

Effectiveness of Policies and Programs to Combat Adult Obesity: a Systematic Review

Eva Tseng, MD MPH^{1,2}, Allen Zhang, BS^{3,4}, Oluwaseun Shogbesan, MD⁵,
Kimberly A. Gudzone, MD MPH^{1,2}, Renee F. Wilson, MS^{3,4}, Hadi Kharrazi, MHI, MD, PhD³,
Lawrence J. Cheskin, MD⁶, Eric B. Bass, MD, MPH^{1,3,4}, and Wendy L. Bennett, MD, MPH^{1,2}

¹Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA; ²Welch Center for Prevention, Epidemiology, & Clinical Research, The Johns Hopkins University, Baltimore, MD, USA; ³Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; ⁴Johns Hopkins University Evidence-based Practice Center, Johns Hopkins Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, MD, USA; ⁵Department of Medicine, Reading Hospital, West Reading, PA, USA; ⁶Department of Health, Behavior & Society, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA.

BACKGROUND: This systematic review identifies programs, policies, and built-environment changes targeting prevention and control of adult obesity and evaluates their effectiveness.

METHODS: We searched PubMed, CINAHL, PsycINFO, and EconLit from January 2000 to March 2018. We included natural experiment studies evaluating a program, policy, or built-environment change targeting adult obesity and reporting weight/body mass index (BMI). Studies were categorized by primary intervention target: physical activity/built environment, food/beverage, messaging, or multiple. Two reviewers independently assessed the risk of bias for each study using the *Effective Public Health Practice Project* tool.

RESULTS: Of 158 natural experiments targeting obesity, 17 reported adult weight/BMI outcomes. Four of 9 studies reporting on physical activity/built environment demonstrated reduced weight/BMI, although effect sizes were small with low strength of evidence and high risk of bias. None of the 5 studies targeting the food/beverage environment decreased weight/BMI; strength of evidence was low, and 2 studies were rated high risk of bias.

DISCUSSION: We identified few natural experiments reporting on the effectiveness of programs, policies, and built-environment changes on adult obesity. Overall, we found no evidence that policies intending to promote physical activity and healthy eating had beneficial effects on weight/BMI and most studies had a high risk of bias. Limitations include few studies met our inclusion criteria; excluded studies in children and those not reporting on weight/BMI outcomes; weight/BMI reporting was very heterogeneous. More high-quality research, including natural experiments studies, is critical for informing the population-level effectiveness of obesity prevention and control initiatives in adults.

KEY WORDS: obesity; prevention; systematic reviews.

J Gen Intern Med 33(11):1990–2001

DOI: 10.1007/s11606-018-4619-z

© Society of General Internal Medicine 2018

Received May 17, 2018

Revised June 30, 2018

Accepted July 31, 2018

Published online September 11, 2018

INTRODUCTION

Forty percent of adults across the world are overweight or obese.¹ Obesity-related conditions, including heart disease, type-2 diabetes, and many cancers, contribute to substantial morbidity and mortality.² The standard approach to obesity management in primary care has been mainly individual-focused behavioral lifestyle programs, such as referrals to nutrition or commercial programs like Weight Watchers.³ Our lack of progress in slowing the rise of obesity is due in part to obesity having multiple complex causes, including the obesogenic environment and the food system.⁴ Population-level approaches to obesity, which can be evaluated in “natural experiment” studies, are necessary for addressing these complex causes.

Natural experiments are events, interventions, or policies that are not controlled by researchers, but can be studied to determine their impact.⁵ Examples include assessing the effects of needle exchange programs on HIV prevention⁶ and the impact of indoor smoking bans on asthma exacerbations.⁷ Because randomized, controlled trials are not always practical or feasible, there is growing interest in using natural experiment studies to understand the effect of population-level policies and programs on chronic diseases like obesity. Examples of obesity policies include the New York City law requiring chain restaurants to post calorie information on menus,⁸ the Danish fat tax (2011–2013) which taxed foods exceeding a certain saturated fat content,^{9, 10} and the ban in schools of sugar-sweetened beverages in several states.¹¹ Examples of built-environment changes amenable to natural experiment studies include the opening of new supermarkets in areas with lower access to healthy foods^{12–14} and installation of new light rail lines to improve public transport infrastructure.^{15–17}

The purpose of this systematic review is to identify programs, policies, and built-environment changes targeting adult obesity that were evaluated in natural experiment studies to evaluate their effectiveness in reducing obesity and related outcomes.

