SPECIAL ARTICLE

Neighborhoods, Obesity, and Diabetes — A Randomized Social Experiment

Jens Ludwig, Ph.D., Lisa Sanbonmatsu, Ph.D., Lisa Gennetian, Ph.D., Emma Adam, Ph.D., Greg J. Duncan, Ph.D., Lawrence F. Katz, Ph.D., Ronald C. Kessler, Ph.D., Jeffrey R. Kling, Ph.D., Stacy Tessler Lindau, M.D., Robert C. Whitaker, M.D., M.P.H., and Thomas W. McDade, Ph.D.

ABSTRACT

BACKGROUND

The question of whether neighborhood environment contributes directly to the development of obesity and diabetes remains unresolved. The study reported on here uses data from a social experiment to assess the association of randomly assigned variation in neighborhood conditions with obesity and diabetes.

METHODS

From 1994 through 1998, the Department of Housing and Urban Development (HUD) randomly assigned 4498 women with children living in public housing in high-poverty urban census tracts (in which ≥40% of residents had incomes below the federal poverty threshold) to one of three groups: 1788 were assigned to receive housing vouchers, which were redeemable only if they moved to a low-poverty census tract (where <10% of residents were poor), and counseling on moving; 1312 were assigned to receive unrestricted, traditional vouchers, with no special counseling on moving; and 1398 were assigned to a control group that was offered neither of these opportunities. From 2008 through 2010, as part of a long-term follow-up survey, we measured data indicating health outcomes, including height, weight, and level of glycated hemoglobin (HbA_{1C}).

RESULTS

As part of our long-term survey, we obtained data on body-mass index (BMI, the weight in kilograms divided by the square of the height in meters) for 84.2% of participants and data on glycated hemoglobin level for 71.3% of participants. Response rates were similar across randomized groups. The prevalences of a BMI of 35 or more, a BMI of 40 or more, and a glycated hemoglobin level of 6.5% or more were lower in the group receiving the low-poverty vouchers than in the control group, with an absolute difference of 4.61 percentage points (95% confidence interval [CI], −8.54 to −0.69), 3.38 percentage points (95% CI, −6.39 to −0.36), and 4.31 percentage points (95% CI, −7.82 to −0.80), respectively. The differences between the group receiving traditional vouchers and the control group were not significant.

CONCLUSIONS

The opportunity to move from a neighborhood with a high level of poverty to one with a lower level of poverty was associated with modest but potentially important reductions in the prevalence of extreme obesity and diabetes. The mechanisms underlying these associations remain unclear but warrant further investigation, given their potential to guide the design of community-level interventions intended to improve health. (Funded by HUD and others.)

From the University of Chicago, Chicago (J.L., S.T.L.); the National Bureau of Economic Research (J.L., L.S., L.F.K., J.R.K.) and Harvard University (L.F.K.) — both in Cambridge, MA; the Brookings Institution (L.G.) and the Congressional Budget Office (J.R.K.) — both in Washington, DC; Northwestern University, Evanston, IL (E.A., T.W.M.); the University of California at Irvine, Irvine (G.J.D.); Harvard Medical School, Boston (R.C.K.); and Temple University, Philadelphia (R.C.W.). Address reprint requests to Dr. Ludwig at the University of Chicago, 1155 E. 60th St., Chicago, IL 60637, or at jludwig@uchicago.edu.

N Engl J Med 2011;365:1509-19.

Many observational studies have shown that neighborhood attributes such as poverty and racial segregation are associated with increased risks of obesity and diabetes, even after adjustment for observed individual and family-related factors.[1-4] In response, the U.S. surgeon general has called for efforts to "create neighborhood communities that are focused on healthy nutrition and regular physical activity, where the healthiest choices are accessible for all citizens."[5]

Previous studies have suggested several pathways through which neighborhoods might influence health. Changes in the built environment (e.g., the addition of grocery stores or spaces where residents can exercise) might affect health-related behaviors and outcomes such as obesity.[4,6-8] Proximity to health care providers might influence the detection or management of health problems. Neighborhood safety might influence exercise level, diet, or level of stress.[4,9] Social norms for health-related behaviors may vary across neighborhoods.[10,11]

It is unclear whether neighborhood environments directly contribute to the development of obesity and diabetes. People living in neighborhoods with high poverty rates differ in many ways from those living in neighborhoods with lower poverty rates, only some of which can be adequately measured in observational studies. These unmeasured individual characteristics may be responsible for variations in health among different neighborhoods. Inferences concerning the influence of neighborhood may be more credible if they are based on randomized studies in which otherwise similar people are encouraged to live in different types of neighborhoods. Using data from Moving to Opportunity (MTO), a large demonstration project intended to uncover the effects of neighborhood characteristics across a range of social and health outcomes in families, we examined the association of randomly assigned variations in neighborhood conditions with obesity and diabetes.

METHODS

STUDY DESIGN

The MTO demonstration project was designed and implemented by the Department of Housing and Urban Development (HUD) with the primary purpose of better understanding the effects of residential location on "employment, income, education, and well-being."[12] Families with children (defined as family members younger than 18 years of age) living in Baltimore, Boston, Chicago, Los Angeles, or New York in selected public housing developments in census tracts with poverty rates of 40% or more in 1990 were eligible. From 1994 through 1998, families were invited by local housing authorities to participate in a randomized lottery to receive a rent-subsidy voucher.[13] One quarter of eligible families applied.[13]

The analysis reported here focuses on one woman from each family, usually the household head, who was interviewed between 2008 and 2010. This research was approved by the Office of Management and Budget and by the institutional review boards at HUD, the National Bureau of Economic Research, and relevant universities. HUD assisted with the design of the data-collection protocol for the long-term MTO study and reviewed the manuscript before submission to ensure that the confidentiality of MTO program participants was not violated; HUD did not screen the manuscript for other purposes.

