eFigure 1: Boxplots of Baseline G_i*statistic by Study Field Center

Boxplots of G-statistic by Study Center at Baseline
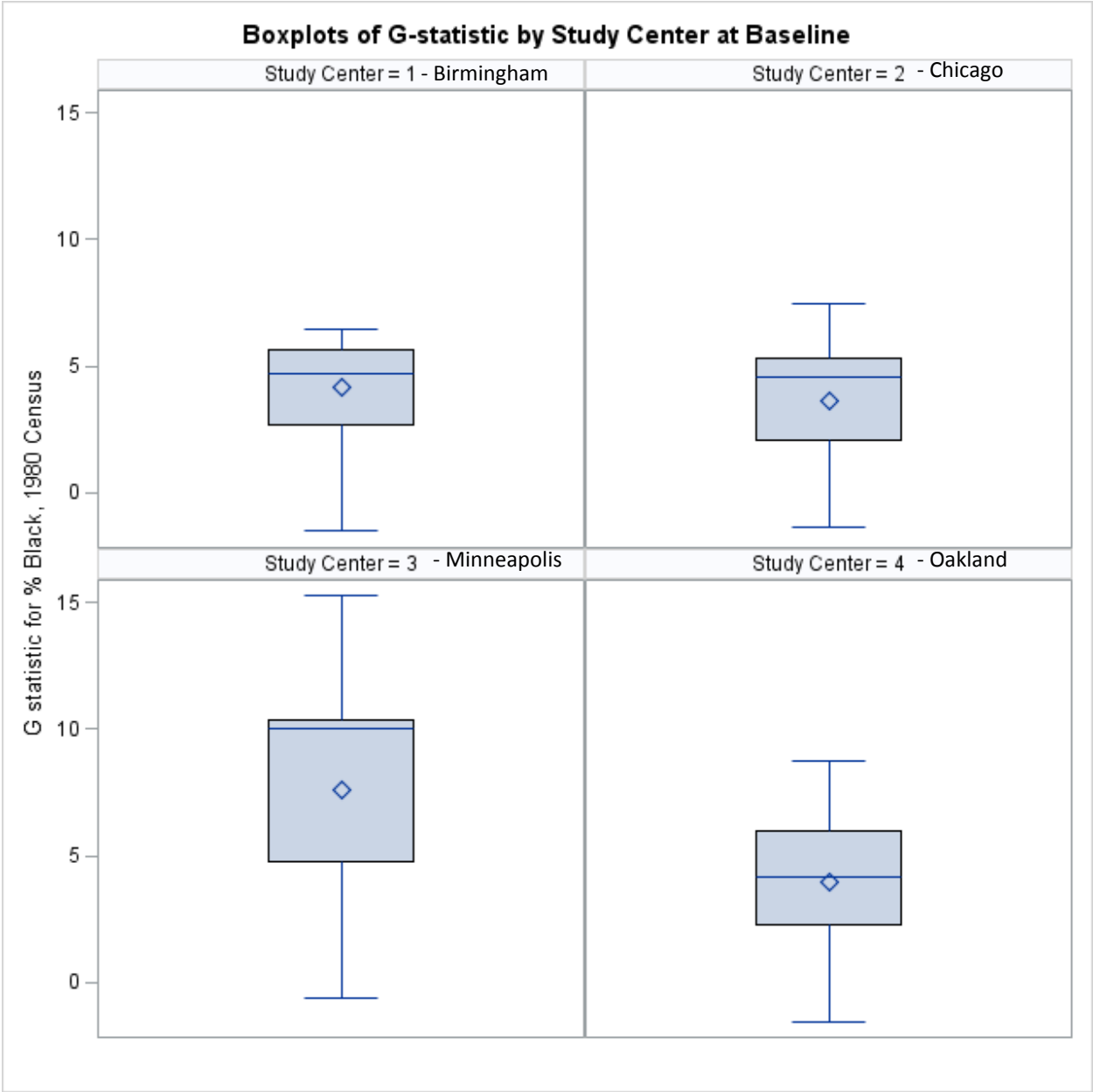
eFigure 2. Directed Acyclic Graph of the Hypothesized Relationships Between Neighborhood-level