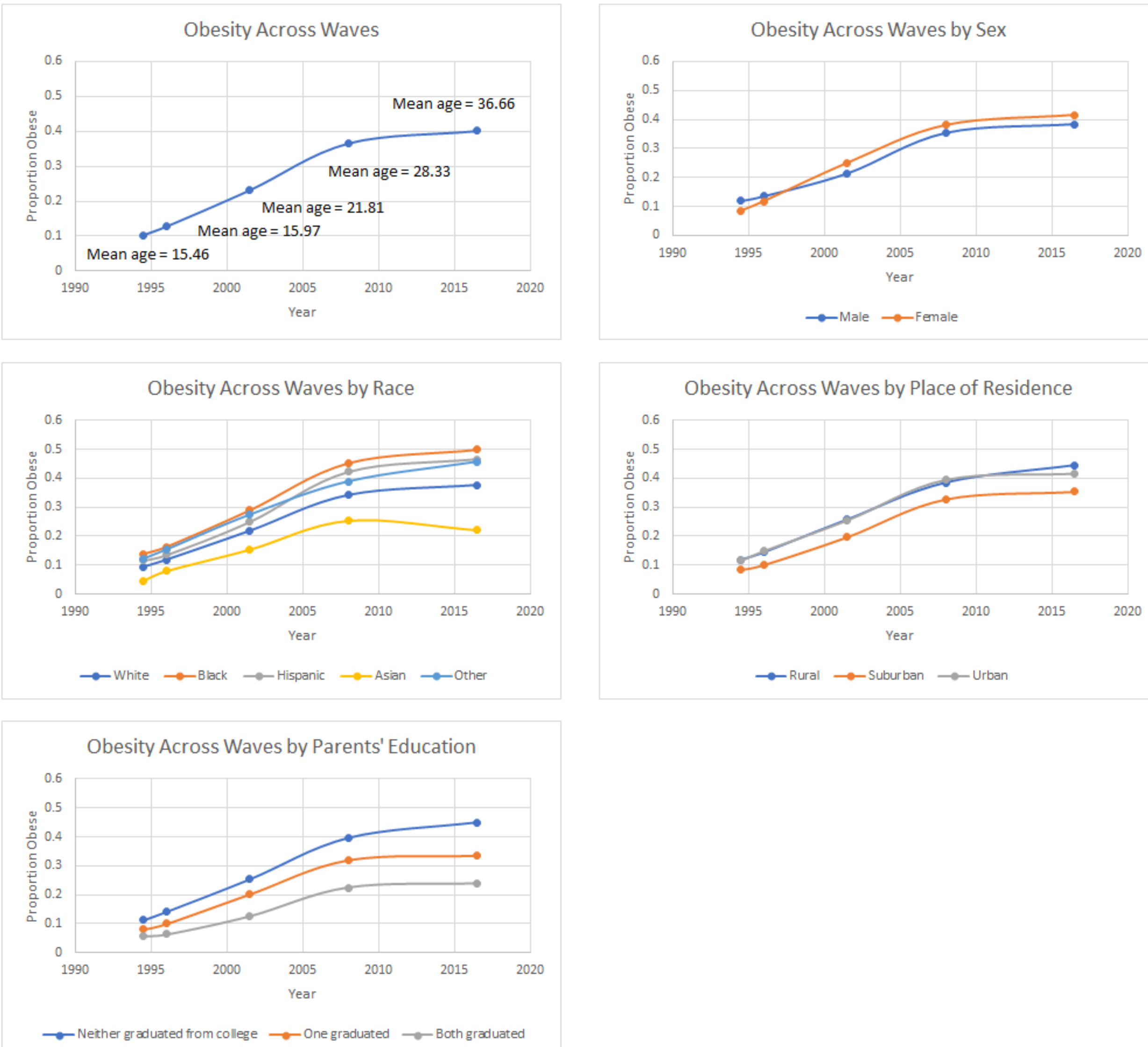


Note: Sub-samples are denoted by a *.