INTERVENTIONS AND RANDOMIZATION

Participating families were randomly assigned to one of three groups. Families assigned to receive low-poverty vouchers were offered a standard rent-subsidy voucher but were required to use it in a census tract with a low poverty rate (<10% in 1990). Vouchers served as subsidies for private-market housing and were equal in value to the difference between a rent threshold minus the family contribution to the rent (30% of income, which is identical to the contribution required for public housing).[14] Families remained eligible for vouchers as long as they met the income criteria and other requirements. Census tracts contain between 2500 and 8000 people and were defined by the Census Bureau as being "homogeneous with respect to population characteristics, economic status, and living conditions."[15] Families that received low-poverty vouchers also received short-term counseling to help with their housing search.[16,17] After 1 year, these families could use the voucher to relocate to a different tract, regardless of the poverty rate in that tract. In the traditional-voucher group, families were given a standard voucher with no restrictions on where they could reside; they were not provided with counseling. This group was included to distinguish the effects of moving with a voucher from the effects of moving to a lower-poverty area. Families in the control group were offered no new assistance.

Randomization was conducted for HUD by Abt Associates with the use of a computerized random-number generator.[16] HUD selected sample sizes for power to detect effects on the primary outcomes of the MTO study (i.e., employment, income, and education).[17] During the study, Abt Associates adjusted the random-assignment rates of later entrants on the basis of acceptance rates among earlier entrants to equalize the statistical power of different cross-group comparisons.[18]

DATA COLLECTION

MTO applicants completed a baseline survey that contained questions concerning "the people who live with you, your housing, your neighborhood, and your work experiences."[19] Among the few baseline measures related to health was the receipt of Supplemental Security Income, a benefit provided for aged, blind, and disabled persons.

After randomization and completion of the baseline survey by participants, HUD engaged our team to follow the families in order to assess long-term outcomes, including some related to health. Data on outcomes were collected by the Survey Research Center at the University of Michigan from June 2008 through April 2010 — an average of 12.6 years after randomization (range, 10.0 to 15.4). The sample frame included one adult from each family in the group that received low-poverty vouchers and the control group and from a randomly selected two thirds of the families in the traditional-voucher group (this group was undersampled for budgetary reasons).

Candidates for study participation were offered $50 to complete our survey[19] and another $25 to undergo height and weight assessments and provide a blood sample. Written informed consent was obtained before the interviews began; the interviews were usually conducted in the participant's home and were completed in 2 hours. Interviewers were unaware of group assignments. The long-term survey design involved two-phase sampling. In phase 1, interviewers sought to interview everyone in the survey sample frame. Once a response rate of 75 to 80% was reached, the interviewers began phase 2, which involved trying to reach a probability subsample of 35% of the families that could not be surveyed in phase 1.[20]

Obesity Assessment

Height and weight were measured with the use of modified protocols from the University of Michigan Health and Retirement Study.[21] Respondents removed heavy outer clothing and items from their pockets and stood with heels and shoulders against a wall. Height was marked on the wall with the use of a rafter angle square and measured to the nearest 0.6 cm (0.25 in.) with a metal tape measure. Weight was measured to the nearest 0.23 kg (0.5 lb) with a digital electronic floor scale (Health o meter [Pelstar], model 800KL), which had a maximum capacity of 180 kg (397 lb).[22] When weight or height could not be measured, that reported by the participant was recorded.

Diabetes Assessment

Up to five drops of whole-blood capillary samples were collected on specimen-collection paper (Whatman no. 903) with an autoretractable lancet finger stick[23] after it had been determined that the participant had no history of a bleeding disorder and was not taking medication that could affect coagulation. Samples were assayed for glycated hemoglobin (HbA_{1C}) at a laboratory with Clinical Laboratory Improvement Amendments certification (FlexSite Diagnostics) with the use of a Roche COBAS Integra immunochemical analyzer that was validated for use with dried blood spots and certified by the National Glycohemoglobin Standardization Program. A single measurement of glycated hemoglobin provides an integrated assessment of a person's average blood glucose levels over the preceding several months; fasting is not required before a sample is obtained.[24]

RESPONSE RATES

To account for two-phase sampling, we calculated effective response rates.[20] For phases 1 and 2, the response rates were calculated as the number of participants with data from each phase, divided by the sum of the number of participants with data and the number with missing data (because the participant declined to provide the data, was incapacitated, had died, or was not contacted) from that phase. Response rates were calculated in accordance with definition RR1w from the American Association for Public Opinion Research.[25] Thus, we calculated the overall response rate as $(P1 \times R1) + (P2 \times R2)$, where P1 and P2 are the share of the total sample from phase 1 and phase 2, respectively, and R1 and R2 are the response rates in phase 1 and phase 2, respectively.

OUTCOME MEASURES

We created dichotomous measures for obesity by applying commonly used criteria based on the body-mass index (BMI, the weight in kilograms divided by the square of the height in meters): 30 or more, 35 or more, and 40 or more.[26] We defined diabetes as a glycated hemoglobin level of 6.5% or more, as recommended by the American Diabetes Association.[27,28]

HUD tracked participants' addresses from baseline to the beginning of long-term follow-up. To illustrate the nature of the change in the neighborhoods where participants lived, we geocoded addresses and linked them to census-tract attributes. In addition, our long-term survey included questions on access to health care, neighborhood safety, and indicators of "collective efficacy" (the social cohesion of the neighborhood).[29]

STATISTICAL ANALYSIS

We first carried out an omnibus F-test to determine whether differences in baseline characteristics across groups were jointly zero.[30] In our main analyses, we used the intention-to-treat principle, comparing differences in average outcomes for controls with those for all members of the two groups receiving vouchers, regardless of whether a family had moved as a result of study participation. The effects on continuous dependent variables were calculated with the use of linear regression, and the effects on dichotomous variables were calculated with the use of logistic regression and are presented as average marginal effects; adjustments were made for baseline covariates to improve precision. All estimates weighted individual participants according to the inverse of the probability of assignment to a particular group, with phase 2 participants also weighted according to the inverse of the likelihood of selection for phase 2 subsampling.[20] We calculated Huber–White robust standard errors to adjust for heteroskedasticity.

We also used instrumental-variable methods to try to estimate the association between health and change in residence with the use of a voucher (the complier average causal effect, which in the MTO demonstration project equals the estimated effect of treatment on the treated)[31] and to estimate a dose–response effect.[32] (For details see

Figure 1 (facing page). Screening and Randomization.
BMI denotes body-mass index. P1 (the share of the total sample in phase 1) = phase 1 subtotal ÷ (phase 1 + phase 2 subtotals). P2 (the share of the total sample in phase 2) = phase 2 subtotal ÷ (phase 1 + phase 2 subtotals). R1 (the response rate from phase 1) = phase 1 analysis sample ÷ phase 1 subtotal. R2 (the response rate from phase 2) = phase 2 analysis sample ÷ (phase 2 subtotal − phase 2 randomly selected for exclusion). The analysis sample refers to the sample for the BMI analysis or the sample for the glycated hemoglobin analysis. The effective response rate = (P1 × R1) + (P2 × R2).