METHODS

This analysis was part of a larger systematic review entitled, “Methods for Evaluating Natural Experiments in Obesity: Systematic Evidence Review”^{18, 19} and expands upon one of the key questions addressed in the report, regarding the 17 natural experiment studies which studied adult obesity outcomes. For a description of detailed methods, see the full report.¹⁸

Data Sources and Searches

We searched PubMed, CINAHL, PsycINFO, and EconLit (1 January 2000 to 28 March 2018) to identify all English-language US and non-US studies of programs, policies, and built-environment changes targeting obesity prevention and control in people of all ages and in any setting (Appendix). We also conducted a gray literature search of websites of key organizations involved in obesity prevention and evaluations of policies to identify white papers and unpublished evaluations. We also searched the reference lists of included articles for missed or unpublished reports.¹⁸

Study Selection

Two reviewers independently screened abstracts and articles to identify natural experiment studies evaluating a program or policy aimed at combating adult obesity that reported body weight/BMI outcomes.¹⁸ Studies were defined as natural experiments based on the UK’s Medical Research Council guidelines, in which exposure to the event or intervention of interest has not been manipulated by a researcher.⁵ Disagreements between reviewers were discussed and, if they could not be resolved, were adjudicated by a third-party reviewer.

Data Extraction and Quality Assessment

Two reviewers serially extracted data on study, population and intervention characteristics, magnitude of effect size and *p* value for weight (BMI or weight change in adults), and physical activity and dietary outcomes, including intake of total daily calories, fruits, and vegetables, sugar-sweetened beverages, fiber, and fast food.

Two reviewers independently assessed the risk of bias for each study using the *Effective Public Health Practice Project* (EPHPP) tool to rate studies on 6 domains.¹⁸ Each study received a global risk of bias rating: low risk of bias if no domains were rated as weak, moderate risk of bias if one domain was rated as weak, or high risk of bias if two or more domains were rated as weak.

Data Synthesis and Analysis

For each study, we categorized the target of the program, policy, or built-environment change based on a classification in the 2012 National Academy of Medicine report, “*Accelerating Progress in Obesity Prevention: Solving the*

Weight of the Nation”²⁰: (1) physical activity and physical and built environments (e.g., building a new light rail line), (2) food and beverage environment (e.g., opening of supermarkets), (3) messaging environment (e.g., calorie counts on menus). Studies with multiple primary intervention targets were categorized as (4) “multiple targets.” We described the overall effect of these programs, policies, or built-environment changes on each outcome of interest (weight, diet, and physical activity behaviors), compared the magnitude of effect sizes for mean BMI change and between-group difference in BMI using a forest plot, and summarized dietary and physical activity behaviors using an evidence map. We were unable to perform a meta-analysis due to study heterogeneity and insufficient number of articles with both the same outcome and intervention. We classified outcomes as *favorable* if a statistically significant difference was reported in the hypothesized direction, *unfavorable* if a statistically significant difference was reported in the opposite direction, and *no difference* if no statistically significant difference was reported.