Racial Residential Segregation and Obesity
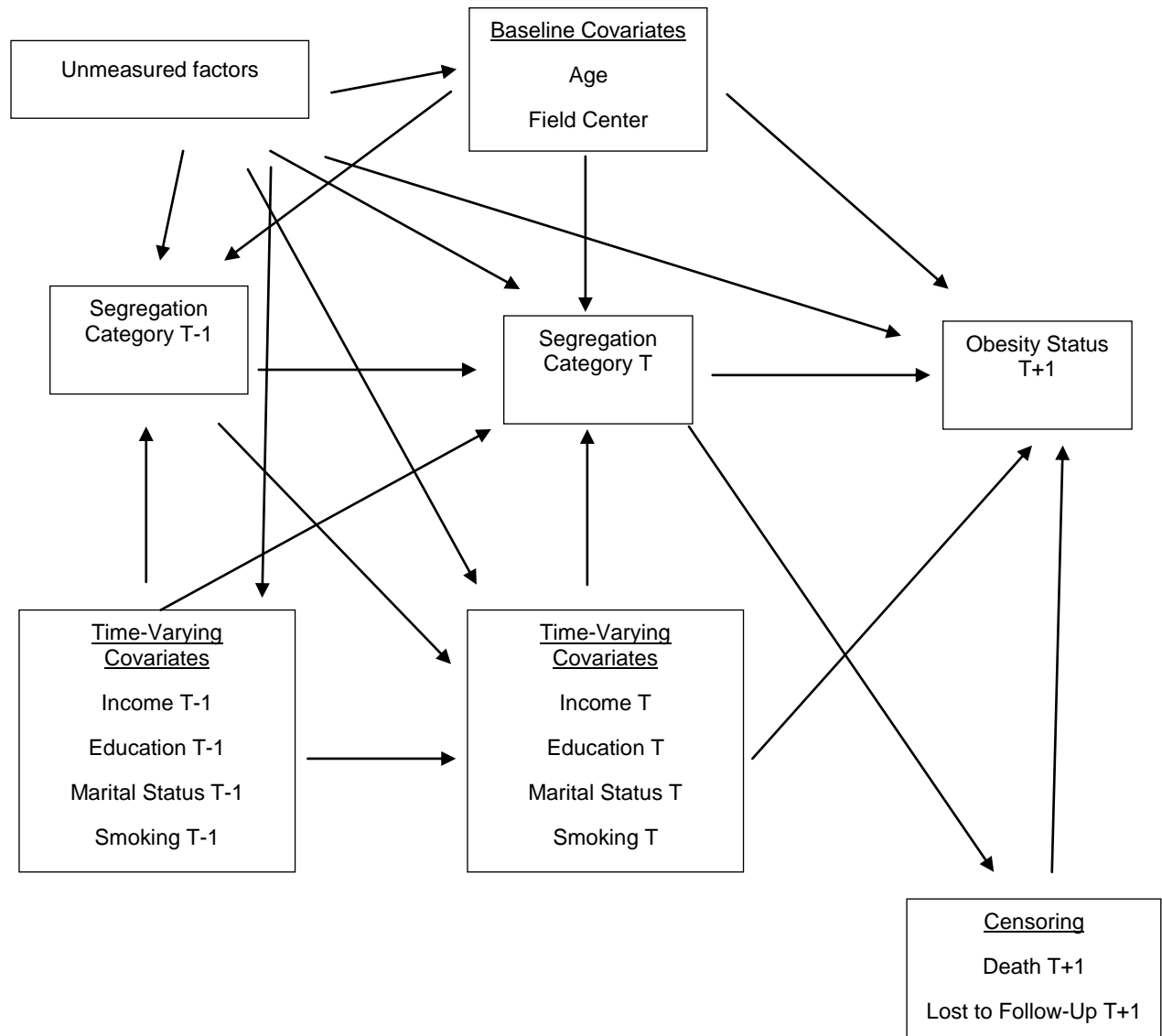
eTable 1: Number of Black Women and Men in the CARDIA Study by Category of Racial Residential Segregated Neighborhoods and Category of Neighborhood Poverty

