Title

Investigating the Role of Childhood Adiposity in the Development of Adult Type 2 Diabetes in a 64-year Follow-up Cohort: An Application of the Parametric G-formula Within an Agent-based Simulation Study.

Authors

Nianogo, Roch A
Arah, Onyebuchi A

Publication Date

2019-11-01

DOI

10.1097/ede.0000000000001062

Peer reviewed

Investigating the Role of Childhood Adiposity in the Development of Adult Type 2 Diabetes in a 64-year Follow-up Cohort

An Application of the Parametric G-formula Within an Agent-based Simulation Study

Roch A. Nianogo,[a,b] *and Onyebuchi A. Arah*[a–c]

Background: The contribution of childhood obesity to adult type 2 diabetes (T2DM), not through adult adiposity, as well as the causal pathways through which childhood obesity increases adult T2DM risk are not well understood. This study investigated the contribution of childhood obesity to incident T2DM including pathways not through adult adiposity, and explored whether race modified this contribution.

Methods: We used data from the Virtual Los Angeles Cohort, an agent-based longitudinal birth cohort composed of 98,230 simulated individuals born in 2009 and followed until age 65 years. We applied the parametric mediational g-formula to the causal mediation analysis investigating the impact of childhood obesity on the development of adult T2DM.

Results: The marginal adjusted odds ratio (aOR) for the total effect of childhood obesity on adult T2DM was 1.37 (95% CI = 1.32, 1.46). Nearly all the effect of childhood obesity on adult T2DM was mostly attributable to pathways other than through adult obesity; the aOR for the pure direct effect was 1.36 (95% CI = 1.31, 1.41). In all racial subpopulations, a similar 3% of the total effect of childhood obesity on adult T2DM was attributable to its effect on adult obesity.

Conclusions: Childhood obesity remains a risk factor for adult T2DM separate from its effects on adult obesity. This study emphasizes the potential benefits of early interventions and illustrates that agent-based simulation models could serve as virtual laboratories for exploring mechanisms in obesity research.

Keywords: Agent-based model; Cohort; Diabetes; G-formula; Mediation; Obesity; Simulation; Synthetic

(*Epidemiology* 2019;30: S101–S109)

Submitted October 6, 2017; accepted July 13, 2019.

From the [a]Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles (UCLA), Los Angeles, CA; [b]California Center for Population Research (CCPR), Los Angeles, CA; and [c]UCLA Center for Health Policy Research, Los Angeles, CA.

R.A.N. was supported by a Burroughs Wellcome Fellowship and the Dissertation Year Fellowship from UCLA. O.A.A. was partly supported by Grant R01-HD072296-01A1 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. In addition, R.A.N. benefited from facilities and resources provided by the California Center for Population Research at UCLA (CCPR), which receives core support (R24-HD041022) from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD). The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

The authors report no conflicts of interest.

The details to construct the data are described in a related article (R. A. Nianogo, O. A. Arah, unpublished data, 2017). Additional details can be made available where technical information is insufficient by contacting Roch Nianogo (niaroch@ucla.edu).

R.A.N. participated in the study conception, design, and analysis and wrote the first draft of the article. O.A.A. supervised and participated in the study conception, design, and analysis and reviewed and revised the manuscript. All authors provided critical input and insights into the development and writing of the article and approved the final manuscript as submitted.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

Correspondence: Roch A. Nianogo, Department of Epidemiology, UCLA Fielding School of Public Health, 650 Charles E. Young Drive South, Los Angeles, CA 90095. E-mail: niaroch@ucla.edu.

ISSN: 1044-3983/19/3002-S101
DOI: 10.1097/EDE.0000000000001062

For decades, obesity has been recognized as a major public health problem affecting millions of Americans, including the most vulnerable segment of the population, namely children, adolescents, and lower-income minorities.[1] Obesity in adulthood is a known risk factor for type 2 diabetes (T2DM)[2] and thus the increasing rates of obesity have substantially played a critical role in the epidemic of T2DM.[3] Studies have shown that compared with normal-weight children, obese children are at a higher risk of becoming obese adults[4] and that obese children who remain obese in adulthood are at increased risk for adult T2DM compared with obese children who become nonobese in adulthood.[2,5] This suggests that childhood obesity may be a risk factor for adult T2DM through adult adiposity and that the increased risk in T2DM due to obesity may be due to the tracking of excess weight from childhood into adulthood. However, how and by which mechanisms this occurs remains relatively unexplored.

In fact, it is unclear whether childhood obesity affects adult T2DM independently of adulthood adiposity.[6] More

generally, the causal pathways through which childhood obesity increases adult T2DM risk are not well understood. The ability to open such a black box can assist policymakers in identifying causal pathways and targeting the ones that would have the greatest impact on reducing T2DM. Investigations of this sort entail mediation and interaction analyses.[7] In the past, methods such as the so-called difference method have been used to estimate mediated (or indirect) effects in the exploration of such mechanisms but they have been unsatisfactory as they can lead to distorted results.[8,9] More recently though, novel methods such as the parametric g-formula[10] have allowed researchers to disentangle the path-specific effects of exposures or interventions.[11–13] Although powerful, the application of such methods are often limited to single existing observational prospective studies.[14,15]

Today, scientific progress has allowed researchers to create complex synthetic populations informed by best evidence, to explore mechanisms, forecast disease burden and evaluate intervention impacts.[15–17] An example of such a technique is an agent-based model (ABM), defined as a computer simulation model representing a system or reality that incorporates evidence about individuals' behaviors and their physical and social environment and could serve as virtual laboratory for testing hypotheses and running experiments.[18,19] Given the paucity of long-running cohorts and trials implemented when obesity reached epidemic proportions among children, developing a virtual cohort where individuals are followed from birth to adulthood to study obesity and its long-term effects can be a suitable alternative. However, the use of the parametric g-formula to untangle complex mechanisms within such synthetic cohort has not yet been explored. A recent study compared ABMs and the parametric g-formula for decision-making and observed that ABMs can result in biased estimates when input parameters—(e.g., baseline risk) are transported from one population to another—owing to (1) the nontransportability of such parameters or to (2) the lack of causal interpretation of regression coefficients.[15]

In the present study, we propose to address these shortcomings by using both an ABM and the parametric g-formula in conjunction, to investigate the contribution of childhood obesity to incident T2DM that is independent of its effect on adult adiposity ("direct effect"), and determine if race modifies this contribution. In particular, we will estimate consistent path-specific effects and decompose the effects of childhood obesity on T2DM by applying the parametric g-formula within a prospective synthetic cohort.

METHODS

Structure of the Agent-based Model

The Virtual Los Angeles (ViLA) Obesity Model is an agent-based model of obesity and T2DM in U.S. individuals born in Los Angeles, CA, and followed from birth to age 65 years (see eAppendix; http://links.lww.com/EDE/B571, for details). Briefly, we simulated 98,230 simulated individuals or agents spread out in 235 simulated neighborhoods from birth to age 65 years in 10 discrete time steps. Each agent was born in 2009 in a specific neighborhood of Los Angeles County and could exhibit healthy and unhealthy behaviors. At each time step, the model updated the individuals' behaviors, their body mass indices, and T2DM status as a function of the agent's current state. The ViLA-Obesity Model was also used to forecast obesity and T2DM incidence and prevalence among U.S. individuals born in Los Angeles County (R. A. Nianogo, MD, MPH, PhD, O. A. Arah, MD, MSc, MPH, DSc, PhD, unpublished data, 2017). This study was deemed not a Human Subject Research by the UCLA IRB.