Strength of Evidence for Weight and BMI Outcomes

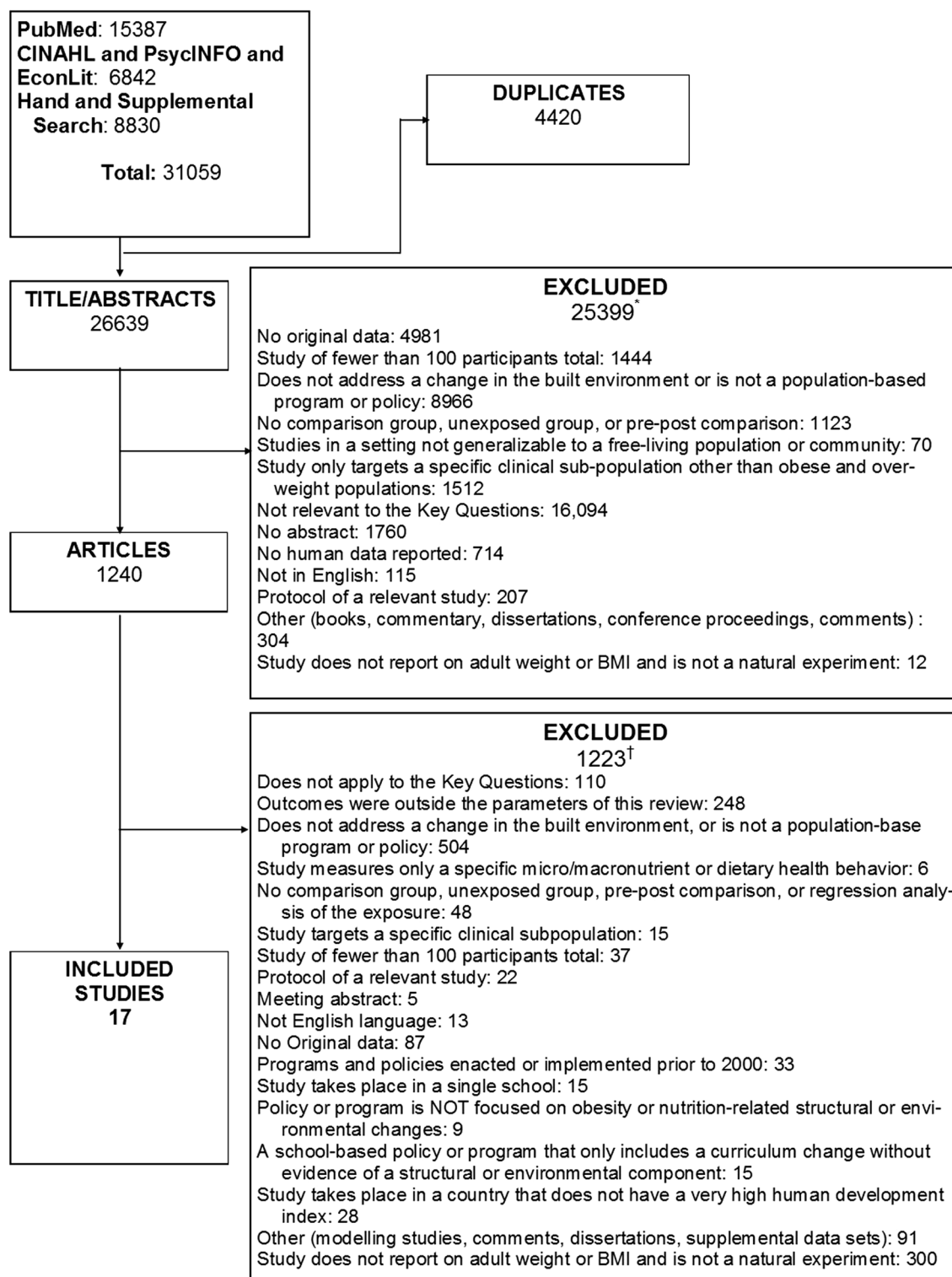
We used the Agency for Healthcare Research and Quality (AHRQ) guide’s strength of evidence grading schema²¹. Two reviewers independently graded the evidence based on the intervention target assessing studies’ limitations, consistency, directness, precision, and potential reporting bias for the evidence on weight/BMI outcomes: “High” strength of evidence indicated that the evidence likely reflected the true effect; “moderate” strength indicated that further research could change the result; “low” strength indicated low confidence that the evidence reflects the true effect; “insufficient” indicated no confidence in the estimate of effect.

Role of the Funding Source

The NIH Office of Disease Prevention funded the larger systematic review through an interagency agreement with AHRQ. A working group convened by the NIH assisted in developing the scope of the review and its key questions. Neither organization had a role in study selection, quality assessment, or synthesis. The investigators are solely responsible for the content. The original review is registered with PROSPERO (#CRD42017055750).