Across the time span of Add Health, individuals on average gained about 0.3 kg/m^2 in BMI annually. However, we should be cautious in interpreting the BMI change between Waves I and II (mean ages 15 and 16) when many youths were still experiencing growth, especially since the oldest members of the Wave I group (the high school seniors) were excluded. Of note, BMI change seemed most pronounced between Waves II and III (mean ages 16 and 22), with smaller changes between Waves III and IV (mean ages 22 and 28) and Waves IV to V (mean ages 28 and 37). It appears that people continuously gained BMI over the two decades, though the rate of change substantially slowed down by the time people were in their 30s.

Figure 2 shows the trajectories of proportion of people obese.

Figure 2: Trajectories of proportion obese across waves



In almost all these graphs, the trends were increasing and concave. That is, proportion obese increased over the two decades, but at a decreasing rate. The general patterns were similar across the subgroups, with some clear distinctions worth noting. Males appeared to start off with a higher proportion obese, but females overtook them early on in this time frame. Non-Hispanic blacks consistently had the highest proportion obese, followed by Hispanics, individuals of other races, non-Hispanic whites, and Asians. Interestingly, Asians were the only subgroup for which the trends in proportion obese actually decreased. The rural and urban subgroups followed similar trajectories until the last wave in which there was some divergence. The level among those in the suburban subgroup was lower than those of the other two residence subgroups. A larger divergence could be seen when stratifying by parental education, with those who had two parents without a college degree at baseline having the highest proportion obese, followed by those who had one parent with a college degree and those who had both parents with a college degree.

To explore these patterns taking all of these characteristics into account collectively, we used GEE. The variables included in the models were those used for stratification in Figure 2, along with survey wave, which was treated as a categorical variable. In Table 2, we display the coefficients from the GEE model with an indicator for obesity as the dependent variable.

Table 2: Coefficients from the GEE model of the obesity indicator as outcome variable on demographic characteristics, Waves I to V (64,704 observations for 19,503 people)

Variable	Coefficient
Wave (reference = I, mean age 15)	
II, mean age 16	0.188 ***
III, mean age 22	0.948 ***
IV, mean age 28	1.575 ***
V, mean age 37	1.733 ***
Sex (reference = male)	
Female	0.061 *
Race (reference = non-Hispanic white)	
Non-Hispanic black	0.318 ***
Hispanic	0.263 ***
Asian	–0.233 ***
Other	0.278 ***
Place of residence (reference = rural)	
Suburban	–0.289 ***
Urban	–0.093 *
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.297 ***
Both parents graduated from college	–0.616 ***
Constant	–2.033 ***