Measures and Variables

Exposure: Childhood Obesity Between Age 6 and 12 Years

The exposure of interest was childhood obesity in middle childhood between the age of 6 and 12 years old. Childhood obesity was defined using the World Health Organization (WHO) guidelines on the basis of the body mass index (BMI) *Z* scores calculated using SAS code provided by the US Centers for Disease Control and Prevention.[20] We used *Z* scores instead of percentiles because *Z* scores are comparable across ages and sex and are better for longitudinal assessments.[21] A child with a BMI *Z* score (BMI*z*) less than −2 was classified as underweight; a BMI*z* greater or equal to −2 but less than 1 was classified as normal weight; a BMI*z* greater or equal to 1 but less than 2 was classified as overweight; and a BMI*z* greater or equal to 2 was classified as obese.[22]

Mediators: Adult Obesity Between Age 30 and 39 Years and Physical Activity Between Age 25 and 39 Years

The primary mediator of interest was adult obesity between the age of 30 and 39 years (binary variable). Using the WHO guidelines, an individual with a BMI less than 18.5 was classified as underweight; a BMI greater or equal to 18.5 but less than 25 was classified as normal weight; a BMI greater or equal to 25 but less than 30 was classified as overweight; and a BMI greater or equal to 30 was classified as obese.[23]

The secondary mediator of interest was the adult physical activity level (i.e., recommended moderate-to-vigorous physical activity) between age 25 and 39 years (binary variable).

Outcome: Adult T2DM Between Age 40 and 49 Years

The outcome of interest was incident adult T2DM between the ages of 40 and 49 years (binary variable).

Covariates and Intermediate Health Behaviors

The following variables were considered in this study: individuals' sociodemographics (age, sex, socioeconomic status, marital status, race), individuals' behaviors

(sugar-sweetened beverage consumption, physical activity, fruit and vegetable consumption, fast-food consumption, smoking, alcohol consumption), neighborhood walkability, and neighborhood access to parks. These variables were binary with the exception of age, which was continuous.

Statistical Analyses

Causal Graph and G-computation Algorithm

We represented our assumptions about the underlying pathways from childhood obesity to adult T2DM and the relations among covariates, exposure, mediators, and outcomes using a directed acyclic causal diagram[24] (Figure 1).

We used g-computation[10] to decompose the effect of childhood obesity on adult T2DM. This method requires correct model specification when modeling all covariates and may be sensitive to violations of assumptions.[25] To conduct our causal mediation analysis, it was assumed that there was conditional exchangeability (i.e., no-uncontrolled confounding assumption), positivity,[26] consistency,[27] no interference (i.e., stable unit treatment value assumption or SUTVA),[28] and no other sources of bias (i.e., no selection bias, no measurement error, and no model misspecification). In the context of mediation analysis, the no-uncontrolled confounding assumption consisted of four parts:[29,30] (i) no-uncontrolled confounding between exposure and outcome, (ii) no-uncontrolled confounding between mediator and outcome, (iii) no-uncontrolled confounding between exposure and mediator, and finally (iv) no exposure-induced mediator-outcome confounder. In our study, the latter assumption (iv) is violated since childhood adiposity was allowed to affect subsequent physical activity levels which in turn could affect subsequent obesity risk and adult T2DM. Fortunately, a recent study has proposed solutions to circumvent this problem.[30] We briefly described the two estimation approaches used to decompose the effect of interest.

Estimation and Effect Decomposition

To circumvent the problem of exposure-induced mediator-outcome confounder (fourth assumption), we applied two analytical approaches described in Vanderweele et al.[30] to compute other natural effects and path-specific effects. Let A denote childhood obesity at age 6–12 years (i.e., the exposure of interest), M adult obesity at age 30–39 years (i.e., the mediator of interest), L adult level of physical activity at age 25–29 years (i.e., an exposure-induced confounder), V adult level of physical activity at age 30–39 years, Y adult T2DM at age 40–49 years (i.e., the outcome of interest), and C a set of baseline covariates not affected by exposure (Figure 1). For any variable W, let $W = w$, $W = 0$, $W = 1$ denote, respectively, the realized mean, reference value, or index value of W. We will let $Y_{A=a}, L_{A=a}, M_{A=a}, V_{A=a}$ denote, respectively, the potential outcomes of Y, L, M, or V had A been set to a. Also, $Y_{A=a,M_{A=a,L_{A=a}},V_{A=a,L_{A=a}}}$ denotes the potential outcome value of Y had A been set to a, M to $M_{A=a,L_{A=a}}$ and V to $V_{A=a,L_{A=a}}$.

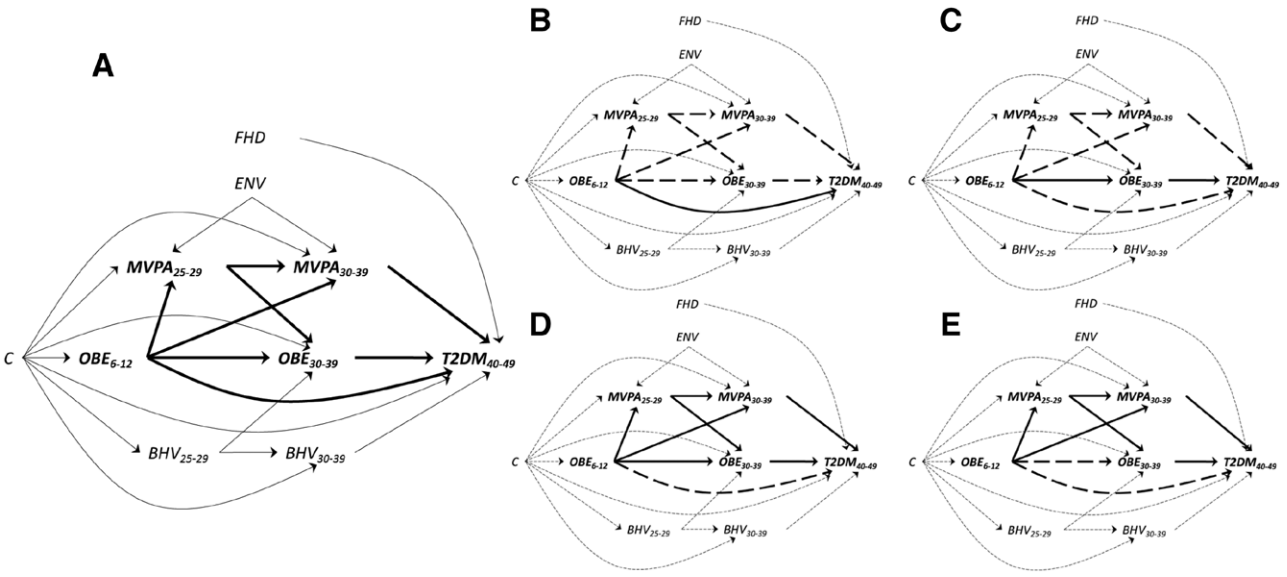


FIGURE 1. Simplified DAG of the assumptions about the data-generating processes between childhood obesity and type 2 diabetes 2. C, Sociodemographics (age, sex, socioeconomic status, marital status); BHV: time-varying behaviors (sugar-sweetened beverage consumption, fast-food consumption, fresh fruit and vegetable consumption, smoking, alcohol drinking); OBE: obesity; ENV: (Neighborhood Access to Parks, Neighborhood walkability); FHD: family history of type 2 diabetes; $MVPA_{25-29}$: moderate-to-vigorous physical activity at age 25–29 years; $MVPA_{30-39}$: moderate-to-vigorous physical activity at age 30–39 years; OBE_{30-39}: Obesity at age 30–39 years; $T2DM_{40-49}$: type 2 diabetes at age 40–49 years. The bold lines depict the pathways from childhood obesity to adult type 2 diabetes. A, General data-generating mechanism from childhood obesity to type 2 diabetes. B, Direct natural effect (PDE, TDE); (C) effect $OBE_{6-12} \rightarrow OBE_{30-39} \rightarrow T2DM_{40-49}$; (D) indirect natural effect (TIE, PIE); (E) effect $OBE_{6-12} \rightarrow MVPA_{25-29}$; $MVPA_{30-39} \rightarrow T2DM_{40-49}$. DAG indicates directed acyclic graph.

Recall that M represents adult obesity at age 30–39 years, L adult level of physical activity at age 25–29 years, and V adult level of physical activity at age 30–39 years. In the first approach "joint mediator approach," we considered the set $Z = \{L,V,M\}$ jointly (i.e., simultaneously) as the mediator of interest. In other words, from childhood obesity to adult T2DM there were essentially two pathways: (i) one direct and (ii) one indirect that combines pathways through adult obesity (M) and pathways through adult physical activity (L, V).