Tables 1 through 9 in the Supplementary Appendix, available with the full text of this article at NEJM.org; these tables also provide data on selected means according to study group and compliance status.) For all end points, a two-sided P value of less than 0.05 was considered to indicate statistical significance, with no adjustment for multiple comparisons. Analyses were performed with the use of Stata software, version 11.0, special edition (StataCorp).[33]

RESULTS

STUDY POPULATION

A total of 4498 families underwent randomization to one of three study groups between 1994 and 1998 (Fig. 1). During the follow-up period, from 2008 through 2010, the effective response rates for data on BMI and glycated hemoglobin level were 84.7% and 70.1%, respectively, for the group that received low-poverty vouchers; 82.8% and 73.7%, respectively, for the group that received traditional vouchers; and 84.4% and 71.3%, respectively, for the control group.

Table 1 presents the baseline characteristics of respondents for whom valid data on BMI or glycated hemoglobin level were collected. (Information on additional baseline characteristics is provided in Table 1 in the Supplementary Appendix.) Most women in the study were unmarried and either black or Hispanic. There were no significant differences in the 57 baseline characteristics between the groups that received low-poverty vouchers or traditional vouchers and the control group ($P=0.93$ and $P=0.35$, respectively).

EFFECTS OF THE INTERVENTION ON NEIGHBORHOOD CONDITIONS

Among the families assigned to receive low-poverty vouchers, 48% used the vouchers; among those assigned to receive traditional vouchers, 63% used

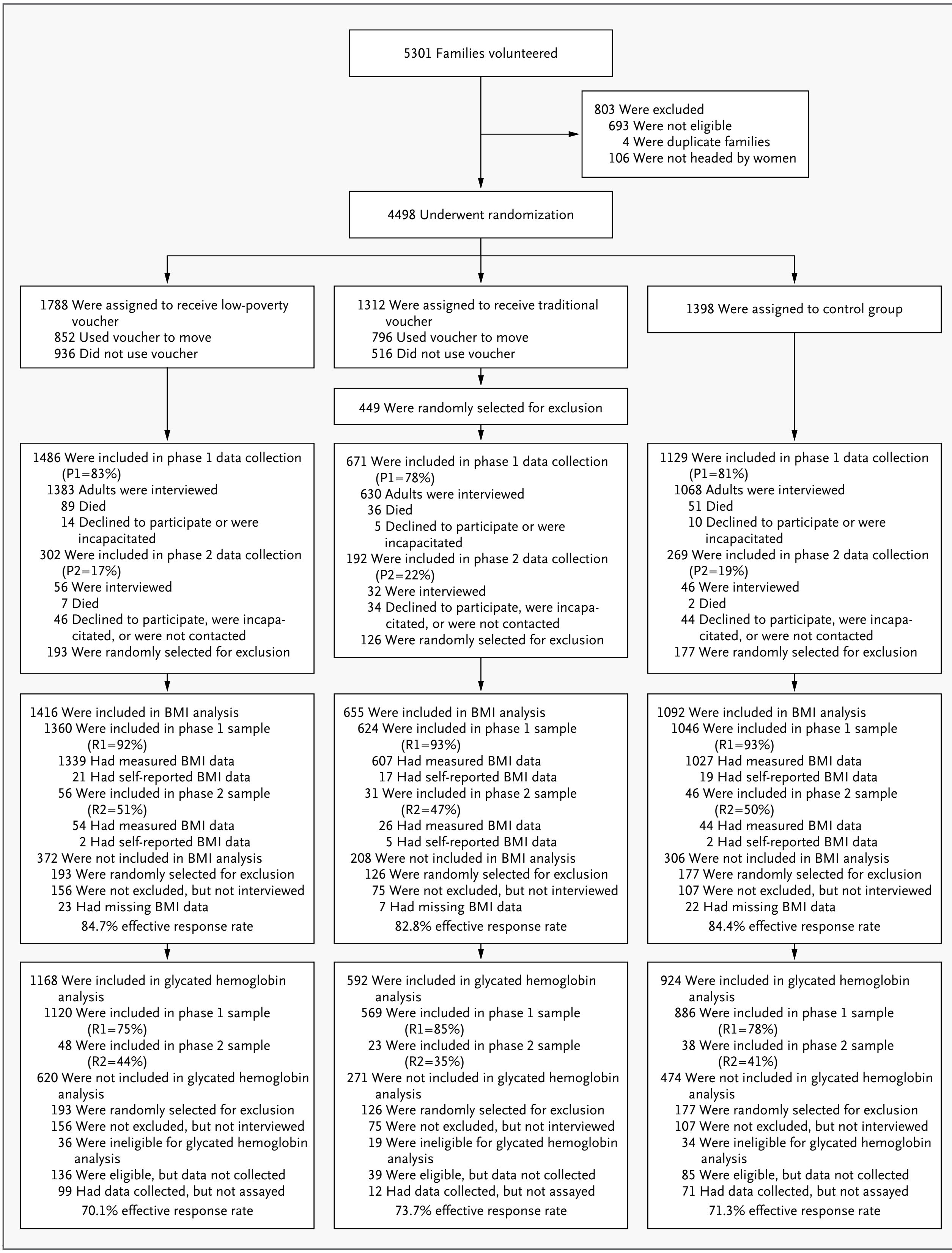
5301 Families volunteered
803 Were excluded
693 Were not eligible
4 Were duplicate families
106 Were not headed by women
4498 Underwent randomization
1788 Were assigned to receive low-poverty voucher
852 Used voucher to move
936 Did not use voucher
1312 Were assigned to receive traditional voucher
796 Used voucher to move
516 Did not use voucher
1398 Were assigned to control group
449 Were randomly selected for exclusion
1486 Were included in phase 1 data collection (P1=83%)
1383 Adults were interviewed
89 Died
14 Declined to participate or were incapacitated
302 Were included in phase 2 data collection (P2=17%)
56 Were interviewed
7 Died
46 Declined to participate, were incapacitated, or were not contacted
193 Were randomly selected for exclusion
671 Were included in phase 1 data collection (P1=78%)
630 Adults were interviewed
36 Died
5 Declined to participate or were incapacitated
192 Were included in phase 2 data collection (P2=22%)
32 Were interviewed
34 Declined to participate, were incapacitated, or were not contacted
126 Were randomly selected for exclusion
1129 Were included in phase 1 data collection (P1=81%)
1068 Adults were interviewed
51 Died
10 Declined to participate or were incapacitated
269 Were included in phase 2 data collection (P2=19%)
46 Were interviewed
2 Died
44 Declined to participate, were incapacitated, or were not contacted
177 Were randomly selected for exclusion
1416 Were included in BMI analysis
1360 Were included in phase 1 sample (R1=92%)
1339 Had measured BMI data
21 Had self-reported BMI data
56 Were included in phase 2 sample (R2=51%)
54 Had measured BMI data
2 Had self-reported BMI data
372 Were not included in BMI analysis
193 Were randomly selected for exclusion
156 Were not excluded, but not interviewed
23 Had missing BMI data
84.7% effective response rate
655 Were included in BMI analysis
624 Were included in phase 1 sample (R1=93%)
607 Had measured BMI data
17 Had self-reported BMI data
31 Were included in phase 2 sample (R2=47%)
26 Had measured BMI data
5 Had self-reported BMI data
208 Were not included in BMI analysis
126 Were randomly selected for exclusion
75 Were not excluded, but not interviewed
7 Had missing BMI data
82.8% effective response rate
1092 Were included in BMI analysis
1046 Were included in phase 1 sample (R1=93%)
1027 Had measured BMI data
19 Had self-reported BMI data
46 Were included in phase 2 sample (R2=50%)
44 Had measured BMI data
2 Had self-reported BMI data
306 Were not included in BMI analysis
177 Were randomly selected for exclusion
107 Were not excluded, but not interviewed
22 Had missing BMI data
84.4% effective response rate
1168 Were included in glycated hemoglobin analysis
1120 Were included in phase 1 sample (R1=75%)
48 Were included in phase 2 sample (R2=44%)
620 Were not included in glycated hemoglobin analysis
193 Were randomly selected for exclusion
156 Were not excluded, but not interviewed
36 Were ineligible for glycated hemoglobin analysis
136 Were eligible, but data not collected
99 Had data collected, but not assayed
70.1% effective response rate
592 Were included in glycated hemoglobin analysis
569 Were included in phase 1 sample (R1=85%)
23 Were included in phase 2 sample (R2=35%)
271 Were not included in glycated hemoglobin analysis
126 Were randomly selected for exclusion
75 Were not excluded, but not interviewed
19 Were ineligible for glycated hemoglobin analysis
39 Were eligible, but data not collected
12 Had data collected, but not assayed
73.7% effective response rate
924 Were included in glycated hemoglobin analysis
886 Were included in phase 1 sample (R1=78%)
38 Were included in phase 2 sample (R2=41%)
474 Were not included in glycated hemoglobin analysis
177 Were randomly selected for exclusion
107 Were not excluded, but not interviewed
34 Were ineligible for glycated hemoglobin analysis
85 Were eligible, but data not collected
71 Had data collected, but not assayed
71.3% effective response rate