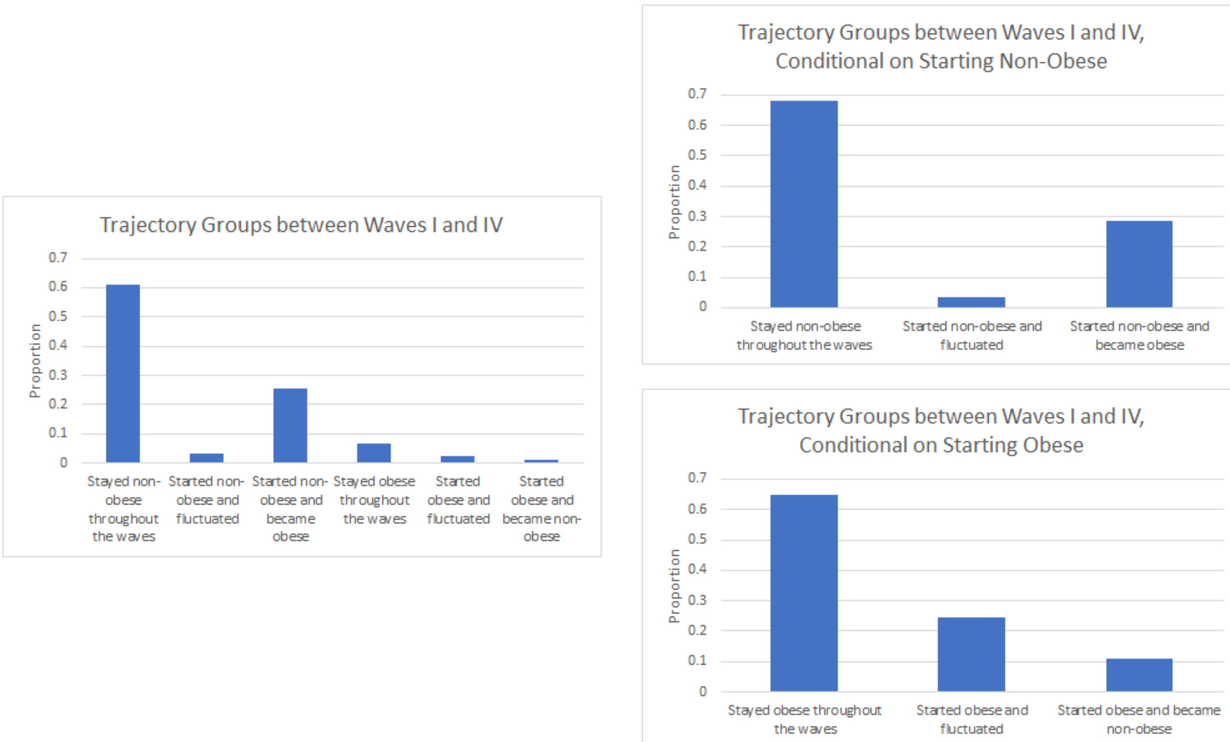
Note: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

The estimated coefficients of Waves II through V were all significant and positive, and they became increasingly positive with each subsequent wave. Thus, the risk of becoming obese was significantly different at Waves II through V as compared to Wave I, and the log odds for obesity increased with each wave. The differences in magnitudes between waves were also telling. Generally, the jumps between consecutive waves decreased, both in absolute and in relative terms, as the survey participants aged. Furthermore, the log odds were pairwise significantly different, as determined by contrasts.

The difference between males and females was only marginally significant at the five-percent level, and females had a higher propensity of obesity than males. The non-Hispanic black, Hispanic, and other groups had a higher risk of obesity, while the Asian group had a lower risk, relative to the non-Hispanic white group. However, the non-Hispanic black, Hispanic, and other groups were not significantly different from each other. Suburban dwellers had a lower likelihood of obesity than both rural and urban dwellers, while urban residents had a slightly but still statistically significantly lower likelihood than rural residents. Having one parent who graduated from college was associated with a lower likelihood of becoming obese relative to having neither parents being college graduates, while having both parents who graduated college was associated with an even lower likelihood of becoming obese.

The above results tell us how these contextual characteristics were associated with the propensity of obesity. However, there are several trajectories that people could take to obesity, and it is important to understand how such a large proportion of the U.S. ended up with such elevated levels of BMI. Figure 3 depicts the proportion of people in each BMI trajectory group, as well as the proportion of people in each group conditional on whether they started as non-obese or obese at Wave I.

Figure 3: BMI trajectory groups between Waves I and IV, overall and conditional on starting BMI category in Wave I



The highest proportion of people were in the "stayed non-obese throughout the waves group." Conditional on one's starting category, staying within that category was the most common. Among those who were non-obese at Wave I, becoming obese was more common than fluctuating. On the other hand, among those who were obese at Wave I, fluctuating was more common than becoming non-obese. Table 3 displays the number of people who followed each of the trajectories, in the form of Table 1. Since those who fluctuated could have a BMI category of non-obese or obese in Wave IV, the fluctuation BMI trajectory groups are split into two across the columns.