In the second approach, "path-specific approach," we considered adult obesity (M) as the actual mediator of interest. Put another way, from childhood obesity to adult T2DM, there were essentially three pathways: (i) pathways involving neither adult obesity nor adult level of physical activity (i.e., $A \rightarrow Y$), (ii) effects not involving adult level of physical activity (i.e., $A \rightarrow M \rightarrow Y$), and (iii) effects involving only adult level of physical activity (i.e., combination of $A \rightarrow L \rightarrow V \rightarrow Y$, $A \rightarrow L \rightarrow M \rightarrow Y$ and $A \rightarrow V \rightarrow Y$) summarized as $A \rightarrow LV \rightarrow Y$. In the second approach, path-specific effects were estimated.

For clarity, the quantities we estimated in this study are briefly defined. More extensive definitions and expressions can be found in Wang and Arah[31] and Vanderweele et al.[30]

The expressions for the natural decomposition are given as follows:

The total effect (TE) measures the overall extent to which childhood obesity causes adult T2DM. It was given by the following expression:

$$E_{\text{TE}} = E\left[Y_{A=1} - Y_{A=0}\right]$$

The pure direct effect (PDE) measures the extent to which childhood obesity causes adult T2DM through pathways other than through the joint mediator set $Z = \{L,V,M\}$ and was given by the following expression:

$$E_{\text{PDE}} = E\left[Y_{A=1,Z_{A=0}} - Y_{A=0,Z_{A=0}}\right]$$

The total direct effect (TDE) measures the extent to which childhood obesity causes adult T2DM through pathways other than through the joint mediator set $Z = \{L,V,M\}$ allowing the joint mediator set to simultaneously boost up or tune down such effect at the same time. It was given by the following expression:

$$E_{\text{TDE}} = E\left[Y_{A=1,Z_{A=1}} - Y_{A=0,Z_{A=1}}\right]$$

The pure indirect effect (PIE) measures the extent to which childhood obesity causes adult T2DM through the joint mediator set $Z = \{L,V,M\}$ only, not accounting for the possible interaction between childhood obesity and the joint mediator set $Z = \{L,V,M\}$. It was given by the following expression:

$$E_{\text{PIE}} = E\left[Y_{A=0,Z_{A=1}} - Y_{A=0,Z_{A=0}}\right]$$

The total indirect effect (TIE) measures the extent to which childhood obesity causes adult T2DM through the joint mediator set $Z = \{L,V,M\}$ only, but accounting for the possible interaction between childhood obesity and the joint mediator set $Z = \{L,V,M\}$. It was given by the following expression:

$$E_{\text{TIE}} = E\left[Y_{A=1,Z_{A=1}} - Y_{A=1,Z_{A=0}}\right]$$

The controlled direct effect (CDE) measures the extent to which childhood obesity causes adult T2DM when fixing the joint mediator set at Z specific value for everyone in the population. There are three types of CDEs: (i) the CDE_{ref} (CDE at the reference level) or CDE when fixing the joint mediator set to the reference level of 0 ($Z = 0$); (ii) the CDE_{idx} (CDE at the index level) or CDE when fixing the joint mediator set to the index level of 1 ($Z = 1$); and (iii) the CDE_{sto} (stochastic CDE) or CDE when allowing the joint mediator set to attain a certain controlled distribution in the population ($Z = z$). These quantities were given by the following expressions:

$$E_{\text{CDE}_{\text{ref}}} = E\left[Y_{A=1,Z=0} - Y_{A=0,Z=0}\right]$$

$$E_{\text{CDE}_{\text{idx}}} = E\left[Y_{A=1,Z=1} - Y_{A=0,Z=1}\right]$$

$$E_{\text{CDE}_{\text{sto}}} = E\left[Y_{A=1,Z=z} - Y_{A=0,Z=z}\right]$$

The expressions for the path-specific effects were also given as follows:

The effect neither involving adult obesity nor adult level of physical activity ($A \rightarrow Y$) was expressed as follows:

$$E_{A\rightarrow Y} = E\left[Y_{A=1,M_{A=0,L_{A=0}},V_{A=0,L_{A=0}}} - Y_{A=0,M_{A=0,L_{A=0}},V_{A=0,L_{A=0}}}\right]$$

The effect not involving adult level of physical activity ($A \rightarrow M \rightarrow Y$) was expressed as follows:

$$E_{A\rightarrow M\rightarrow Y} = E\left[Y_{A=1,M_{A=1,L_{A=0}},V_{A=0,L_{A=0}}} - Y_{A=1,M_{A=0,L_{A=0}},V_{A=0,L_{A=0}}}\right]$$

The effect involving only adult level of physical activity ($A \rightarrow LV \rightarrow Y$) was expressed as follows:

$$E_{A\rightarrow LV\rightarrow Y} = E\left[Y_{A=1,M_{A=1,L_{A=1}},V_{A=1,L_{A=1}}} - Y_{A=1,M_{A=1,L_{A=0}},V_{A=0,L_{A=0}}}\right]$$

We completed all data preparation, parametric modeling, simulations, and analysis in SAS version 9.4 (SAS Institute, Inc., Cary, NC).

Sensitivity Analysis

We conducted three groups of sensitivity analyses. First, we assessed whether the time at which childhood obesity and T2DM were assessed impacted our conclusion. Second, we varied the potential remission rate of T2DM due to lifestyle and weight loss management with or without medication to

TABLE 1. Baseline and Follow-up Characteristics of Simulated Individuals in the ViLA-obesity Model (n = 98,230)

	Childhood (6–12)	Adulthood (30–39)	Adulthood (40–49)
Age in years (mean, SD)	9 (1.78)	34.52 (2.63)	44.48 (2.63)
Male (%)		49	49
Low income (i.e., below or at FPL) (%)	22	22	22
Married (%)	0	44	44
Non-White (%)	63	63	63
Has family history of type 2 diabetes (%)	8	8	8
Ate fast-food ≥1 times in past week (%)	76	74	52
Physically active at least 1 hour per day (%)	23	27	24
Drank ≥1 glasses of SSB (%)	66	45	23
Eat ≥5 fresh fruits and vegetables (%)	45	54	51
Current smoker (%)	0	12	9
Binge drank alcohol the past month (%)	0	17	13
High neighborhood walkability (%)	28	27	27
High neighborhood access to parks (%)	54	56	55
Body mass index (kg/m^2 mean, SD)	(20.68) 4.39	27.24 (6.26)	26.24 (6.97)
Obese (%)	24	32	30
Has type 2 diabetes (%)	0	3	25

All categorical variables are binary.
FPL indicates federal poverty level; SSB, sugar-sweetened beverage consumption.

assess the robustness of our results. This was to account for the fact that T2DM could be cured in its early stages.[32,33] Third, we investigated the direct and mediated effects of childhood obesity on prevalent T2DM.

RESULTS

Table 1 describes the baseline and follow-up characteristics of the simulated cohort. Two-thirds of our population were non-White and about one-fourth had an income below or at the federal poverty level. Consumption of fast-food was found in 75% of children and in about 50% of adults in their 40s. About one-fourth and one-third of individuals were obese in childhood and adulthood, respectively. One in four individuals had T2DM in their 40s.

Table 2 presents the decomposition of the effects of childhood obesity on adult T2DM estimated using g-computation. The marginal adjusted odds ratio (aOR) for the total effect of childhood obesity on adult T2DM was 1.37 (95% CI = 1.32, 1.46). The results were similar using the joint mediator or the path-specific approach. Under both approaches, nearly all the effect of childhood obesity on adult T2DM was mostly attributable to pathways other than through adult obesity (e.g., pure direct effect aOR: 1·36 [95% CI = 1.31, 1.41]). Only 3% of the total effect of childhood obesity on adult T2DM was attributable to childhood obesity affecting adult obesity and subsequently affecting adult T2DM (Figure 2).