Table 1. Baseline Characteristics of the Study Population.*

Characteristic	Low-Poverty Voucher (N=1425)	Traditional Voucher (N=657)	Control (N=1104)
		number (percent)	
Age†			
≤35 yr	196 (14.6)	94 (13.5)	163 (14.7)
36–40 yr	310 (21.5)	156 (23.9)	253 (23.3)
41–45 yr	347 (23.5)	143 (21.7)	257 (23.2)
46–50 yr	273 (18.6)	124 (20.5)	194 (17.1)
>50 yr	299 (21.7)	140 (20.4)	237 (21.7)
Race or ethnic group‡			
Black	973 (65.0)	393 (63.9)	706 (66.1)
Other nonwhite	339 (28.1)	194 (27.6)	288 (26.8)
White	92 (8.5)	52 (7.1)	88 (6.9)
Hispanic	404 (31.5)	235 (33.0)	346 (30.3)
Never married	874 (62.6)	395 (63.5)	692 (64.3)
Age <18 yr at birth of first child	347 (25.1)	163 (28.0)	265 (25.0)
Employed	368 (27.1)	176 (26.0)	258 (23.9)
Enrolled in school	216 (16.0)	113 (17.7)	172 (16.9)
Received high-school diploma	565 (38.3)	233 (34.3)	407 (35.9)
Received certificate of General Educational Development (GED)	235 (16.2)	124 (18.7)	204 (19.9)
Receives Supplemental Security Income§	221 (15.9)	107 (17.1)	171 (16.3)

* Numbers are raw, unweighted data. Percentages were calculated with the use of sample weights to account for changes in random-assignment ratios across randomized groups and for subsample interviews. Percentages include imputed values. The sample consisted of women for whom valid data on body-mass index or glycated hemoglobin level were available in the long-term follow-up study. An omnibus F-test failed to reject the null hypothesis that the baseline characteristics reported were the same across study groups. (P=0.41 for the comparison of the characteristics of the low-poverty–voucher group with the control group; P=0.77 for the comparison of the traditional-voucher group with the control group.) See Table 1 in the Supplementary Appendix for additional baseline characteristics and related P values.

† The age listed was that calculated as of December 31, 2007, just before the long-term follow-up began in June 2008.

‡ Race categories do not sum to the total number because of missing data (for 21 women in the low-poverty–voucher group, 18 in the traditional-voucher group, and 22 in the control group). A Hispanic person could be a member of any race.