Table 3: Transitions between Wave I to IV, based on the trajectory groups

		Wave IV (mean age 28) BMI category	
		Non-obese (normal + overweight)	Obese
Wave I (mean age 15) BMI category	Non-obese (normal + overweight)	Stayed non-obese: 5621 Non-obese but fluctuated: 227	Non-obese to obese: 2322 Non-obese but fluctuated: 62
	Obese	Obese to non-obese: 93 Obese but fluctuated: 12	Stayed obese: 613 Obese but fluctuated: 210

Next, we explored how the contextual characteristics used in the GEE model were associated with the trajectories from Waves I to IV, conditional on the baseline category. The ordinal logistic regression results for those who were non-obese at Wave I are presented in Table 4, and the corresponding results for those who were obese at Wave I are presented in Table 5.

Table 4: Ordinal logistic regression of trajectory groups among those who started out in the non-obese BMI category on demographic characteristics, Waves I to IV (n = 7,331)

Variable	Coefficient
Sex (reference = male)	
Female	0.188 **
Race (reference = non-Hispanic white)	
Non-Hispanic black	0.303 ***
Hispanic	0.218 *
Asian	–0.303
Other	0.117
Place of residence (reference = rural)	
Suburban	–0.166 *
Urban	0.001
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.209 *
Both parents graduated from college	–0.572 ***
Cut 1	0.743
Cut 2	0.915

Notes: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Categories from lowest to highest: stayed non-obese throughout the waves, started non-obese and fluctuated, started non-obese and became obese.

Brant test p-value was 0.893, fail to reject the proportional odds assumption.

The reference category was staying non-obese throughout the waves and the other groups represented higher categories, or more severe trajectories. The ordering of categories from low to high was stayed non-obese, started non-obese and fluctuated, and started non-obese and became obese. Females were more likely to be in a higher category than males, non-Hispanic blacks and Hispanics were more likely to be in a higher category than non-Hispanic whites, suburban residents were less likely to be in a higher category than rural dwellers, and those with more educated parents were less likely to be in a higher category than those who had neither parent graduate from college. These results are similar to the ones from the GEE model. That is, similar characteristics were associated in the same direction with both the time-varying obesity indicator and the likelihood of being in more severe BMI trajectories, conditional on starting off as non-obese.

Table 5: Ordinal logistic regression of trajectory groups among those who started out in the obese BMI category on demographic characteristics, Waves I to IV (n = 817)

Variable	Coefficient
Sex (reference = male)	
Female	0.407
Race (reference = non-Hispanic white)	
Non-Hispanic black	–0.020
Hispanic	–0.201
Asian	0.701
Other	0.099
Place of residence (reference = rural)	
Suburban	–0.109
Urban	0.074
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.168
Both parents graduated from college	–0.501
Cut 1	–2.001
Cut 2	–0.517

Notes: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Categories: started obese and became non-obese, started obese and fluctuated, stayed obese throughout the waves.

Brant test p-value was 0.624, fail to reject the proportional odds assumption.

The ordering of categories from low to high here was started obese and became non-obese (reference category), started obese and fluctuated, and stayed obese. None of the results in Table 4 were significant. That is, none of these contextual variables were associated with group trajectories among those who were obese at Wave I, in stark contrast to the results from the ordinal logistic regression of BMI trajectories conditional on starting off as non-obese.

Discussion

This paper explores BMI transitions and trajectories over more than two decades, during the transformative period from adolescence to adulthood. Between Waves I and V (mean ages 15 and 37) of Add Health, obesity increased four-fold. The majority of those who had a normal

BMI in adolescence became overweight or obese by adulthood and very few youths who were overweight or obese in adolescence achieved normal BMI by adulthood. However, BMI was not increasing at a constant rate throughout this duration. BMI trajectories witnessed an increasing but concave pathway over time.