Segregation Category[b], %	Women (n=1,167)			Men (n=1,040)		
	Neighborhood Poverty Category:[a]			Neighborhood Poverty Category:[a]		
	High	Medium	Low	High	Medium	Low
High	340	367	240	363	292	186
Medium	10	49	85	11	51	64
Low	5	8	63	2	5	66

CARDIA, Coronary Artery Risk Development in Young Adults Study

[a] Levels of neighborhood poverty categorized into high, medium, and low based on tertiles of the percentage of neighborhood residents living in poverty.

[b] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

eTable 2: Hazard of Incident Obesity by Sociodemographic Characteristics and Cardiovascular Risk Factors among Black Women and Men in the CARDIA Study, 1985-2011

	Incident Obesity HR (95% CI)	
	Women (n=1,167)	Men (n=1,040)
Model 1: Baseline Covariates[a]		
Neighborhood Poverty		
High	1.0 (0.79,1.3)	1.0 (0.78,1.4)
Medium	1.1 (0.87,1.3)	0.91 (0.68,1.2)
Low (Ref)	1.0	1.0
Educational Attainment		
Less than high school (Ref)	1.0	1.0
High school diploma	1.0 (0.76,1.4)	0.91 (0.65,1.3)
Some college	0.85 (0.61,1.2)	1.0 (0.70,1.5)
Household Income	0.98 (0.93,1.0)	1.1 (0.98,1.1)
Married/Partnered	0.97 (0.79,1.2)	1.0 (0.77,1.3)
Current Smoker	0.91 (0.74,1.1)	0.79 (0.62,1.0)
Physical Activity	1.0 (1.0,1.0)	1.0 (1.0,1.0)
Model 2: Time-varying Covariates[b]		
Neighborhood Poverty		
High	0.98 (0.78,1.3)	1.1 (0.80,1.4)
Medium	1.0 (0.81,1.3)	1.0 (0.77,1.4)
Low	1.0	1.0
Educational Attainment		
Less than high school	1.0	1.0
High school diploma	1.2 (0.8,1.7)	1.3 (0.86,2.0)
Some college	0.99 (0.68,1.4)	1.3 (0.85,2.1)
Household Income	0.98 (0.93,1.1)	1.1 (0.99,1.2)
Married/Partnered	0.96 (0.79,1.2)	1.0 (0.79,1.3)
Current Smoker	0.91 (0.73,1.1)	0.75 (0.57,0.98)
Physical Activity	1.0 (1.0,1.0)	1.0 (1.0,1.0)

CARDIA, Coronary Artery Risk Development in Young Adults Study; CI, Confidence Interval; HR, Hazard Ratio

[a] Discrete time hazard model predicting hazard of incident obesity, simultaneously adjusted for all variables listed (as recorded at baseline), as well as age at baseline and field center at baseline

[b] Discrete time hazard model predicting hazard of incident obesity, simultaneously adjusted for all variables listed (as recorded at each exam), as well as age at baseline and field center at baseline

eAppendix 1: CARDIA Data Access and SAS Programming Code Used to Run Marginal Structural Models

Data from the CARDIA Study can be accessed with an approved proposal request and sponsorship by a CARDIA investigator. More information on the proposal process, and a list of CARDIA investigators is available online at http://www.cardia.dopm.uab.edu.