TABLE 2. Decomposition of the Effect of Childhood Obesity on Adult Type 2 Diabetes in the ViLA-obesity Model Using D-computation in a Marginal Structural Model

Method	OR[a] (95% CI)
Joint mediator approach ($MVPA_{25–29}$, $MVPA_{30–39}$, $OBE_{30–39}$) as the joint mediator set	
PDE	1.36 (1.31, 1.41)
TIE	1.01 (1.00, 1.02)
PIE	1.01 (1.00, 1.02)
TDE	1.36 (1.31, 1.41)
CDE_{sto} (marginal)	1.37 (1.37, 1.38)
CDE_{ref}	1.39 (1.33, 1.46)
CDE_{idx}	1.38 (1.31, 1.44)
Total Effect	1.37 (1.32, 1.46)
Path-specific approach ($OBE_{30–39}$ as the actual mediator)	
Effect involving neither adult obesity nor PA ($OBE_{6–12}$→ $T2DM_{40–49}$) (PSDE)	1.36 (1.31, 1.41)
Effect not involving PA ($OBE_{6–12}$→ $OBE_{30–39}$→ $T2DM_{40–49}$) (PSIE-A)	1.00 (0.99, 1.01)
Effect involving only PA ($OBE_{6–12}$→ $MVPA_{25–29;}$ $MVPA_{30–39}$ → $T2DM_{40–49}$) (PSIE-B)	1.01 (1.01, 1.01)

95% CI obtained via bootstrapping.
[a]Marginal odds ratio.
CI indicates confidence interval; $MVPA_{25–29}$, moderate-to-vigorous physical activity, between age 25 and 29 years; $MVPA_{30–39}$, moderate-to-vigorous physical activity at age 30–39 years; $OBE_{30–39}$, obesity between age 30 and 39 years; PA, short for adult level of physical activity; PSDE, path-specific direct effect; PSIE, path-specific indirect effect; $T2DM_{40–49}$, type 2 diabetes at age 40–49 years.

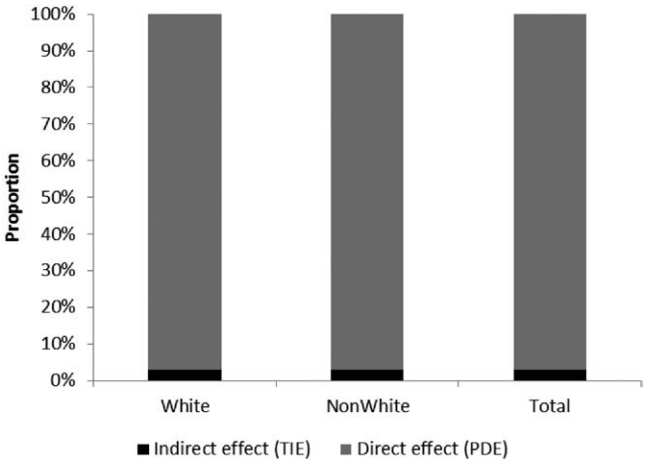


FIGURE 2. Proportion of the effect of childhood adiposity on adult type 2 diabetes that is mediated through adult adiposity by race in ViLA.

In a sensitivity analysis, the effect of early obesity on adult T2DM risk not due to adult adiposity was much greater in individuals who were obese in adolescence compared with individuals who were obese in childhood (eFigure 2; http://links.lww.com/EDE/B571). Varying the remission rates from 0% to 46% had a negligible effect on the "direct" and mediated effects of childhood obesity on adult incident and prevalent T2DM. Likewise, the direct and mediated effects of childhood obesity on adult prevalent T2DM was not substantially different from those obtained from the analysis of adult incident T2DM (eTables 2–7; http://links.lww.com/EDE/B571).

DISCUSSION

The purpose of this study was to investigate the overall contribution of childhood obesity and racial differences in the contribution of childhood obesity to incident adult T2DM. Using the g-computation algorithm[10] within the virtual cohort of Los Angelinos, we consistently examined and quantified the pathways through which childhood obesity affected T2DM. Our findings suggest that nearly all of the effect attributable to childhood obesity in the development of incident T2DM was due to pathways other than through adult obesity (so-called "direct effect"). In other words, childhood obesity affected the risk of incident T2DM but not through adult adiposity, with the effect of childhood obesity through adult obesity and adult level of physical activity appearing to be minimal in this study. In addition, we did not find the presence of racial disparities in the effect of childhood obesity on T2DM.

Our findings suggest that childhood obesity could affect metabolic health (including insulin resistance) early in life and throughout the life course without affecting adult weight. In other words, the metabolic disturbances already present in childhood[34,35] may track from childhood to adulthood without affecting adult weight. This would result in adults who have normal weight but have metabolic disturbances.[36,37] This phenotype has sometimes been termed "metabolically unhealthy normal weight" (MUNW), with individuals in this phenotype being more likely to be male, less physically active, hypertensive, and former smokers compared with normal-weight individuals without incident diabetes.[36,37] Studies have also reported that individuals who were overweight or obese during childhood but became normal weight during adulthood had a persistent increased risk of adult T2DM compared with individuals who were never overweight or obese,[2,5] suggesting that childhood obesity may have long-term effects that are independent of its effect on adult obesity. Two plausible mechanisms can explain this phenomenon. First, obesity during childhood and adolescence could lead to metabolic changes via inflammatory and hormonal pathways through the production of adipokines (e.g., tumor necrosis factor α [TNF-α], interleukin-6 [IL-6]), the release of free fatty acids (FFA) from adipocyte lipolysis, and the decrease in adiponectin level.[38–41] These changes in turn are responsible for inducing inflammation and insulin resistance, which have been incriminated in the development of T2DM.[38–41] When these hormonal and inflammatory changes are sustained outside of adult adiposity, it could result in T2DM among normal-weight individuals and the MUNW phenotype.[36,37] Second, other risk factors present early in life (i.e., before and during pregnancy and infancy) have been suggested to increase the risk of childhood obesity in children and T2DM later in life.[42–45] These include maternal smoking during pregnancy, maternal weight gain, maternal diabetes or gestational diabetes, and environmental exposures to endocrine-disrupting chemicals.[42–45] These factors may play a critical role in obesity and diabetes development through epigenetic factors and fetal metabolic programming.[46–49] Furthermore, metabolic disturbances that occur in childhood and more specifically β-cell capacity may be set early in life[50] testifying to their potential to increase T2DM risk throughout an individual life course. Further modeling and longitudinal studies should consider tracking physiologic changes over the life course to better understand the life course causes and consequences of childhood obesity.

Our findings also suggest that childhood obesity could also increase the risk of adult adiposity without necessarily impacting metabolic health; this could result in the minimal effect of childhood obesity on adult T2DM risk from adult adiposity. In other words, the excess weight gain may tract from childhood to adulthood[4,51,52] without the tracking of the metabolic disturbances. This would result in adults who are obese but are "metabolically healthy."[37,53] This phenotype has sometimes been termed "metabolically healthy obese" (MHO) and is associated with the non-Hispanic Black race, moderate alcohol intake, and higher leisure-time physical activity.[37,53] In fact, it has been shown that the long-term incidence of T2DM is reduced in MHO compared with metabolically unhealthy obese individuals.[54] Two phenomena may be at play here. First, there is a view ("the adipose tissue expandability"

hypothesis[55]) that stipulates that any excess nutrition or energy beyond the limit to which adipose tissue can expand will be stored as ectopic fat in sites such as muscles.[55] According to this hypothesis, the metabolic disturbances such as insulin resistance would only occur when deposition exceeds the capacity of the natural adipose tissue stores.[55] In other words, individuals could gain weight until a certain weight threshold but remain metabolically healthy. The second phenomenon is related to the inability of BMI to distinguish between fat mass and fat-free mass. For instance, a large muscle mass (and consequently a high BMI) would likely promote insulin sensitivity and protect against the metabolic syndrome.[56] This would result in individuals with high BMI but normal metabolic profile. That is why the use of BMI as a measure of adiposity is inadequate in correctly identifying individuals and understanding the tracking of metabolic disturbances over time.[57] In fact, BMI poorly captures body composition (i.e., whole-body fat, liver fat, leg fat, visceral fat, skeletal muscle) and the ratio of fat mass and fat-free mass (FFM) over time.[58] Further studies should incorporate better physical (e.g., measures of body fat from Dual Energy X-ray Absorptiometry [DXA]), metabolic, and hormonal measures to better assess the development of childhood obesity. The existence of the MHNW and MHO phenotypes reinforces the notion that the natural history of T2DM and its relation with adiposity may not be linear over time.[58]