These trends were more problematic for some demographic subgroups than for others. Sex, race, place of residence, and parental educational attainment were all significant in the GEE model. Females were more likely to be obese than males, though the differences between sexes was only marginally significant. Non-Hispanic blacks and Hispanics had the highest propensity for obesity, followed by non-Hispanic whites, and Asians. Interestingly, Asians actually saw a decrease between Waves IV and V (mean ages 28 and 37). This was the only subgroup and the only pair of consecutive waves for which a decrease took place. Those residing in suburban areas had a lower propensity for obesity than both those in rural and those in urban areas, and those in urban areas had a marginally lower propensity for obesity than those in rural areas. Those with educated parents also had a lower propensity for obesity.

While similar variables were significant for the ordinal logistic regression exploring the likelihood of being a in a certain BMI trajectory, conditional on starting in the non-obese category, no variables were significant for the corresponding model exploring the likelihood of being in a certain BMI trajectory, conditional on starting in the obese category. That is, there are some demographic subgroups which seem to be targetable for prevention among those who were not obese in adolescence, but it is difficult to target subgroups for management or reversal among those who were already obese. This difference could have been an issue of sample size, since there was a much larger proportion of people who started off as non-obese at Wave I, the ratio of the non-obese group to the obese group at Wave I being about 9:1. As a sensitivity check, the BMI classifications were split between normal versus non-normal (overweight or obese) instead, to increase the sample size of the group that started off in the heavier category and to reduce the ratio of the lighter to heavier groups to about 3:1. However, substantive conclusions were similar; for the two ordinal logistic regressions, variables generally registered the same significance results.

Other variable combinations were assessed. Wave was used as a categorical variable in our GEE model. Other models replacing wave by (1) age or (2) age and age-squared were also tested. The quasi-information criteria (QIC) were calculated for these models, and models using wave or both age and age-squared were superior to models using age. This highlighted the importance of the curvature of the BMI trajectories, as using age linearly resulted in the worst model. We ultimately opted for wave as a categorical variable for ease of interpretation.

Behavioral variables from Wave I were also included in the GEE model and the ordinal logistic regressions. These behavioral variables were screen time (hours watching television, watching videos, and playing computer or video games), regular exercise (at least three times a week), hours of sleep per night, and whether respondents did not eat breakfast. Results from these models are included in the Appendix. While some of these variables registered significance, we have opted not to include them in the main results, as these variables were not as robust as the contextual variables and were more likely to change over time. In addition, their values at Wave I many years ago might not have a lasting influence on BMI. However, it is noteworthy that the substantive conclusions of the contextual characteristics remain the same, with or without the inclusion of such behavioral characteristics.

Aside from variable sensitivity checks, various models were tested as well. The GEE results presented here used the exchangeable correlation structure, but different correlation structures were tested for sensitivity. Additionally, corresponding GEE models were run with continuous BMI as a dependent variable, instead of an indicator variable for obesity. In both cases, significance conclusions and magnitudes of the coefficient estimates were similar across models. In addition, although the Brant test did not reject the proportional odds assumption for our ordinal logistic regression models, we also relaxed this assumption and ran multinomial logistic regression models for our two trajectory groups conditional on baseline classification. Significance conclusions and magnitudes of the coefficient estimates were similar in corresponding models. The ordinal logistic regression models were presented here for ease of interpretation.

There are a couple of caveats that should be noted. First, there was no question consistently asked across the waves of Add Health regarding the survey participants' pregnancy status. Women pregnant during a survey might have had their BMIs over-estimated. Second, although some of the BMI trajectory groups involved fluctuations across waves, those in the other trajectory groups might have also seen fluctuations as well; it might just be that these fluctuations were not observable at the times that the surveys were conducted. More and frequent waves might improve the accuracy of these groupings. When the full sample of Wave V is released, we plan on including its data in the BMI trajectory group classifications.

Conclusion

While there are national estimates on the prevalence of obesity and studies documenting the characteristics associated with obesity, less is known about its dynamics in the years from adolescence to adulthood. Obesity has long-term social, economic, and health implications, so a better understanding of its dynamics is crucial. In an effort to explore such dynamics from approximately ages 15 to 37, we analyzed BMI transitions, changes, and trajectories across five waves of a nationally representative longitudinal survey of adolescent and adult health, with aims of studying overall changes across the duration, pinpointing certain windows as being particularly critical, and identifying population segments most susceptible to becoming obese.