§ Supplemental Security Income is a federal assistance program for aged, blind, and disabled people.

the vouchers. The association between study-group assignment and neighborhood poverty rate was significant. One year after randomization, the census-tract poverty rate for the group that received low-poverty vouchers was 17.1 percentage points lower than that for the control group, for which the poverty rate was 50.0% (95% confidence interval [CI], −18.6 to −15.6) (Table 2), a change of 1.4 SD in the national census-tract poverty distribution (Table 2 in the Supplementary Appendix). This association between low-poverty vouchers and a reduced poverty rate attenuated over time, in part because families in the control group eventually moved to lower-poverty areas without assistance from the MTO program. Ten years after randomization, the mean poverty rate in the group that received low-poverty vouchers was 4.9 percentage points lower than the rate in the control group, which was 33.0%. Estimates of the effect of treatment on the treated were twice as large as the intention-to-treat estimates for the group that received low-poverty vouchers and were 1.5 times

Table 2. Residential Mobility, Poverty Rate, and Census-Tract Characteristics, According to Study Group.*

Variable	Control	Low-Poverty Voucher			Traditional Voucher		
	Mean	Intention-to-Treat Estimate (95% CI)†	P Value	Mean	Intention-to-Treat Estimate (95% CI)†	P Value	Mean
Mean no. of moves‡	2.1	0.57 (0.42 to 0.71)	<0.001	2.7	0.58 (0.38 to 0.79)	<0.001	2.7
Poverty rate in census tract (%)§							
Baseline	53.1	−0.37 (−1.23 to 0.50)	0.41	52.5	−0.37 (−1.55 to 0.81)	0.54	52.9
At 1 yr	50.0	−17.06 (−18.57 to −15.56)	<0.001	32.7	−13.50 (−15.33 to −11.67)	<0.001	36.6
At 5 yr	39.9	−9.78 (−11.25 to −8.31)	<0.001	30.0	−6.26 (−8.41 to −4.11)	<0.001	33.0
At 10 yr	33.0	−4.86 (−6.23 to −3.48)	<0.001	28.3	−2.87 (−4.80 to −0.95)	0.003	29.2
Mean census-tract characteristics (%)¶							
Poor‖	39.6	−9.14 (−10.26 to −8.02)	<0.001	30.4	−6.07 (−7.53 to −4.61)	<0.001	32.9
Minorities	88.0	−6.23 (−7.58 to −4.89)	<0.001	81.9	−0.99 (−2.88 to 0.90)	0.30	85.8
Household headed by a woman	54.3	−7.95 (−9.08 to −6.82)	<0.001	46.2	−5.03 (−6.55 to −3.51)	<0.001	48.7
College graduate	16.1	4.49 (3.68 to 5.30)	<0.001	20.5	1.41 (0.29 to 2.52)	0.01	18.4
Respondents reporting collective efficacy (%)**							
At 4–7 yr	54.0	10.61 (6.46 to 14.76)	<0.001	65.4	5.30 (0.53 to 10.07)	0.03	59.9
At 10–15 yr	58.9	8.20 (4.20 to 12.21)	<0.001	67.2	0.80 (−5.16 to 6.76)	0.79	62.4
Respondents reporting feeling safe or very safe on streets near home during the day (%)							
At 4–7 yr	74.9	9.14 (5.77 to 12.52)	<0.001	84.6	8.95 (5.16 to 12.73)	<0.001	84.4
At 10–15 yr	80.7	3.70 (0.52 to 6.87)	0.02	84.2	5.00 (0.50 to 9.50)	0.03	85.1
Respondents reporting having at least one friend who graduated from college (%)							
At 4–7 yr	40.8	6.90 (2.63 to 11.17)	0.002	48.0	4.55 (−0.22 to 9.33)	0.06	45.3
At 10–15 yr	53.4	6.90 (2.74 to 11.06)	0.001	60.4	−2.11 (−8.33 to 4.11)	0.51	53.2
Respondents reporting access to local health care services, excluding emergency room (%)							
At 4–7 yr	89.7	−1.35 (−4.13 to 1.43)	0.34	88.8	−0.21 (−3.15 to 2.73)	0.89	89.5
At 10–15 yr	93.4	−1.36 (−3.49 to 0.77)	0.21	92.1	0.64 (−2.11 to 3.40)	0.65	95.2

* The analysis sample consisted of women with a valid BMI or glycated hemoglobin measurement. Analyses of number of moves and census-tract characteristics were further limited to participants with valid addresses at baseline and years 1, 5, and 10. The intention-to-treat estimates come from a regression that compares average outcomes across randomly assigned groups, with statistical control for baseline characteristics, which may differ slightly from the difference in raw group means presented here. See the Supplementary Appendix for the sample sizes used.

† Intention-to-treat estimates compare the average of the outcomes for everyone assigned to the intervention group with the average of the outcomes for controls, with adjustment for the set of baseline covariates shown in Table 1 and indicators for survey-sample release (families were randomly selected with regard to the time at which they would first be contacted about participation in the long-term follow-up study), site, and random-assignment periods. The effects on continuous dependent variables were calculated with the use of linear regression; the effects on dichotomous variables were calculated with the use of logistic regression and are presented as average marginal effects.

‡ The total number of moves is the number from the time of randomization (1994 through 1998) to the beginning of long-term follow-up (May 2008).

§ Census-tract characteristics were recorded as of the time when a family lived in the tract and were interpolated with the use of 1990 and 2000 decennial census data and data from the American Community Survey, 2005 to 2009.

¶ Average duration-weighted census-tract characteristics give more weight to tracts in which families spent relatively more time during the study period.

‖ The term “poor” is defined as having an annual income below the federal government’s poverty threshold.

** Collective efficacy is defined as the likelihood that adults will take action in response to youth spraying graffiti on local buildings. See Sampson et al. for more details on collective efficacy.[29]

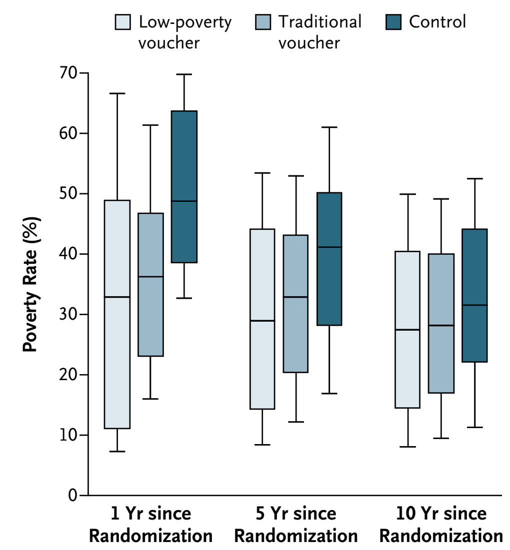

Figure 2. Census-Tract Poverty Rate According to Study Group and Years since Randomization.