The sample program below provides code to run discrete time hazard models and marginal structural models. Data should be imputed and in long form to properly run these models. This is an example of cumulative segregation models – the baseline (non-MSM only) and time-varying models could be run by substituting the exposure variables in the model statement.

```
/*Table 3: Hazard of Incident Obesity by Residential Segregation Among Black
CARDIA Participants, 1985-2011*/


/*************Discrete Time Hazard Models*****************/

proc sort data=analysis_long;
by _Imputation_ MALE ID year;
run;


/*produce estimates using imputed data*/
title "Non-MSM - Cumulative Segregation and Time-Varying Covariates";
proc genmod data=analysis_long descending;
class ID FFIPS2010 year (ref="0") CENTER (ref="1") educat_tv(ref="1");
model incid_obese=cuml_dummyseg_2 cuml_dummyseg_3 year CENTER EXAMAGE
educat_tv marstat_tv smoking_tv physact_tv income_tv
                / dist=bin link=cloglog type3 covb;
repeated subject=ID(FFIPS2010) / type=exch printmle;
by _Imputation_ male;
  ods output ParameterEstimates=gmparms
          ParmInfo=gmpinfo
          CovB=gmcovb;
run;

/*prepare data output by model for proc mianalyze – delete reference
variables*/
data test;
set gmparms;
if DF=0 then delete;
run;

data test2;
set gmpinfo;
if Parameter in ("Prm8", "Prm12", "Prm16") then delete;
run;

proc sort data=test; by male; run;
proc sort data=test2; by male; run;
proc sort data=gmcovb; by male; run;

/*produce final estimates*/
proc mianalyze parms=test covb=gmcovb parminfo=test2;
MODELEFFECTS cuml_dummyseg_2 cuml_dummyseg_3;
by male;
```

```
run;


/******* Marginal structural model for Cumulative Segregation - Women*******/

/*run separately by gender*/

data female2;
set analysis_long;
if male=0;
run;


title "IPW Numerator Model";
proc logistic data=female2 descending;
class year (ref="0") CENTER (ref="1") educat_bl (ref="1") /param=ref;
model racial_seg_cuml=year CENTER EXAMAGE educat_bl marstat_bl smoking_bl
physact_bl income_bl
/link=glogit;
output out=model1 pred=seg_0;
where cens=0 and died=0;
run;

title "IPW Demoninator Model";
proc logistic data=female2 descending;
class year (ref="0") CENTER (ref="1") educat_tv (ref="1") /param=ref;
model racial_seg_cuml=year CENTER EXAMAGE educat_tv marstat_tv smoking_tv
physact_tv income_tv
/link=glogit;
output out=model2 pred=seg_w;
where cens=0 and died=0;
run;

title "IPW Lost to Follow Up Numerator Model";
proc logistic data=female2;
class racial_seg_cuml (ref="1") year (ref="0") CENTER (ref="1") educat_bl
(ref="1") /param=ref;
model cens = racial_seg_cuml year CENTER EXAMAGE educat_bl marstat_bl
smoking_bl physact_bl income_bl ;
output out=model3 pred=cens_0;
run;

title "IPW Lost to Follow Up Model";
proc logistic data=female2;
class racial_seg_cuml (ref="1") year (ref="0") CENTER (ref="1") educat_tv
(ref="1") /param=ref;
model cens = racial_seg_cuml year CENTER EXAMAGE educat_tv marstat_tv
smoking_tv physact_tv income_tv ;
output out=model4 pred=cens_w;
run;

title "IPW Death Numerator Model";
proc logistic data=female2;
class racial_seg_cuml (ref="1") year (ref="0") CENTER (ref="1") educat_bl
(ref="1") /param=ref;
model died = racial_seg_cuml year CENTER EXAMAGE educat_bl marstat_bl
smoking_bl physact_bl income_bl ;
```

```
output out=model5 pred=died_0;
run;


title "IPW Death Demoninator Model";
proc logistic data=female2;
class racial_seg_cuml (ref="1") year (ref="0") CENTER (ref="1") educat_tv
(ref="1") /param=ref;
model died = racial_seg_cuml year CENTER EXAMAGE educat_tv marstat_tv