RESULTS

Of the 158 natural experiment studies in the larger systematic review,^{18, 19} 17 reported weight/BMI outcomes in adults (Fig. 1 PRISMA diagram). Table 1 displays the individual study and population characteristics, main results for weight/BMI outcomes, and risk of bias. Study



* Sum of excluded abstracts exceeds 25399 because reviewers were not required to agree on reasons for abstract exclusion.

† Sum of excluded articles exceeds 1223 because reviewers were not required to agree on reasons for exclusion

Figure 1 Evidence search and selection.

durations varied from 1 to 20 years. The mean age of the study population ranged from 38 to 80+ years. One study was an outlier for study duration and mean age because it followed children to adulthood for 20 years.²⁴ The mean

baseline BMI was 17–30 kg/m². Only 7 studies reported on race/ethnicity. Most natural experiment studies had a high risk of bias ($n = 9$) (Table 3), particularly in terms of handling withdrawals and dropouts, and study design.

Table 1 Study and Population Characteristics, Main Results for Weight/BMI Outcomes, and Overall Rating for Quality of Study for Included Natural Experiment Studies ($n = 17$), by Intervention Target

Author, year, US or non-US	Study characteristics		Range of baseline population characteristics within studies						Main result for weight/BMI outcomes		Overall risk of bias
	Program, policy, or built-environment change	Comparator group	Study length	N	Mean or range of mean age, years	Female, %	White, % Black, %	Mean or range of mean BMI, kg/m ²			
Physical activity/physical and built environments, <i>n</i> = 8 Bekemeier 2018 US	Patterns of PA interventions	No activities vs. PA interventions (built environment, personal health, comprehensive, school-based)	NR	NR	NR	NR	White: NR Black: 5–11	NR	Reduced obesity • Built environment, <i>p</i> ≤ 0.05 • Personal health, NR • Comprehensive, <i>p</i> ≤ 0.01 • School-based, <i>p</i> ≤ 0.05 Reduced BMI • Cross-sectional comparison of mean BMI	Moderate	
Brown 2008 US	New urbanist neighborhoods in North Carolina	Conventional suburban vs. new urbanist neighborhoods	NA	453	37.2–46.8	NR	NR	18.5–24.9	• Lower BMI in new urbanist multi-family units vs. conventional neighborhoods (<i>p</i> = 0.03) • No diff in BMI between urbanist single-family units vs. conventional neighborhoods (<i>p</i> = 0.43) Reduced BMI	High	
Brown 2015 US	Building of light rail extension in Salt Lake City, Utah	Pre/post for never riders vs. new riders	12	537	37.8–43.9	51	NR	29	Reduced BMI for new riders (mean change −0.29 kg/m ² , <i>p</i> < 0.01) • Increased BMI for non-riders (mean change 0.19 kg/m ² , NR) No difference in BMI (NR) • Cross-sectional comparison of mean BMI	High	
Camacho-Rivera 2017 US	Housing Choice Voucher Program (HCVP) in New York City	Voucher program vs. unassisted vs. public housing residents	NA	362	46.4	74.5	NR	NR	• Public housing, 32.03 kg/m ² • HCVP, 30.28 kg/m ² • Unassisted, 30.04 kg/m ² Women: increased BMI • Between-group diff, 0.269 kg/m ² , <i>p</i> < 0.01 Men: no difference in BMI • Between-group diff, −0.001 kg/m ² , <i>p</i> value NR	High	
Chen 2015 Non-US	China's Home Appliances Going to Countryside policy • Subsidy for purchase of select home appliances	Pre/post for women and men	36	18,769: Women, 8112 Men, 7032	45.4–47.5	NR	NR	23.2–23.6	No difference in BMI (NR) • Cross-sectional comparison of mean BMI	Moderate	
Cleland 2008 Non-US	Categories of compulsory PA in schools in Australia	Low vs. medium vs. high compulsory PA schools	240	5170	7–15	NR	NR	17–20	No difference in BMI (NR) • No diff in mean BMI after 20 years among adults who attended different compulsory PA schools	High	

(continued on next page)

Table 1. (continued)