The horizontal line in the middle of each vertical bar indicates the median of the census-tract poverty rates within each randomly assigned group, the upper and lower boundaries of each bar mark the 75th and 25th percentiles, and the I bars (whiskers) mark the 90th and 10th percentiles. Census tracts are small geographic areas that usually contain between 2500 and 8000 people and were defined by the Census Bureau to correspond to local communities that have relatively homogeneous population characteristics. The census-tract poverty rate for families in the study 1, 5, or 10 years after randomization was linearly interpolated from data in the 1990 and 2000 decennial censuses and the American Community Survey for 2005 through 2009. The sample includes 3026 women for whom there was a valid measure of body-mass index or a valid measure of the glycated hemoglobin level in addition to valid addresses at baseline and at the three time points shown.

as large for the group that received traditional vouchers (see the Supplementary Appendix). In an analysis of the 25th percentile of each group's census-tract poverty distribution (Fig. 2), the differences across groups were even larger.

Study-group assignment was also associated with other neighborhood attributes, including safety and collective efficacy. However, there was no significant association between study-group assignment and access to routine medical care.

PRIMARY OUTCOMES

At 10 to 15 years of follow-up, assignment to the low-poverty–voucher group was associated with a decreased risk of extreme obesity and diabetes. Among the women in the control group, 58.6% had a BMI of 30 or more, 35.5% had a BMI of 35 or more, 17.7% had a BMI of 40 or more, and 20.0% had a glycated hemoglobin level of 6.5% or more. In the intention-to-treat analysis, the women in the group that received low-poverty vouchers, as compared with the women in the control group, had lower prevalences of a BMI of 35 or more (−4.61 percentage points; 95% CI, −8.54 to −0.69; P=0.02, calculated without adjustment for multiple comparisons) and of a BMI of 40 or more (−3.38 points; 95% CI, −6.39 to −0.36; P=0.03), representing relative reductions of 13.0% and 19.1%, respectively (Table 3). The women in the group that received low-poverty vouchers also had a lower prevalence of glycated hemoglobin levels of 6.5% or more, as compared with the women in the control group (−4.31 percentage points; 95% CI, −7.82 to −0.80; P=0.02), a relative reduction of 21.6%.

The differences in outcomes for BMI and diabetes between the group that received traditional vouchers and the control group were not significant at the level of 0.05. The difference in outcomes between the two voucher groups was not significant for any BMI threshold, but there was a trend toward a significant difference in the prevalence of glycated hemoglobin levels of 6.5% or more (P=0.05).

We found no significant differences across subgroups defined by baseline characteristics in effects on health in post hoc analyses, including baseline age or demonstration site (Tables 6 and 7 in the Supplementary Appendix).

Our dose–response model revealed that adults who spent more time in lower-poverty census tracts had greater improvements in diabetes and BMI outcomes (Table 9 in the Supplementary Appendix). We tested for the presence of nonlinear relationships between neighborhood attributes and these health outcomes, but these tests had low statistical power.

DISCUSSION

As compared with the control group, the group with a randomly assigned opportunity to use a voucher to move to a neighborhood with a lower poverty rate had lower prevalences of a BMI of 35 or more, a BMI of 40 or more, and a glycated hemoglobin level of 6.5% or more, representing

Table 3. Body-Mass Index (BMI) and Glycated Hemoglobin Level at Follow-up, According to Study Group.*

Variable	Control	Low-Poverty Voucher			Traditional Voucher		
	Prevalence (%)	Intention-to-Treat Estimate (95% CI)†	P Value	Prevalence (%)	Intention-to-Treat Estimate (95% CI)†	P Value	Prevalence (%)
BMI‡							
≥30	58.6	−1.19 (−5.41 to 3.02)	0.58	57.5	−0.14 (−6.27 to 5.98)	0.96	58.4
≥35	35.5	−4.61 (−8.54 to −0.69)	0.02	31.1	−5.34 (−11.02 to 0.34)	0.07	30.8
≥40	17.7	−3.38 (−6.39 to −0.36)	0.03	14.4	−3.58 (−7.95 to 0.80)	0.11	15.4
Glycated hemoglobin§							
≥6.5%	20.0	−4.31 (−7.82 to −0.80)	0.02	16.3	−0.08 (−5.18 to 5.02)	0.98	20.6

* The analysis sample consisted of women with a valid BMI measurement (for the BMI analysis) or a valid glycated hemoglobin measurement (for the glycated hemoglobin analysis) in the long-term follow-up data collection. See the Supplementary Appendix for the sample sizes used.
† Intention-to-treat estimates compare the average outcomes for all participants assigned to an intervention group with the average outcomes for controls, with adjustment for the set of baseline covariates shown in Table 1 and indicators for survey-sample release and random-assignment periods. The effects are calculated with the use of logistic regression and are presented as average marginal effects.
‡ BMI (the weight in kilograms divided by the square of the height in meters) was calculated from measured height and weight for most adults as part of the long-term follow-up data collection. Self-reported values were used for 23 observations in the low-poverty–voucher group, 22 observations in the traditional-voucher group, and 21 observations in the control group.
§ Glycated hemoglobin (HbA_{1c}) was assayed from dried blood spots collected as part of the long-term follow-up data collection.

relative reductions of 13.0%, 19.1%, and 21.6%, respectively. The magnitudes of the associations with health were larger still for participants who moved with a voucher that was restricted to use in a low-poverty area than they were for the intention-to-treat estimates for all participants who received the restricted voucher and are consistent with the effect sizes reported in previous observational studies.[3] Because we generated estimates for several BMI cutoff points, our estimates for the associations between program participation and extreme obesity may be marginally significant.

Approximately half the participants randomly assigned to receive low-poverty vouchers used these vouchers, and many of the families in the control group subsequently moved to areas with lower poverty rates. Neither imperfect program compliance nor crossover compromises the internal validity of our intention-to-treat estimates, but these factors may reduce the statistical power of the analyses.

Although we could not reject the null hypothesis that the association of the traditional voucher with obesity is equal to zero or that the association is the same as that for the low-poverty voucher, the difference between the prevalence of a glycated hemoglobin level of 6.5% or more in the group that received low-poverty vouchers and the prevalence in the group that received traditional vouchers approached significance. This finding is consistent with that of previous MTO studies in which outcomes not involving health suggested that changes in the neighborhood environment, rather than the act of moving itself, are responsible for these effects[32]; it is also consistent with our finding that low-poverty vouchers and traditional vouchers had different associations with neighborhood attributes that may affect health (Table 2).