smoking_tv physact_tv income_tv ;
output out=model6 pred=died_w;
run;


/*reorganize data to produce stabilized IPWs*/

proc sort data=model1;
by _Imputation_ ID year _level_;
run;


proc transpose DATA=model1 OUT=model1a (drop=_name_ _label_);
by _Imputation_ ID year;
var seg_0;
id _level_;
run;


data model1a;
set model1a;
seg_01=_1;
seg_02=_2;
seg_03=_3;
run;


proc sort data=model2;
by _Imputation_ ID year _level_;
run;


proc transpose DATA=model2 OUT=model2a (drop=_name_ _label_);
by _Imputation_ ID year;
var seg_w;
id _level_;
run;


data model2a;
set model2a;
seg_w1=_1;
seg_w2=_2;
seg_w3=_3;
run;


proc sort data=female2; by _Imputation_ ID year; run;
proc sort data=model1a; by _Imputation_ ID year; run;
proc sort data=model2a; by _Imputation_ ID year; run;
proc sort data=model3; by _Imputation_ ID year; run;
proc sort data=model4; by _Imputation_ ID year; run;
proc sort data=model5; by _Imputation_ ID year; run;
proc sort data=model6; by _Imputation_ ID year; run;
```

```
/*calculate IPWs*/

data IPW_seta;
merge female2 model1a model2a model3 model4 model5 model6;
by _Imputation_ ID year;
if racial_seg_cuml=3 then do; num=seg_03; den=seg_w3; end;
if racial_seg_cuml=2 then do; num=seg_02; den=seg_w2; end;
if racial_seg_cuml=1 then do; num=seg_01; den=seg_w1; end;
/*W2=round(num/den,.001);*/

if first.ID then do;
k1_0=1; k1_w=1; kc1_0=1; kc1_w=1; kd1_0=1; kd1_w=1;
end;
retain k1_0 k1_w kc1_0 kc1_w kd1_0 kd1_w;

/*inverse probability of treatment weights*/
k1_0=k1_0*num;
k1_w=k1_w*den;

/*inverse probability of censoring weights*/
kc1_0=kc1_0*cens_0;
kc1_w=kc1_0*cens_w;

/*inverse probability of death weights*/
kd1_0=kd1_0*died_0;
kd1_w=kd1_0*died_w;

/*stabilized weights*/
stabwt=(k1_0*kc1_0*kd1_0)/(k1_w*kc1_w*kd1_w);
run;


/*check IPW distribution*/

title "IPW Means and Standard Deviations";
proc means data=IPW_seta n nmiss mean std min max median;
var stabwt;
run;


proc means data=IPW_seta n mean std;
class year;
var stabwt;
run;


/*Truncate Outlier IPWs*/

/*check the distribution and determine 1% 99% cut points*/
title "IPW Weight Distribution";
proc univariate data=IPW_seta;
var stabwt;
run;


data IPW_setb;
set IPW_seta;
stabwt2=stabwt;
/*1% Truncate*/
if stabwt>2.3715827 then stabwt2=2.3715827;
if stabwt<0.4756949 and stabwt ne . then stabwt2=0.4756949;
```

```
run;


title "IPW Means and Standard Deviations";
proc means data=IPW_setb n nmiss mean std min max median;
var stabwt2;
run;

proc means data=IPW_setb n mean std;
class year;
var stabwt2;
run;


/*Apply IPWs in Discrete Time Hazard Model*/

title "Marginal Structural Model Cumulative Segregation - Women";
proc genmod data=IPW_setb descending;
class ID FFIPS2010 year (ref="0") CENTER (ref="1") educat_bl (ref="1") ;
model incid_obese=cuml_dummyseg_2 cuml_dummyseg_3 year CENTER EXAMAGE
educat_bl marstat_bl smoking_bl physact_bl income_bl
                / dist=bin link=cloglog type3 covb;
scwgt stabwt2;
repeated subject=ID(FFIPS2010) / type=exch printmle;
by _Imputation_;
  ods output ParameterEstimates=gmparms
            ParmInfo=gmpinfo
            CovB=gmcovb;
run;


/*prepare data output by model for proc mianalyze – delete reference
variables*/

data test;
set gmparms;
if DF=0 then delete;
run;


data test2;
set gmpinfo;
if Parameter in ("Prm8", "Prm12", "Prm16") then delete;
run;


proc sort data=test; run;
proc sort data=test2; run;
proc sort data=gmcovb; run;


/*produce final estimates*/

proc mianalyze parms=test covb=gmcovb parminfo=test2;
MODELEFFECTS cuml_dummyseg_2 cuml_dummyseg_3;
run;
```