Author, year, US or non-US	Study characteristics	Range of baseline population characteristics within studies					Main result for weight/ BMI outcomes	Overall risk of bias
		Program, policy, or built-environment change	Comparator group	Study length	N	Mean or range of mean age, years	White, % Black, %	Mean or range of mean BMI, kg/m ²
MacDonald 2010 US	New light rail transit system in Charlotte, NC	Users vs. non-users	24	801	52	57.7	White: 71.3 Black: 21.2	27
Webb 2012 Non-US	UK policy of free bus passes to older adults • 2004–2008	Users vs. non-users of public transport	48	8773	NR	55.3	NR	27.9
Webb 2016 Non-US	UK policy of free bus passes to older adults • 2012–2013	Bus pass holder vs. non-bus pass holder	NA	4650	62–80+	89.1	NR	28
Food and beverage environment, <i>n</i> = 4	Opening of supermarket in Philadelphia	Intervention vs. control neighborhood	48	1440	54	79	White: 9 Black: 84	NR
Dubowitz 2015 US	Opening of supermarket in Pittsburgh	Intervention vs. control neighborhood	36	1372	53.3	75	White: NR Black: 95.2	30.5
Rigdon 2017 US	Supplemental Nutrition Assistance Program (SNAP) • Food-purchasing assistance to low- and no-income people	SNAP participation vs. no participation	12	10,360	42	53.9	White: NR Black: 15.1	NR
Sadler 2013 US	Opening of independent grocery store in Flint, MI	Pre/post in control vs. intervention towns	24	186	55	55	White: NR Black: 61	29.5
Sturm 2015 US	Los Angeles (LA) Fast Food ban • 2008 zoning regulation restricting opening/expanding of standalone fast-food restaurants	South LA vs. control (remainder city of LA or LA County)	48	12,058	NR	NR	NR	26.4–27.9
Messaging environment, <i>n</i> = 1	Calorie Labeling Law in New York • Requires chain restaurants post calorie counts on menus	Counties that implemented law vs. no implementation	96	103,220	53.2	59	White: NR Black: 12	27.1

(continued on next page)

Reduced BMI

- Between-group diff. –1.18 kg/m², 95%CI -2.22 to -0.13, *p* = 0.015
- Reduced odds of obesity
- OR 0.79, 95%CI 0.63 to 0.98, *p* = 0.033

Women: reduced BMI (*p* = 0.001)

Men: reduced BMI (*p* = 0.305)

- Cross-sectional comparison of mean BMI

Reduced BMI

- Between-group diff. –0.46 kg/m², *p* = 0.56
- Reduced BMI
- Between-group diff. –0.31 kg/m², NR

No difference in BMI (NR)

Reduced BMI (NR) in intervention town

Increased BMI (NR) in control town

Increased BMI in South LA more than comparison areas (NR)

Reduced BMI in counties implementing the policy compared to counties that did not

- Decrease in BMI of –0.4 kg/m², *p* = 0.01

Table 1. (continued)

Author, year, US or non-US	Study characteristics	Range of baseline population characteristics within studies						Main result for weight/ BMI outcomes	Overall risk of bias	
		Program, policy, or built-environment change	Comparator group	Study length	N	Mean or range of mean age, years	Female, %			White, % Black %
Multiple environments*, n = 2										
Bolton 2017 Non-US	Health-Promoting Communities: Being Active Eating Well • Community and workplace programs for healthy eating & PA in areas of socioeconomic disadvantage in Australia	Control vs. intervention community	24	3648	NR	NR	NR	NR	No difference in BMI (NR) at follow up • Control, 25.6 kg/m ² • Intervention, 26.0 kg/m ²	High
Hu 2016 Non-US	UK Reducing Health Inequalities program • Broad front including raising living standards, family support policies, tax reduction, improved early education	UK vs. Finland vs. Italy	240	260,054	NR	52	NR	NR	UK: no difference in improvement of inequalities of obesity • OR 1.25, NR Finland: unfavorable trend • OR 1.90, <i>p</i> < 0.1 Italy: favorable trend • OR 0.76, <i>p</i> < 0.1	Moderate