At first glance, our findings may seem counterintuitive given the existing evidence of a positive association between childhood obesity and adult obesity[4,51,52] and of a positive association between adult obesity and adult T2DM.[2,5] However, assuming that there would be a non-null indirect effect through adult obesity is not necessarily guaranteed. The reasoning behind this expectation is likely based on the "product heuristics and method" for mediation analysis.[8] The product heuristics stipulates that if an exposure A (e.g., childhood obesity) affects a mediator M (e.g., adult adiposity) and the same mediator M (e.g., adult adiposity) affects an outcome Y (e.g., T2DM), then the indirect or mediated effects of A on Y through M is roughly equal to the product of the effect of A on M and the effect of M on Y. This is not necessarily true and one may need to invoke additional assumptions or perform a different analysis to consistently estimate the indirect effect using the potential outcomes framework.[8,9,31,59] As noted by Glynn,[8] a more robust analysis may result in a null indirect effect even when it appears that there is an indirect effect using the product method. Implementing a more robust analysis such as using g-computation[31,59] requires knowledge and proficiency, and this may explain why investigating the contribution of childhood obesity to chronic diseases that is independent of adult adiposity using such methods has seldom been seen in the literature.[60]

Nonetheless, studies that have attempted to tackle this issue have found no effect of childhood obesity on T2DM risk that is independent of adult obesity (i.e., no direct effect), suggesting that nearly all the increased risk in T2DM due to childhood obesity might be due to its effect on adult obesity.[60–64]These studies differ from ours in that they have conducted their mediation analysis using standard adjustment for adult BMI or weight status, a method that is closely related to the "difference method."[8,59,65] As mentioned, this simple way of adjustment to decompose effects can result in misleading estimates including false null results, especially in the presence of heterogeneous effects and interactions, and the threat of collider-stratification bias.[8,9,66] Because there seems to exist a possible interaction between childhood obesity and adult obesity in the effect of childhood obesity on T2DM,[2,67] these estimates based on adjusting for adult BMI could be biased. Furthermore, a systematic review also reported inconsistent findings,[6] and only one study seemed to be in line with our current study when adjusting for adult BMI.[68] In contrast, our study has applied the G-computation algorithm to causal mediation analysis and incorporated interactions between exposure and mediator to provide an appropriate and robust estimation of effects.[31,59]

Alternatively, our modeling could have produced artifactual associations if our model calibration did not reflect human physiology. In this iteration of the model, we calibrated subsequent variables such that the outcome at time t (e.g., BMI_t) would be a function of the lagged version of the dependent variable (e.g., BMI_{t-1}) and sociodemographics and other variables (see Table 5 in the adjacent article). This is based on the premise that weight gain can track from childhood to adulthood,[4,51,52] and on the life course perspective.[69] This ensures, in general, that BMI at a certain time-point would be related to BMI at a subsequent time-point (i.e., "short-term" tracking of BMI). However, it may not ensure that BMI, say, at time $t - 3$ (or time $t - 4$) would be related to BMI at time t after adjusting for other factors (i.e., medium to long-term tracking of BMI). This phenomenon should be further explored in longitudinal studies to help describe the short, medium, and long-term tracking of BMI over the life course. Also, although we have strived to calibrate our model to the best of our ability by careful review of the literature (e.g., randomized trials and other studies), certain model causal input parameters may be biased. Despite these potential threats, other simulation models such as Archimedes[70] and the Cardiovascular Disease Policy Model[71] have paved the way for incorporating best available evidence to replicate the pathophysiology of chronic diseases and have been used for informing clinical and policy making.

Additional limitations in this study should be considered. First, the use of a virtual cohort as opposed to a real cohort is limited by our imperfect replication of the real world. Nevertheless, the model has been validated whenever possible against external sources of data representative of Los Angeles County and is continuously being updated. Second, our analysis is based on a hypothetical birth cohort of Los Angeles County that was "born" in 2009 and thus our findings are not generalizable to the current U.S. population, but a future population of Los

Angeles. Third, another limitation related to the first is that in the ViLA-Obesity model, we do not allow for new individuals to enter the cohort once it started or for current individuals to be lost-to-follow-up, die before the end of follow-up, or experience competing risks that can prevent them from experiencing T2DM in adulthood. In essence, our model assumes that the simulated population is closed even though this is not true in the real population. Fourth, as mentioned earlier, the use of BMI as a measure of adiposity may not support the tracking of body fat from childhood to adulthood. Fifth, our model did not include early life risk factors such as prenatal and neonatal factors in this iteration of the model and so was suboptimally parametrized to track their effects from childhood to adulthood. In future iterations of the ViLA model, we may allow the simulated agents to have an offspring so that we can study the prenatal and intergenerational effects of obesity on diabetes. Finally, as is common in projection studies, our results could be biased if our assumptions are different from the real world. The high incidence of T2DM we projected could be due to the racially diverse composition of Los Angeles County and to the increasing prevalence of obesity. This high incidence in addition to a closed population and the fact that the T2DM has been modeled as a life-long condition could explain the high prevalence of T2DM found in our virtual population. In sensitivity analyses, varying the potential remission rate of T2DM due to lifestyle and weight loss management with or without medication did not substantially impact the effects of childhood obesity on incident or prevalent T2DM that were and were not mediated by adult adiposity.

In conclusion, to our knowledge, this study is one of the first to use a formal mediation analysis to look at the contribution of childhood obesity on adult T2DM independently of adult adiposity. Whether or not childhood obesity impacts T2DM through its effect on adult obesity or other mechanisms, it is difficult to reverse once established and is a strong marker for T2DM.[72,73] Therefore, prevention is warranted early in life and throughout the life course.[74] Our study used a virtual experiment to implement a formal causal mediation analysis and illustrated the utility of g-computation for effect decomposition. As demonstrated in this study, agent-based modeling can be used as a virtual laboratory for integrating best available knowledge to explore new mechanisms and heterogeneity, test novel methods, and generate new hypotheses in obesity and diabetes research.

ACKNOWLEDGMENTS

R.A.N. wishes to thank his doctoral dissertation committee for providing constructive feedback that helped improve this article.