An MTO study published in 2007, which measured self-reported outcomes 4 to 7 years after randomization, showed that the prevalence of obesity (defined as a BMI of 30 or more) among adults assigned to receive low-poverty vouchers was 42.0%, as compared with 46.8% for the control group.[32] Use of self-reported measures raises concerns about the Hawthorne effect and the possibility that the neighborhood environment could affect self-reporting. The 2007 study was not informative with regard to long-term health effects because the problem of fade-out (attenuation in the differences in outcomes between treatment groups and control groups) is pervasive in social experiments, and the study did not show results for the most costly condition associated with obesity — diabetes.

The present study has several strengths, including the use of a large social experiment to overcome concerns about selection bias associated with epidemiologic studies and the collection of physical measurements for health outcomes 10 to

15 years after randomization. The study also had the effect of causing a relatively homogeneous group of people to live in a wider range of neighborhoods than is usual for epidemiologic studies. Because the moves led to changes in neighborhoods as defined by the most commonly used markers of neighborhood areas (e.g., tracts and ZIP Codes), the study inherently addresses the potential for measurement error that can result when epidemiologic studies use the wrong geographic proxy for "neighborhood."[34]

Our study also has several limitations. First, it is possible that the participants for whom outcomes were not available in our long-term study would have differed systematically across the randomized groups in unobservable attributes. Second, our use of a glycated hemoglobin level of 6.5% or more does not account for people with successfully treated diabetes. Third, the baseline surveys conducted by HUD included little information about health. This restriction limits our ability to determine whether the association between a move to a lower-poverty neighborhood and reductions in the prevalence of obesity and diabetes reflects a change in onset or persistence, but it does not affect the internal validity of our intention-to-treat estimates.

A further limitation of the study is the fact that the participants volunteered. More than 90% of the households in the study were headed by a black or Hispanic woman and included children. Among the 1.2 million households in public housing nationwide, 50% are nonwhite and 38% headed by women with children.[35] Our sample also had a higher prevalence of obesity than national samples of all U.S. families.

Although care should be taken in applying these results to populations with different attributes, our finding that neighborhood environments are associated with the prevalence of obesity and diabetes may have implications for understanding trends and disparities in overall health across the United States. The increase in U.S. residential segregation according to income in recent decades[36] suggests that a larger proportion of the population is being exposed to distressed neighborhood environments. Minorities are also more likely than whites to live in distressed areas.[37]

The results of this study, together with those of previous studies documenting the large social costs of obesity[38] and diabetes,[39] raise the possibility that clinical or public health interventions that ameliorate the effects of neighborhood environment on obesity and diabetes could generate substantial social benefits. The mechanisms accounting for these associations remain unclear, but further investigation is warranted to provide guidance in designing neighborhood-level interventions to improve health.

The views expressed in this article are those of the authors and should not be interpreted as those of the Congressional Budget Office, HUD, or any other federal agency or private foundation that provided support for the project.

Supported by grants from HUD (C-CHI-00808), the National Science Foundation (SES-0527615), the National Institute of Child Health and Human Development (NICHD) (R01-HD040404 and R01-HD040444), the Centers for Disease Control and Prevention (R49-CE000906), the National Institute of Mental Health (R01-MH077026), the National Institute on Aging (R56-AG031259 and P01-AG005842-22S1), the National Institutes of Health (to Dr. Lindau) through NORC (5P30 AG012857) and the University of Chicago Center on Demography and Economics of Aging Core on Biomeasures in Population Based Aging Research (1K23AG032870-01A1), and the Institute of Education Sciences at the Department of Education (R305U070006) and by the Population Research Center at the National Opinion Research Center (through a grant [R24-HD051152-04] from the NICHD), the Center for Health Administration Studies at the University of Chicago, the John D. and Catherine T. MacArthur Foundation, the Smith Richardson Foundation, the Spencer Foundation, the Annie E. Casey Foundation, the Bill and Melinda Gates Foundation, and the Russell Sage Foundation.

Dr. Kessler reports receiving fees for board membership from Eli Lilly, Mindsite, and Wyeth-Ayerst; receiving consulting fees from Wellness and Prevention, GlaxoSmithKline, Sanofi-Aventis, Kaiser Permanente, Merck, Ortho-McNeil Janssen Scientific Affairs, Pfizer, Shire US, SRA International, Takeda Global Research and Development, Transcept Pharmaceuticals, Wyeth-Ayerst, and Plus One Health Management; and holding stock in DataStat. Dr. Kessler's institution, Harvard Medical School, has received grant support from Analysis Group, Bristol-Myers Squibb, Eli Lilly, EPI-Q, Ortho-McNeil Janssen Scientific Affairs, Pfizer, Sanofi-Aventis, Shire US, and Walgreens. Drs. Lindau and Ludwig's institution, the University of Chicago, has received grant support from PepsiCo. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the members of the research team at the National Bureau of Economic Research, Joe Amick, Ryan Gillette, Ijun Lai, Jordan Marvakov, Matt Sciandra, Fanghua Yang, Sabrina Yusuf, and Michael Zabek, for assisting with the data preparation and analysis; Nancy Gebler (working under subcontract to our research team) of the Survey Research Center at the University of Michigan for leading the data-collection effort for the survey; and Todd Richardson and Mark Shroder of HUD and Kathleen Cagney, Elbert Huang, and Harold Pollack of the University of Chicago for their helpful comments on an earlier version of this article.