BMI, body mass index; CI, confidence interval; HCVP, housing choice voucher program; NA, not applicable; NR, not reported; PA, physical activity; SNAP, supplemental nutrition assistance program; UK, United Kingdom
 *²² included the physical activity and built environment, school environment, and food and beverage environment²³; included the food and beverage environment, physical activity environment, school and work environments, and messaging environment

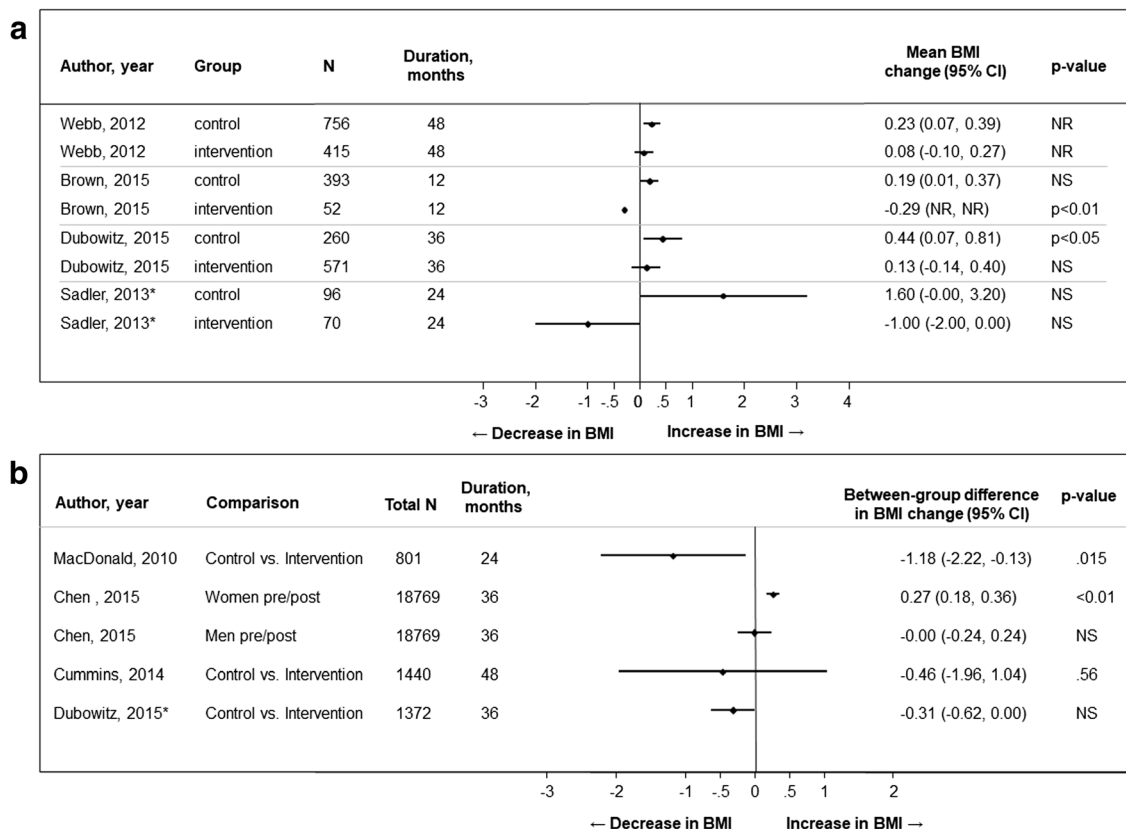
Nine studies targeted the physical activity and built environment, including the building of new light rail systems or extensions,^{15, 16} free bus pass eligibility,^{25, 26} and participation in the *Housing Choice Voucher Program* in New York City.²⁷ Five studies targeted food and beverage environments, including the building of grocery stores^{12, 13, 28} and the Los Angeles fast food ban, which restricted opening/expanding stand-alone fast-food restaurants.²⁹ One study targeted an obesity-related messaging environment, the calorie labeling law in New York requiring chain restaurants to post calorie counts on menus.³⁰ Two studies targeted multiple environments (Table 1), one focusing on healthy eating and physical activity programs in the workplace²³ and the other examining the effect of the English national strategy to reduce health inequalities with 3 comparison countries.²²