REFERENCES

1. Flegal KM, Carroll MD, Kit BK, Ogden CL. Prevalence of obesity and trends in the distribution of body mass index among US adults, 1999–2010. *JAMA*. 2012;307:491–497.
2. Juonala M, Magnussen CG, Berenson GS, et al. Childhood adiposity, adult adiposity, and cardiovascular risk factors. *N Engl J Med*. 2011;365:1876–1885.
3. Scheen AJ, Van Gaal LF. Combating the dual burden: therapeutic targeting of common pathways in obesity and type 2 diabetes. *Lancet Diabetes Endocrinol*. 2014;2:911–922.
4. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: a systematic review of the literature. *Obes Rev*. 2008;9:474–488.
5. Park MH, Sovio U, Viner RM, Hardy RJ, Kinra S. Overweight in childhood, adolescence and adulthood and cardiovascular risk in later life: pooled analysis of three British birth cohorts. *PLoS One*. 2013;8:3–8.
6. Lloyd LJ, Langley-Evans SC, McMullen S. Childhood obesity and risk of the adult metabolic syndrome: a systematic review. *Int J Obes (Lond)*. 2012;36:1–11.
7. VanderWeele TJ. *Explanation in Causal Inference : Methods for Mediation and Interaction*. New York: Oxford University Press; 2015. https://global.oup.com/academic/product/explanation-in-causal-inference-9780199325870?cc=us&lang=en&. Accessed April 17, 2017.
8. Glynn AN. The product and difference fallacies for indirect effects. *Am J Pol Sci*. 2012;56:257–269.
9. Pearl J. The mediation formula : a guide to the assessment of causal pathways in nonlinear models mediation : direct and indirect effects. In: Berzuini C, Dawid P, Bernardinelli L, eds. *Causality: Statistical Perspectives and Applications*. Chichester, UK: John Wiley and Sons, Ltd; 2012:151–179.
10. Robins JM. A new approach to causal inference in mortality studies with sustained exposure periods--application to control of the healthy worker survivor effect. *Mathe Model*. 1986;7:1393–1512.
11. Victora CG, Horta BL, Loret de Mola C, et al. Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil. *Lancet Glob Health*. 2015;3:e199–e205.
12. Zhang YT, Laraia BA, Mujahid MS, et al. Does food vendor density mediate the association between neighborhood deprivation and BMI? *Epidemiology*. 2015;26:344–352.
13. Loret de Mola C, Hartwig FP, Gonçalves H, et al. Genomic ancestry and the social pathways leading to major depression in adulthood: the mediating effect of socioeconomic position and discrimination. *BMC Psychiatry*. 2016;16:308.
14. Nianogo RA, Wang MC, Wang A, et al. Projecting the impact of hypothetical early life interventions on adiposity in children living in low-income households. *Pediatr Obes*. 2017;12:398–405.
15. Murray EJ, Robins JM, Seage GR, Freedberg KA, Hernán MA. A comparison of agent-based models and the parametric G-formula for causal inference. *Am J Epidemiol*. 2017;186:131–142.
16. Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. *Am J Epidemiol*. 2015;181:92–99.
17. Nianogo RA, Arah OA. Agent-based modeling of noncommunicable diseases: a systematic review. *Am J Public Health*. 2015;105:e20–e31.
18. Luke DA, Stamatakis KA. Systems science methods in public health: dynamics, networks, and agents. *Annu Rev Public Health*. 2012;33:357–376.
19. Bonabeau E. Agent-based modeling: methods and techniques for simulating human systems. *Proc Natl Acad Sci U S A*. 2002;99(suppl 3):7280–7287.
20. CDC. A SAS Program for the 2000 CDC Growth Charts (ages 0 to <20 years). CDC. Available at:http://www.cdc.gov/nccdphp/dnpao/growthcharts/resources/sas.htm. Published 2016. Accessed 16 September 2016.
21. Wang Y, Chen HJ. Use of percentiles and Z-scores in anthropometry. *Handb Anthr*. 2012:91–114.
22. World Health Organization (WHO). WHO BMI-for-age growth charts. WHO. Available at: http://www.who.int/growthref/who2007_bmi_for_age/en/. Published 2015. Accessed 29 September 2016.
23. World Health Organization. BMI classification. 2016. Available at: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html. Published 2016. Accessed 6 October 2016.
24. Pearl J. Causal diagrams for emprirical research. *Biometrika*. 1995;82:669–688.
25. Daniel RM, Cousens SN, De Stavola BL, Kenward MG, Sterne JA. Methods for dealing with time-dependent confounding. *Stat Med*. 2013;32:1584–1618.

26. Westreich D, Cole SR. Invited commentary: positivity in practice. *Am J Epidemiol*. 2010;171:674–677; discussion 678.
27. Cole SR, Frangakis CE. The consistency statement in causal inference: a definition or an assumption? *Epidemiology*. 2009;20:3–5.
28. Rubin DB. Randomization analysis of experimental data: the fisher randomization test comment. *Source J Am Stat Assoc*. 1980;75:591–593. Available at: http://www.jstor.org/stable/2287653. Accessed 6 October 2016.
29. Pearl J. Direct and indirect effects. In: *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufmann; 2001:411–420.
30. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*. 2014;25:300–306.
31. Wang A, Arah OA. G-computation demonstration in causal mediation analysis. *Eur J Epidemiol*. 2015;30:1119–1127.
32. Gregg EW, Chen H, Wagenknecht LE, et al.; Look AHEAD Research Group. Association of an intensive lifestyle intervention with remission of type 2 diabetes. *JAMA*. 2012;308:2489–2496.
33. Bohula EA, Scirica BM, Inzucchi SE, et al. Effect of lorcaserin on prevention and remission of type 2 diabetes in overweight and obese patients (CAMELLIA-TIMI 61): a randomised, placebo-controlled trial. *Lancet*. 2018;392:2269–2279.
34. Cote AT, Harris KC, Panagiotopoulos C, Sandor GG, Devlin AM. Childhood obesity and cardiovascular dysfunction. *J Am Coll Cardiol*. 2013;62:1309–1319.
35. Mattsson N, Nnemaa TR, Juonala M, Viikari JSA, Raitakari OT. Childhood predictors of the metabolic syndrome in adulthood. The Cardiovascular Risk in Young Finns Study. *Ann Med*. 2008;40:542–552.
36. Eckel N, Mühlenbruch K, Meidtner K, Boeing H, Stefan N, Schulze MB. Characterization of metabolically unhealthy normal-weight individuals: risk factors and their associations with type 2 diabetes. *Metabolism*. 2015;64:862–871.
37. Wildman RP, Muntner P, Reynolds K, Mcginn AP. The obese without cardiometabolic risk factor clustering and the normal weight with cardiometabolic risk factor clustering. *Arch Intern Med*. 2013;168:1617–1624.
38. Ahmad SI. *Diabetes: An Old Disease, a New Insight*. In Ahmad SI, ed. New York, NY: Springer New York; 2013. doi:10.1007/978-1-4614-5441-0
39. Ouchi N, Kihara S, Funahashi T, Matsuzawa Y, Walsh K. Obesity, adiponectin and vascular inflammatory disease. *Curr Opin Lipidol*. 2003;14:561–566.
40. Lara-Castro C, Fu Y, Chung BH, Garvey WT. Adiponectin and the metabolic syndrome: mechanisms mediating risk for metabolic and cardiovascular disease. *Curr Opin Lipidol*. 2007;18:263–270.
41. van Kruijsdijk RC, van der Wall E, Visseren FL. Obesity and cancer: the role of dysfunctional adipose tissue. *Cancer Epidemiol Biomarkers Prev*. 2009;18:2569–2578.
42. Gillman MW, Rifas-Shiman SL, Kleinman K, Oken E, Rich-Edwards JW, Taveras EM. Developmental origins of childhood overweight: potential public health impact. *Obesity (Silver Spring)*. 2008;16:1651–1656.
43. Power C, Jefferis BJ. Fetal environment and subsequent obesity: a study of maternal smoking. *Int J Epidemiol*. 2002;31:413–419.
44. Daniels SR, Arnett DK, Eckel RH, et al. Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment. *Circulation*. 2005;111:1999–2012.
45. Newbold RR, Padilla-Banks E, Jefferson WN, Heindel JJ. Effects of endocrine disruptors on obesity. *Int J Androl*. 2008;31:201–208.
46. Slomko H, Heo HJ, Einstein FH. Minireview: epigenetics of obesity and diabetes in humans. *Endocrinology*. 2012;153:1025–1030.
47. Ruchat SM, Hivert MF, Bouchard L. Epigenetic programming of obesity and diabetes by in utero exposure to gestational diabetes mellitus. *Nutr Rev*. 2013;71(suppl 1):S88–S94.
48. Heerwagen MJ, Miller MR, Barbour LA, Friedman JE. Maternal obesity and fetal metabolic programming: a fertile epigenetic soil. *Am J Physiol Regul Integr Comp Physiol*. 2010;299:R711–R722.
49. Dabelea D, Crume T. Maternal environment and the transgenerational cycle of obesity and diabetes. *Diabetes*. 2011;60:1849–1855.
50. Thearle MS, Bunt JC, Knowler WC, Krakoff J. Childhood predictors of adult acute insulin response and insulin action. *Diabetes Care*. 2009;32:938–943.
51. Guo SS, Chumlea WC. Tracking of body mass index in children in relation to overweight in adulthood. *Am J Clin Nutr*. 1999;70:145S–148S.
52. Fuentes RM, Notkola IL, Shemeikka S, Tuomilehto J, Nissinen A. Tracking of body mass index during childhood: a 15-year prospective population-based family study in eastern Finland. *Int J Obes Relat Metab Disord*. 2003;27:716–721.
53. Stefan N, Kantartzis K, Machann J, et al. Identification and characterization of metabolically benign obesity in humans. *Arch Intern Med*. 2008;168:1609–1616.
54. Meigs JB, Wilson PW, Fox CS, et al. Body mass index, metabolic syndrome, and risk of type 2 diabetes or cardiovascular disease. *J Clin Endocrinol Metab*. 2006;91:2906–2912.
55. Virtue S, Vidal-Puig A. Adipose tissue expandability, lipotoxicity and the metabolic syndrome–an allostatic perspective. *Biochim Biophys Acta*. 2010;1801:338–349.
56. Dulloo AG, Jacquet J, Solinas G, Montani JP, Schutz Y. Body composition phenotypes in pathways to obesity and the metabolic syndrome. *Int J Obes (Lond)*. 2010;34(suppl 2):S4–17.
57. Blundell JE, Dulloo AG, Salvador J, Frühbeck G; EASO SAB Working Group on BMI. Beyond BMI–phenotyping the obesities. *Obes Facts*. 2014;7:322–328.
58. Müller MJ, Lagerpusch M, Enderle J, Schautz B, Heller M, Bosy-Westphal A. Beyond the body mass index: tracking body composition in the pathogenesis of obesity and the metabolic syndrome. *Obes Rev*. 2012;13(Suppl 2):6–13.
59. Robins JM, Greenland S. Identifiability and exchangeability for direct and indirect effects. *Epidemiology*. 1992;3:143–155.
60. Park MH, Falconer C, Viner RM, Kinra S. The impact of childhood obesity on morbidity and mortality in adulthood: a systematic review. *Obes Rev*. 2012;13:985–1000.
61. Tirosh A, Shai I, Afek A, et al. Adolescent BMI trajectory and risk of diabetes versus coronary disease. *N Engl J Med*. 2011;364:1315–1325.
62. Lawlor DA, Davey Smith G, Clark H, Leon DA. The associations of birthweight, gestational age and childhood BMI with type 2 diabetes: findings from the aberdeen children of the 1950s cohort. *Diabetologia*. 2006;49:2614–2617.
63. Hyppönen E, Power C, Smith GD. Prenatal growth, BMI, and risk of type 2 diabetes by early midlife. *Diabetes Care*. 2003;26:2512–2517. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12941711. Accessed 1 May 2014.
64. Must A, Jacques PF, Dallal GE, Bajema CJ, Dietz WH. Long-term morbidity and mortality of overweight adolescents. *N Engl J Med*. 1992;327:1350–1355.
65. Pearl J. The causal mediation formula–a guide to the assessment of pathways and mechanisms. *Prev Sci*. 2012;13:426–436.
66. VanderWeele TJ. Explanation in Causal Inference : Methods for Mediation and Interaction. Available at: https://global.oup.com/academic/product/explanation-in-causal-inference-9780199325870?cc=us&lang=en&. Accessed 11 May 2018.
67. Liang Y, Hou D, Zhao X, et al. Childhood obesity affects adult metabolic syndrome and diabetes. *Endocrine*. 2015;50:87–92.
68. Al Mamun A, Cramb SM, O'Callaghan MJ, Williams GM, Najman JM. Childhood overweight status predicts diabetes at age 21 years: a follow-up study. *Obesity (Silver Spring)*. 2009;17:1255–1261.
69. Ben-Shlomo Y, Kuh D. A life course approach to chronic disease epidemiology: conceptual models, empirical challenges and interdisciplinary perspectives. *Int J Epidemiol*. 2002;31:285–293.
70. Eddy DM, Schkessinger L. A trial-validated model of diabetes. *Diabetes Care*. 2003;26:3093–3101.
71. Wang M, Moran AE, Liu J, et al. Projected impact of salt restriction on prevention of cardiovascular disease in China: a modeling study. *PLoS One*. 2016;11:1–16.
72. Twig G, Tirosh A, Leiba A, et al. BMI at age 17 years and diabetes mortality in midlife: a nationwide cohort of 2.3 million adolescents. *Diabetes Care*. 2016;39:3093–3101.
73. Kelsey MM, Zaepfel A, Bjornstad P, Nadeau KJ. Age-related consequences of childhood obesity. *Gerontology*. 2014;60:222–228.
74. Nianogo RA, Arah OA. Impact of public health interventions on obesity and type 2 diabetes prevention: a Simulation Study. *Am J Prev Med*. 2018;55(6):795–802.