eTable 3: Baseline Participant Characteristics of Black Women and Men in the CARDIA Study, by Racial Residential Segregation Category for BIRMINGHAM Field Center

	Women (n=277)			Men (n=253)		
	Segregation Category:[a]			Segregation Category:[a]		
	High n=235	Medium n=37	Low n=5	High n=217	Medium n=25	Low n=11
Age in years, mean (SD)	24.2 (3.7)	23.7 (4.1)	22.4 (3.8)	24.0 (3.7)	24.6 (4.0)	23.8 (3.7)
Married, %	23.0	29.7	40.0	24.4	28.0	45.5
Educational Attainment, %						
Less than high school	9.8	10.8	0.0	11.1	12.0	9.1
High school diploma	38.7	37.8	0.0	43.8	24.0	63.6
Some college	51.5	51.4	100.0	45.2	64.0	27.3
Income Category, %						
<$16,000	41.9	32.1	20.0	31.1	27.3	16.7
$16,000-$49,999	47.9	57.1	80.0	59.8	68.2	16.7
≥$50,000	10.2	10.7	0.0	9.2	4.6	66.7
Neighborhood Poverty,[b] mean (SD)	30.6 (14.1)	17.7 (10.7)	24.4 (23.2)	31.0 (13.8)	20.6 (11.4)	7.2 (5.1)
Current Smoker, %	22.6	10.8	0.0	33.3	24.0	36.4
Physical Activity,[c] mean (SD)	216 (205)	220 (193)	235 (246)	443 (300)	580 (348)	416 (307)
Lost to Follow-Up, %	17.0	13.5	40.0	21.7	8.0	27.3
Died During Follow-Up, %	3.4	0.0	0.0	8.3	8.0	0.0

CARDIA, Coronary Artery Risk Development in Young Adults Study; SD, Standard Deviation.
[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.
[b] Neighborhood poverty measured as the percent of neighborhood residents living below the federal poverty threshold in the participant's census tract of residence (ranging from 0 to 100), using US Census data.
[c] Physical activity recorded in standardized units that represent both the duration and intensity of the activity.

eTable 4: Baseline Participant Characteristics of Black Women and Men in the CARDIA Study, by Racial Residential Segregation Category for CHICAGO Field Center

	Women (n=237)			Men (n=212)		
	Segregation Category:[a]			Segregation Category:[a]		
	High n=182	Medium n=24	Low n=31	High n=162	Medium n=18	Low n=32
Age in years, mean (SD)	23.7 (3.9)	24.4 (4.1)	24.5 (3.7)	23.3 (3.8)	23.8 (4.0)	24.8 (3.3)
Married, %	22.5	16.7	45.2	21.6	33.3	25.0
Educational Attainment, %						
Less than high school	11.5	8.3	6.5	32.7	22.2	12.5
High school diploma	40.1	41.7	22.6	32.7	44.5	31.3
Some college	48.4	50.0	71.0	34.6	33.3	56.3
Income Category, %						
<$16,000	30.9	22.7	20.8	24.6	20.0	12.0
$16,000-$49,999	55.9	59.1	62.5	69.1	53.3	60.0
≥$50,000	13.2	18.2	16.7	16.4	26.7	28.0
Neighborhood Poverty,[b] mean (SD)	32.2 (14.3)	14.9 (5.1)	12.0 (6.3)	32.5 (13.0)	14.0 (6.0)	12.7 (6.4)
Current Smoker, %	32.4	20.8	25.8	44.7	33.3	18.8
Physical Activity,[c] mean (SD)	314 (232)	370 (179)	316 (258)	628 (369)	566 (319)	601 (335)
Lost to Follow-Up, %	18.9	25.0	34.4	36.4	16.7	31.3
Died During Follow-Up, %	0.0	0.0	3.1	6.8	0.9	3.1

CARDIA, Coronary Artery Risk Development in Young Adults Study; SD, Standard Deviation.