Effectiveness of Policies, Programs, and Built-Environment Changes on Obesity, Weight, or BMI

Table 1 shows weight/BMI outcomes. Figure 2a displays mean pre-post BMI change within each group for 4 studies, and Fig. 2b displays the between-group difference in BMI in 4 other studies. All 17 studies reported on BMI. Strength of evidence for the outcome of weight/BMI was low due to an overall high risk of bias (4 of 9 studies) and inconsistency in

the direction of effect (Tables 2 and 3). Among the 9 studies focused on the physical activity/built environment, 4 showed weight/BMI reduction, 3 had inconsistent findings by subgroups (i.e., some subgroups showed favorable outcomes and others did not), and 2 showed no difference. Three of the 4 studies that showed weight/BMI reduction focused on transit use.^{15, 16, 25} For example, users of a new light rail transit system in Charlotte, NC, had a reduced BMI compared to non-users, though this study had a high risk of bias.¹⁶ The fourth study examined physical activity intervention clusters (e.g., community health education campaigns, individual health behavior change) finding a reduction in obesity in 3 of the 4 clusters.³¹ Three of 9 studies showed inconsistent results by subgroup^{26, 32, 33} (Table 1). Two studies targeting the physical activity and built environment, one examining different categories of compulsory school physical activity in Australia and the other evaluating the *Housing Choice Voucher Program* in New York City, showed no changes in weight/BMI, although both studies had a high risk of bias.^{24, 27}

None of the 5 studies targeting the food and beverage environment showed a weight/BMI reduction (i.e., one study increased weight/BMI and 4 showed no difference) (Table 1, Fig. 2). Strength of evidence was low due to high risk of bias (3 of 5 studies) and inconsistency in the direction of effect (Tables 2 and 3). Rigdon et al. showed that participation in the



Abbreviations: BMI, Body Mass Index; CI, Confidence Interval; NS, non-significant; NR, not reported

*P-value reported as "non-significant" so confidence interval calculated using p=0.05 (42)

Figure 2 Forest plot displaying mean pre/post BMI change within each group (a) and between groups (b).

Table 2 Summary of the Strength of Evidence for BMI/Weight Outcomes

Intervention target	No. studies	Study limitations	Directness	Consistency	Precision	Reporting Bias	Strength of evidence	Summary of key outcomes
Physical activity and built environment	9	High	Direct	Inconsistent	Unknown	Undetected	Low	Low strength of evidence that these programs or policies have any impact on BMI/weight outcomes
Food and beverage	5	High	Direct	Inconsistent	Unknown	Undetected	Low	Low strength of evidence that these programs or policies have any impact on BMI/weight outcomes
Messaging	1	High	Indirect	Unknown	Unknown	Undetected	Insufficient	Insufficient strength of evidence that these programs or policies have any impact on BMI/weight outcomes
Multiple	2	High	Indirect	Unknown	Unknown	Undetected	Insufficient	Insufficient strength of evidence that these programs or policies have any impact on BMI/weight outcomes

federal Supplemental Nutrition Assistance Program had no effect on BMI vs. non-participation after 1 year, but also had a high risk of bias.³⁴

One study by Restrepo et al. focused on the messaging environment.³⁰ Calorie labeling was associated with a BMI reduction in 11 counties in New York that implemented a law. However, strength of evidence was insufficient because only one study was included (Table 2).³⁰

Among the 2 studies targeting multiple interventions, one study by Bolton et al. compared changes in BMI in Australia after implementing community and workplace programs promoting healthy eating and physical activity²³ (Table 1). The study showed no difference in BMI between the control and intervention communities after

2 years and was rated as having a high risk of bias due to lack of information on blinding and withdrawals and dropout rates²³ (Table 3). The second study by Hu et al. examined whether changes in trends in health inequalities in England after implementation of its broad national program (e.g., family support policies, tax-reduction) were more favorable vs. other countries without such a program, and found inconsistent direction of effects (England: no difference; Finland: unfavorable trend; Italy: favorable trend).²² The strength of evidence was rated insufficient due to a high risk of bias (1 of 2 studies), indirectness of evidence (i.e., self-reported height and weight data), and inconsistency in the direction of effect (Tables 2 and 3).