Investigating the Role of Childhood Adiposity in the Development of Adult Type 2 Diabetes in a 64-Year Follow-Up Cohort: An Application of the Parametric G-formula within an Agent-Based Simulation Study

Supplemental file

Contents

General Overview of the ViLA Obesity Model

The ViLA is a stochastic dynamic discrete-time agent-based computer simulation model of obesity and type 2 diabetes incidence and prevalence in the population of Los Angeles County. Each agent was born in 2009 in a specific neighborhood of Los Angeles County defined

by socio-demographics (i.e. proportion of individual self-identified as non-White, the proportion of individuals living below the federal poverty level and the proportion of individuals who had a bachelor's degree or higher), physical activity opportunities (i.e. walkability, access to parks), food environment (i.e. supermarket, fast-food densities). The agent could exhibit healthy and unhealthy behaviors (i.e. breastfeeding, fast-food consumption, sugar-sweetened beverage consumption, fresh fruit and vegetable consumption, moderate-to-vigorous physical activity, cigarette smoking and alcohol binge drinking). In this synthetic cohort, 98,000 US adults residing in Los Angeles were born in 2009 and followed until age 65 in 10 discrete time steps in order to study obesity and type 2 diabetes.

Risk functions for incident type 2 diabetes and other endogenous variables

At each time step, the model updated the individuals' behaviors, their body mass indices and type 2 diabetes status as a function of the agent's current state. Lifestyle behaviors had a similar form and were a function of the behavior in the previous step and the agent's socio-demographics. In addition, some behaviors such as physical activity were also a function of the neighborhood physical opportunities. BMI and type 2 diabetes were calculated as a function of lifestyle behaviors in the previous steps, the BMI at the previous step and the agent's socio-demographics. For example, the equation model for type 2 diabetes can be written as follows: $logitPr(Y_{t+1} = 1|A_t,\ Age_t, C, Y_t = 0) = \beta_0 + \beta_{At}A_t + \beta_{Aget}Age_t\ + \beta_C C$ where *Y* represents type 2 diabetes, *A* represents lifestyle behaviors and BMI, and *C* represents time-invariant socio-demographics. The relative risks (i.e. β's) were obtained from published evidence and from the National Health and Nutrition Examination Survey (NHANES). (See Appendix 1)

Calibration and Validation

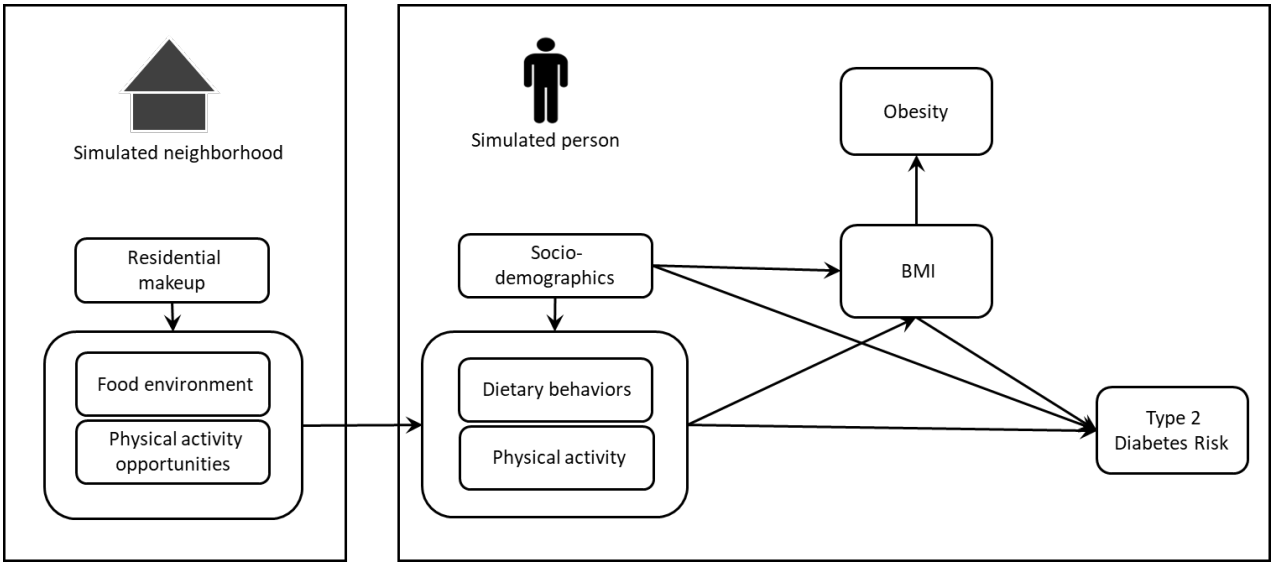
The socio-demographics and risk factor distributions for the population and the physical environment were pulled from the American Community Survey (ACS), the California Health Interview Survey (CHIS), the Centers for Disease Control and Prevention (CDC) and the World Health Organization (WHO). The relative risks (i.e. regression coefficients relating any two variables) came from published evidence or from parameters computed using publicly and privately available data (e.g. American Community Survey, National Establishment Time-Series [NETS], Walkscore.com, WHO, National Health and Nutrition Examination Survey [NHANES]). (see Appendix table 1)

The ViLA Obesity Model was calibrated to be representative of the American population residing in Los Angeles, California. This was done through the fine tuning of the baseline risk (i.e. intercept) and a subsequent internal validation was accomplished by ensuring that the predicted mean outcome approximately matched the observed mean outcome.