REFERENCES

1. Chang VW. Racial residential segregation and weight status among US adults. Soc Sci Med 2006;63:1289-303.
2. Black JL, Macinko J. The changing distribution and determinants of obesity in the neighborhoods of New York City, 2003-2007. Am J Epidemiol 2010;171:765-75.
3. Krishnan S, Cozier YC, Rosenberg L, Palmer JR. Socioeconomic status and incidence of type 2 diabetes: results from the Black Women's Health Study. Am J Epidemiol 2010;171:564-70.
4. Morenoff JD, Diez Roux AV, Hansen BB, Osypuk TL. Residential environments and obesity: what can we learn about policy interventions from observational studies? In: Schoeni RF, House JS, Kaplan GA, Pollack H, eds. Making Americans healthier: social and economic policy as health policy. New York: Russell Sage Foundation Press, 2008:309-43.
5. Office of the Surgeon General. The surgeon general's vision for a healthy and fit nation. Rockville, MD: Department of Health and Human Services, 2010.
6. Franco M, Diez Roux A, Glass TA, Caballero B, Brancati FL. Neighborhood characteristics and availability of healthy foods in Baltimore. Am J Prev Med 2008; 35:561-7.
7. Papas MA, Alberg AJ, Ewing R, Helzlsouer KJ, Gary TL, Klassen AC. The built environment and obesity. Epidemiol Rev 2007;29:129-43.
8. Lovasi GS, Neckerman KM, Quinn JW, Weiss CC, Rundle A. Effect of individual or neighborhood disadvantage on the association between neighborhood walkability and body mass index. Am J Public Health 2009;99:279-84.
9. Fowler-Brown AG, Bennett GG, Goodman MS, Wee CC, Corbie-Smith GM, James SA. Psychosocial stress and 13-year BMI change among blacks: the Pitt County study. Obesity (Silver Spring) 2009;17: 2106-9.
10. Cohen DA, Finch BK, Bower A, Sastry N. Collective efficacy and obesity: the potential influence of social factors on health. Soc Sci Med 2006;62:769-78.
11. Christakis NA, Fowler JH. The spread of obesity in a large social network over 32 years. N Engl J Med 2007;357:370-9.
12. Expanding housing choices for HUD-assisted families. Washington, DC: Department of Housing and Urban Development, April 1996. (http://www.huduser.org/portal/publications/affhsg/choices.html.)
13. Goering J, Feins JD, Richardson TM. What have we learned from housing mobility and poverty deconcentration? In: Goering J, Feins JD, eds. Choosing a better life? Evaluating the Moving to Opportunity social experiment. Washington, DC: Urban Institute Press, 2000:3-36.
14. Olsen EO. Housing programs for low-income households. In: Moffitt RA, ed. Means-tested transfer programs in the United States. Chicago: University of Chicago Press, 2003:365-442.
15. Bureau of the Census. Census tracts and block numbering areas. (http://www.census.gov/geo/www/cen_tract.html.)
16. Feins JD, Holin MJ, Phipps AA. Moving to Opportunity for fair housing demonstration: program operations manual (revised). Cambridge, MA: Abt, September 1996. (http://www.abtassociates.com/reports/D19960002.pdf.)
17. Feins JD, Holin MJ, Phipps AA, Magri D. Implementation assistance and evaluation for the Moving to Opportunity demonstration: final report. Cambridge, MA: Abt, April 1995. (http://www.abtassociates.com/reports/D19950045.pdf.)
18. Orr L, Feins JD, Jacob R, et al. Moving to Opportunity interim impacts evaluation: final report. Washington, DC: Department of Housing and Urban Development, Office of Policy Development and Research, June 2003. (http://www.abtassociates.com/reports/2003302754569_71451.pdf.)
19. Moving to Opportunity. Survey instruments. (http://www.nber.org/mtopublic/instruments.html.)
20. Groves RM, Fowler FJ Jr, Couper MP, Lepkowski JM, Singer E, Tourangeau R. Survey methodology. Hoboken, NJ: John Wiley, 2004.
21. University of Michigan, Institute for Social Research. Health and retirement study: physical measures and biomarkers. 2008. (http://hrsonline.isr.umich.edu/modules/meta/2008/core/qnaire/online/2008PhysicalMeasuresBiomarkers.pdf.)
22. Health o meter. Professional home care digital scales. (http://www.homscales.com/products/Professional%20Home%20Care%20Scales/1.aspx.)
23. Whatman. Protein saver cards. (http://www.whatman.com/903ProteinSaverCards.aspx#10548236.)
24. Saudek CD, Herman WH, Sacks DB, Bergenstal RM, Edelman D, Davidson MB. A new look at screening and diagnosing diabetes mellitus. J Clin Endocrinol Metab 2008;93:2447-53.
25. American Association for Public Opinion Research. Standard definitions. (http://www.aapor.org/Standard_Definitions/1481.htm.)
26. Obesity Education Initiative. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: the evidence report. Bethesda, MD: National Heart, Lung, and Blood Institute, September 1998. (NIH publication no. 98-4083.) (http://www.nhlbi.nih.gov/guidelines/obesity/ob_gdlns.pdf.)
27. International Expert Committee. International Expert Committee report on the role of the A1C assay in the diagnosis of diabetes. Diabetes Care 2009;32:1327-34.
28. American Diabetes Association. Standards of medical care in diabetes — 2010. Diabetes Care 2010;33:Suppl 1:S11-S61.
29. Sampson RJ, Raudenbush SW, Earls F. Neighborhoods and violent crime: a multilevel study of collective efficacy. Science 1997;277:918-24.
30. Jacob BA, Ludwig J. The effects of housing assistance on labor supply: evidence from a voucher lottery. Am Econ Rev (in press).
31. Bloom HS. Accounting for no-shows in experimental evaluation designs. Eval Rev 1984;8:225-46.
32. Kling JR, Liebman JB, Katz LF. Experimental analysis of neighborhood effects. Econometrica 2007;75:83-119.
33. Stata 11.0 special edition. College Station, TX: Stata, 2009. (http://www.stata.com.)
34. Ludwig J, Liebman JB, Kling JR, et al. What can we learn about neighborhood effects from the Moving to Opportunity experiment? Am J Sociol 2008;114:144-88.
35. U.S. House Ways and Means Committee. Green book, 2008. (http://democrats.waysandmeans.house.gov/singlepages.aspx?NewsID=10490.)
36. Watson T. Inequality and the measurement of residential segregation by income in American neighborhoods. Rev Income Wealth 2009;55:820-44.
37. Jargowsky PA. Stunning progress, hidden problems: the dramatic decline of concentrated poverty in the 1990s. Washington, DC: Brookings Institution, May 2003. (http://www.brookings.edu/~/media/Files/rc/reports/2003/05demographics_jargowsky/jargowskypoverty.pdf.)
38. Finkelstein EA, Trogdon JG, Brown DS, Allaire BT, Dellea PS, Kamal-Bahl SJ. The lifetime medical cost burden of overweight and obesity: implications for obesity prevention. Obesity (Silver Spring) 2008;16:1843-8.
39. Trogdon JG, Hylands T. Nationally representative medical costs of diabetes by time since diagnosis. Diabetes Care 2008;31:2307-11.