[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

[b] Neighborhood poverty measured as the percent of neighborhood residents living below the federal poverty threshold in the participant's census tract of residence (ranging from 0 to 100), using US Census data.

[c] Physical activity recorded in standardized units that represent both the duration and intensity of the activity.

eTable5: Baseline Participant Characteristics of Black Women and Men in the CARDIA Study, by Racial Residential Segregation Category for MINNEAPOLIS Field Center

	Women (n=266)			Men (n=277)		
	Segregation Category:[a]			Segregation Category:[a]		
	High n=226	Medium n=31	Low n=9	High n=236	Medium n=30	Low n=11
Age in years, mean (SD)	23.9 (4.1)	24.3 (3.8)	24.0 (4.3)	23.9 (3.9)	24.7 (3.8)	25.4 (2.4)
Married, %	27.1	22.6	33.3	25.2	13.3	36.4
Educational Attainment, %						
Less than high school	22.7	16.1	11.1	27.1	16.7	9.1
High school diploma	40.0	35.5	55.6	48.7	50.0	36.4
Some college	37.3	48.4	33.3	24.2	33.3	54.6
Income Category, %						
<$16,000	47.0	45.8	50.0	44.6	35.0	11.1
$16,000-$49,999	40.4	33.3	50.0	49.7	65.0	77.8
≥$50,000	12.7	20.8	0.0	5.7	0.0	11.1
Neighborhood Poverty,[b] mean (SD)	19.7 (8.1)	18.2 (9.5)	7.5 (6.7)	21.4 (9.2)	21.5 (8.6)	4.2 (1.9)
Current Smoker, %	43.6	36.7	55.6	52.2	50.0	50.0
Physical Activity,[c] mean (SD)	322 (244)	378 (204)	329 (151)	580 (370)	519 (305)	687 (186)
Lost to Follow-Up, %	29.2	45.2	44.4	33.9	53.3	36.4
Died During Follow-Up, %	3.5	3.2	0.0	3.8	3.3	0.0

CARDIA, Coronary Artery Risk Development in Young Adults Study; SD, Standard Deviation.

[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

[b] Neighborhood poverty measured as the percent of neighborhood residents living below the federal poverty threshold in the participant's census tract of residence (ranging from 0 to 100), using US Census data.

[c] Physical activity recorded in standardized units that represent both the duration and intensity of the activity.

eTable 6: Baseline Participant Characteristics of Black Women and Men in the CARDIA Study, by Racial Residential Segregation Category for OAKLAND Field Center

	Women (n=383)			Men (n=298)		
	Segregation Category:[a]			Segregation Category:[a]		
	High n=301	Medium n=52	Low n=30	High n=226	Medium n=53	Low n=19
Age in years, mean (SD)	24.4 (3.8)	25.4 (3.4)	25.0 (3.9)	23.9 (3.8)	24.7 (3.4)	25.4 (3.3)
Married, %	29.0	34.6	30.0	35.4	35.9	21.1
Educational Attainment, %						
Less than high school	3.3	0.0	3.3	6.6	1.9	0.0
High school diploma	33.3	40.4	26.7	41.1	26.4	15.8
Some college	63.3	59.6	70.0	52.2	71.7	84.2
Income Category, %						
<$16,000	18.6	8.9	20.0	20.0	8.9	7.1
$16,000-$49,999	61.2	68.9	52.0	53.5	62.2	71.4
≥$50,000	20.2	22.2	28.0	26.5	28.9	21.4
Neighborhood Poverty,[b] mean (SD)	21.5 (9.2)	13.5 (5.7)	11.6 (10.7)	22.0 (9.0)	14.4 (6.0)	11.7 (10.7)
Current Smoker, %	34.8	27.5	26.7	30.1	17.0	15.8
Physical Activity,[c] mean (SD)	273 (240)	322 (282)	340 (222)	517 (338)	566 (320)	521 (418)
Lost to Follow-Up, %	22.6	30.8	23.3	29.7	30.2	26.3
Died During Follow-Up, %	1.3	0.0	0.0	5.8	7.6	5.3

CARDIA, Coronary Artery Risk Development in Young Adults Study; SD, Standard Deviation.