References

1. Hales CM, Margaret D. Carroll, Cheryl D. Fryar, Cynthia L. Ogden. *Prevalence of Obesity Among Adults and Youth: United States, 2015–2016.* 2017. 288.
2. Power C, Lake JK, Cole TJ. Measurement and long-term health risks of child and adolescent fatness. *Int J Obes Relat Metab Disord.* 1997;21(7):507-526.
3. Serdula MK, Ivery D, Coates RJ, Freedman DS, Williamson DF, Byers T. Do obese children become obese adults? A review of the literature. *Prev Med.* 1993;22(2):167-177.
4. Herman KM, Craig CL, Gauvin L, Katzmarzyk PT. Tracking of obesity and physical activity from childhood to adulthood: the Physical Activity Longitudinal Study. *Int J Pediatr Obes.* 2009;4(4):281-288.
5. Johannsson E, Arngrimsson SA, Thorsdottir I, Sveinsson T. Tracking of overweight from early childhood to adolescence in cohorts born 1988 and 1994: overweight in a high birth weight population. *Int J Obes (Lond).* 2006;30(8):1265-1271.
6. Magarey AM, Daniels LA, Boulton TJ, Cockington RA. Predicting obesity in early adulthood from childhood and parental obesity. *Int J Obes Relat Metab Disord.* 2003;27(4):505-513.
7. Niclasen BV, Petzold MG, Schnohr C. Overweight and obesity at school entry as predictor of overweight in adolescence in an Arctic child population. *Eur J Public Health.* 2007;17(1):17-20.
8. Williams S. Overweight at age 21: the association with body mass index in childhood and adolescence and parents' body mass index. A cohort study of New Zealanders born in 1972-1973. *Int J Obes Relat Metab Disord.* 2001;25(2):158-163.
9. Wright CM, Emmett PM, Ness AR, Reilly JJ, Sherriff A. Tracking of obesity and body fatness through mid-childhood. *Arch Dis Child.* 2010;95(8):612-617.
10. Mei Z, Grummer-Strawn LM, Scanlon KS. Does overweight in infancy persist through the preschool years? An analysis of CDC Pediatric Nutrition Surveillance System data. *Soz Praventivmed.* 2003;48(3):161-167.
11. Stettler N, Zemel BS, Kumanyika S, Stallings VA. Infant weight gain and childhood overweight status in a multicenter, cohort study. *Pediatrics.* 2002;109(2):194-199.
12. Cunningham SA, Kramer MR, Narayan KM. Incidence of childhood obesity in the United States. *N Engl J Med.* 2014;370(5):403-411.
13. Li C, Goran MI, Kaur H, Nollen N, Ahluwalia JS. Developmental trajectories of overweight during childhood: role of early life factors. *Obesity (Silver Spring).* 2007;15(3):760-771.
14. Rendall MS, Weden MM, Fernandes M, Vaynman I. Hispanic and black US children's paths to high adolescent obesity prevalence. *Pediatr Obes.* 2012;7(6):423-435.
15. Gordon-Larsen P, Adair LS, Nelson MC, Popkin BM. Five-year obesity incidence in the transition period between adolescence and adulthood: the National Longitudinal Study of Adolescent Health. *Am J Clin Nutr.* 2004;80(3):569-575.