Appendix Table 1. Data sources for the socio-demographics and risk factor distribution

Socio-demographics and risk factor distribution	Data sources

Neighborhood socio-demographics (percent non-White, percent below federal poverty level, percent bachelor graduates graduate or above)	American Community Survey [1]
Neighborhood food environment (fast-food density, supermarket density)	National Establishment Time-Series [2]
Neighborhood physical activity opportunities (park density, walkability)	National Establishment Time-Series [2] Walkscore.com[3]
Individual socio-demographics (sex, race, income, marital status)	American Community Survey [1]
Breastfeeding	Centers for Disease Control and Prevention [4]
Individual behaviors (i.e. fast-food consumption Moderate-to-vigorous physical activity, Sugar sweetened beverage consumption, Fresh fruit and vegetable consumption Smoking, Alcohol drinking)	California Health Interview Survey [5]
Type 2 diabetes	California Health Interview Survey [5]
Body Mass Index	WHO[6] Los Angeles Health and Nutrition Examination Survey [7] National Health and Nutrition Examination Survey [8]

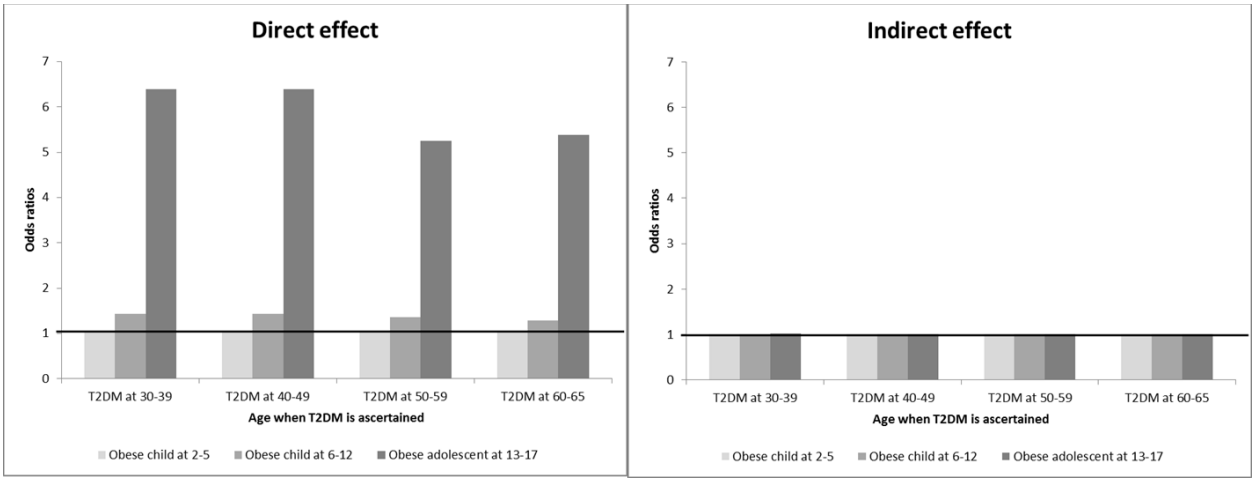


Appendix Figure 1. Simplified ViLA Model diagram

The model simulates individuals living in specific simulated neighborhood from birth until age 65. Each neighborhood is simulated with specified residential makeup, food environment and physical activity opportunities. Likewise, simulated individuals have characteristics such as socio-demographics, dietary behaviors and physical activity that can affect their subsequent body mass index (BMI), obesity and type 2 diabetes risk.

Sensitivity Analysis

- **Impact of changing the time at which the childhood obesity and adult T2DM were assessed on the direct and mediated effects of childhood obesity on incident T2DM**



Appendix Figure 2. Sensivity analysis for the decomposition of the effect of childhood obesity on adult type 2 diabetes in the ViLA-Obesity model.

The direct effect and the indirect effect represent the pure (PDE) and total effect (TIE) respectively. T2DM, Type 2 diabetes

- **Impact of varying the remission rates of T2DM on the direct and mediated effect of childhood obesity on incident and prevalent T2DM**

In the main analysis, we had assumed that T2DM was a life-long condition and that individuals who are get T2DM will remain diabetic until the end of follow-up. Given increasing literature that remission of T2DM is possible through various means, we conducted a sensitivity analysis varying potential remission rates to assess the robustness of our results. We assessed whether the direct and mediated effect of childhood obesity on T2DM would be affected if we were to T2DM were reversed in its early stages due to lifestyle and medication interventions.[9–11] We did not include bariatric surgery because of the high financial associated cost, the long-term complications and the fact that not everyone would want to undergo surgery for treating type 2 diabetes. Most studies defined partial remission rate as having a level of HbA1c between 5.7% and 6.5% (i.e. transition from diabetes to prediabetes) and complete remission as having a level of HbA1c <5.7% (i.e. normalization of the hyperglycemia without taking any medication). [9–11] For the purpose of the sensitivity analysis in this study, we defined the remission rate as the percent of individuals with T2DM who experience a partial or complete remission of their type 2 diabetes. In particular, we calculated an “effective remission rate” defined as the percent of individuals with T2DM who seek, obtain and comply with the intervention and who experience a partial or complete remission of their type 2 diabetes. This is calculated as the remission rate times the percent hypothetical uptake of lifestyle and/or medication intervention. This takes into account the fact that not every person with type 2 diabetes will successfully seek, obtain and

comply with the intervention. Essentially, the effective remission rates are as follow: 46%, 35%, 23%, 12%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 0%. We simulated these remission rates among people with type 2 diabetes at each time step starting with time step 4 (Age 18-24) until time step 9 (60-65). Since the results were similar when using the joint mediator approach and the path-specific approach, we only focused the sensitivity analysis on the pure and controlled direct effect of childhood obesity on adult T2DM. We reported the Total Indirect Effect (TIE), Pure Direct Effect (PDE), the Total Effect (TE), the Stochastic Controlled Direct (CDEsto) for both adult incident and prevalent T2DM. We reported the effects on prevalent T2DM since we expected the prevalence of T2DM to be different when applying the remission rates. Furthermore, seeing that the biggest change in T2DM prevalence occurred after 40-49, we conducted the sensitivity analysis on three outcomes T2DM at 40-49, T2DM at 50-59 and T2DM at 60-65.

As expected, simulating remission among people who had diabetes in ViLA did not alter the incidence proportion of T2DM as initially calculated. This was to be expected since we did not change the data generating process as far as the incidence of T2DM in the model. Therefore, varying the rates of remission of T2DM negligibly impacted the direct and mediated of effects of childhood obesity on adult incident T2DM. (See Appendix Table 2 to 4). We recall here that our outcome of interest in the manuscript was incident T2DM defined as the first occurrence of T2DM among at-risk individuals.

However, as expected, varying the remission rate of T2DM would impact the prevalence of T2DM in ViLA. Nevertheless, as seen in Appendix Table 5 to 7, when varying the effective remission rates from 0% to 46%, the mediated and direct effects of childhood obesity on adult prevalent T2DM was not substantially different from those obtained from the analysis of adult incident T2DM. This suggested that remission of type 2 diabetes although it had a great impact on the prevalence of T2DM, did not substantially impact the direct and mediated effects of childhood obesity on both incident and prevalent T2DM.