[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

[b] Neighborhood poverty measured as the percent of neighborhood residents living below the federal poverty threshold in the participant's census tract of residence (ranging from 0 to 100), using US Census data.

[c] Physical activity recorded in standardized units that represent both the duration and intensity of the activity.

eTable 7: Baseline Participant Characteristics of Black Women and Men in the CARDIA Study, by Racial Residential Segregation Category and Baseline Obesity Status

	Women (n=1469)		Men (n=1153)	
	Analysis Sample n=1156	Prevalent Obesity n=313	Analysis Sample n=1036	Prevalent Obesity n=117
Segregation Category %				
High	81.2	87.5	80.9	85.5
Medium	12.4	9.0	12.1	11.1
Low	6.4	3.5	7.1	3.4
Age in years, mean (SD)	24.2 (3.9)	25.3 (3.7)	24.0 (3.7)	25.4 (3.8)
Married, %	26.8	34.9	27.3	43.1
Educational Attainment, %				
Less than high school	10.4	13.7	16.9	15.4
High school diploma	37.1	34.5	40.7	30.8
Some college	52.5	51.8	42.4	53.9
Income Category, [a]	4.4 (1.9)	3.9 (1.9)	4.5 (1.8)	4.7 (1.5)
<$16,000	31.0	41.2	27.7	19.5
$16,000-$49,999	53.5	47.1	56.5	64.4
≥$50,000	15.5	11.7	15.8	16.1
Neighborhood Poverty,[b] mean (SD)	23.4 (12.8)	25.9 (13.4)	24.0 (12.6)	22.7 (11.5)
Current Smoker, %	31.9	29.5	37.5	32.8
Physical Activity,[c] mean (SD)	285 (235)	248 (192)	541 (345)	467 (316)
Field Center, %				
Birmingham, AL	24.0	25.2	24.4	34.2
Chicago, IL	20.5	24.3	20.5	19.7
Minneapolis, MN	22.8	16.3	26.6	21.4
Oakland, CA	32.7	34.2	28.5	24.8

CARDIA, Coronary Artery Risk Development in Young Adults Study; SD, Standard Deviation.

[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i^*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i^*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

[b] Neighborhood poverty measured as the percent of neighborhood residents living below the federal poverty threshold in the participant's census tract of residence (ranging from 0 to 100), using US Census data.

[c] Physical activity recorded in standardized units that represent both the duration and intensity of the activity.

eTable 8: Odds of Obesity at Baseline by Racial Residential Segregation[a] among Black Women and Men in the CARDIA Study, 1985/1986

	Women (n=1469) OR (95% CI) [b]	Men (n=1153) OR (95% CI) [b]
Baseline Segregation[a]		
High	2.1 (0.96, 4.4)	2.7 (0.79, 8.9)
Medium	1.5 (0.64, 3.7)	2.0 (0.52, 7.7)
Low	1.0	1.0

CARDIA, Coronary Artery Risk Development in Young Adults Study; CI, Confidence Interval; OR, Odds Ratio

[a] Levels of racial residential segregation categorized into high, medium, and low based on the value of the local G_i*statistic, which measures the deviation of the racial composition of the census tract from the larger area. Segregation levels based on a G_i*statistic z-score of less than zero, zero to 1.96, and greater than 1.96, represent low, medium, and high categories, respectively.

[b] Odds of obesity by neighborhood segregation level as recorded at baseline, simultaneously adjusted for all covariates as recorded at baseline: age, field center, marital status, education, physical activity, current smoking, income.