16. Gordon-Larsen P, The NS, Adair LS. Longitudinal trends in obesity in the United States from adolescence to the third decade of life. *Obesity (Silver Spring).* 2010;18(9):1801-1804.
17. Befort CA, Nazir N, Perri MG. Prevalence of obesity among adults from rural and urban areas of the United States: findings from NHANES (2005-2008). *J Rural Health.* 2012;28(4):392-397.
18. Hill JL, You W, Zoellner JM. Disparities in obesity among rural and urban residents in a health disparate region. *BMC Public Health.* 2014;14:1051.
19. Dietz WH. Health consequences of obesity in youth: childhood predictors of adult disease. *Pediatrics.* 1998;101(3 Pt 2):518-525.
20. Must A, Strauss RS. Risks and consequences of childhood and adolescent obesity. *Int J Obes Relat Metab Disord.* 1999;23 Suppl 2:S2-11.
21. Reilly JJ, Kelly J. Long-term impact of overweight and obesity in childhood and adolescence on morbidity and premature mortality in adulthood: systematic review. *Int J Obes (Lond).* 2011;35(7):891-898.
22. Hanley JA, Negassa A, Edwardes MD, Forrester JE. Statistical analysis of correlated data using generalized estimating equations: an orientation. *Am J Epidemiol.* 2003;157(4):364-375.
23. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics.* 1986;42(1):121-130.
24. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. *Biometrics.* 1988;44(4):1049-1060.

Appendix

Table 1: Coefficients from the GEE model of the obesity indicator on demographic and behavioral characteristics, Waves I to V (63,696 observations for 19,174 people)

Variable	Coefficient
Wave (reference = I, mean age 15)	
II, mean age 16	0.191 ***
III, mean age 22	0.956 ***
IV, mean age 28	1.592 ***
V, mean age 37	1.742 ***
Sex (reference = male)	
Female	0.045
Race (reference = non-Hispanic white)	
Non-Hispanic black	0.288 ***
Hispanic	0.268 ***
Asian	–0.267 ***
Other	0.256 **
Place of residence (reference = rural)	
Suburban	–0.288 ***
Urban	–0.091 *
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.294 ***
Both parents graduated from college	–0.575 ***
Hours of screen time	0.005 ***
Regular exercise (reference = no)	
Yes	–0.025
Hours of sleep	–0.013
No regular breakfast (reference = yes)	
No	0.440 ***
Constant	–2.125 ***

Note: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Table 2: Ordinal logistic regression of trajectory groups among those who started out in the non-obese BMI category on demographic and behavioral characteristics, Waves I to IV (n = 7,236)

Variable	Coefficient
Sex (reference = male)	
Female	0.186 **
Race (reference = non-Hispanic white)	
Non-Hispanic black	0.267 **
Hispanic	0.237 *
Asian	–0.329
Other	0.097
Place of residence (reference = rural)	
Suburban	–0.168 *
Urban	–0.019
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.190 *
Both parents graduated from college	–0.529 ***
Hours of screen time	0.006 **
Regular exercise (reference = no)	
Yes	–0.015
Hours of sleep	–0.027
No regular breakfast (reference = yes)	
No	0.352 ***
Cut 1	0.708
Cut 2	0.882

Notes: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Categories from lowest to highest: stayed non-obese throughout the waves, started non-obese and fluctuated, started non-obese and became obese.

Brant test p-value was 0.918, fail to reject the proportional odds assumption.

Table 3: Ordinal logistic regression of trajectory groups among those who started out in the obese BMI category on demographic and behavioral characteristics, Waves I to IV (n = 803)

Variable	Coefficient
Sex (reference = male)	
Female	0.403
Race (reference = non-Hispanic white)	
Non-Hispanic black	–0.157
Hispanic	–0.235
Asian	0.811
Other	–0.058
Place of residence (reference = rural)	
Suburban	–0.091
Urban	0.095
Parents' education (reference = neither parent graduated from college)	
One parent graduated from college	–0.216
Both parents graduated from college	–0.504
Hours of screen time	0.012 *
Regular exercise (reference = no)	
Yes	–0.353
Hours of sleep	–0.162 *
No regular breakfast (reference = yes)	
No	0.012
Cut 1	–3.234
Cut 2	–1.697

Notes: Significance is denoted by *** $p < 0.001$, ** $p < 0.01$, and * $p < 0.05$.

Categories: started obese and became non-obese, started obese and fluctuated, stayed obese throughout the waves.

Brant test p-value was 0.808, fail to reject the proportional odds assumption.