Table 3 Risk of Bias for Each Study Assessed Using the EPHPP Tool

Author, year	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and drop outs	Overall rating
Physical activity/physical and built environment, <i>n</i> = 9							
Bekemeier 2018	M	M	H	M	M	M	Moderate
Brown 2008	H	H	H	M	M	M	High
Brown 2015	L	M	L	H	L	H	High
Camacho-Rivera 2017	H	H	L	M	L	H	High
Chen 2015	L	M	L	M	L	H	Moderate
Cleland 2008	M	H	M	M	L	H	High
MacDonald 2010	M	M	L	M	L	L	High
Webb 2012	L	H	L	M	M	M	Moderate
Webb 2016	M	H	L	M	M	M	Moderate
Food and beverage environment, <i>n</i> = 5							
Cummins 2014	M	M	L	M	M	M	Low
Dubowitz 2015	L	M	L	M	M	M	Low
Rigdon 2017	L	H	L	M	L	H	High
Sadler 2013	H	M	L	H	H	H	High
Sturm 2015	M	H	L	M	H	M	High
Messaging environment, <i>n</i> = 1							
Restrepo 2016	M	H	M	H	H	H	High
Multiple environments, <i>n</i> = 2							
Bolton 2017	M	M	L	H	L	H	High
Hu 2016	M	M	L	M	H	M	Moderate

EPHPP Effective Public Health Practice Project; H high risk of bias; L low risk of bias; M medium risk of bias

*Studies were given a “low risk of bias if none of the domains were rated as weak, moderate risk of bias if only one of the domains was rated as weak, or high risk of bias if two or more domains were rated as weak

Effectiveness of Policies, Programs, and Built-Environment Changes on Dietary Behaviors




















Figure 3 displays a summary of results for dietary outcomes, stratified by intervention target. Each outcome symbol represents an individual study, and the main result for the outcome. Reported dietary behaviors include intake of fruit and vegetables ($n = 6$), sugar-sweetened beverages ($n = 1$), total daily calories ($n = 2$), and fast food ($n = 1$). Studies reporting on fruit and vegetable intake generally found a small (0.1–0.3 servings/day) increase, or no difference. One study by Dubowitz et al., rated as low risk of bias, investigated the impact of a new supermarket in Pittsburgh, PA, and showed no difference in fruit/vegetable intake between the supermarket and control neighborhoods (between-group difference, -0.14 servings/day, p value not reported), but demonstrated a decrease in caloric intake (between-group difference, -178 kcal/day, $p < 0.005$).¹² The largest included study, by Restrepo et al., evaluated the calorie labeling law in New York and found no difference in fruit and vegetable intake comparing counties that implemented the law vs. did not.²⁷ Only one study reported on sugar-sweetened beverage intake and found no difference.²³






Two studies reported on caloric intake.^{12, 32} One, rated as moderate risk of bias, showed an increase in caloric intake for women and men after implementation of a Chinese policy providing a subsidy for select home appliances in rural communities.³²

Only one study reported on fast food intake and demonstrated no difference among residents living in public housing using a Voucher Program vs. unassisted housing units, but had a high risk of bias.²⁷

Effectiveness of Policies, Programs, and Built Environment Changes on Physical Activity Behaviors

Eight of the 17 studies reported on physical activity behaviors. Figure 3 displays a summary of physical activity outcomes, stratified by intervention target. Of the 9 studies targeting the physical activity and built environment, 2 showed an increase in physical activity,^{15, 26} 2 showed no difference,^{16, 24} and one large study found that physical leisure activity declined among women and did not change among men.³² One study, rated as high risk of bias, showed that after the building of a light rail extension in Salt Lake City, riders increased physical activity

Intervention Target	Does Not Favor Improved Outcomes	No Difference	Favors Improved Outcomes
Physical Activity & Built Environment ($n=9$)†		   	 
Food and Beverage ($n=5$)		   	
Messaging ($n=1$)		 	
Multiple ($n=2$)§		   	

Types of Outcomes:	 Sugar sweetened beverage intake	 Fast Food intake	 Physical activity
	 Fruit/vegetable intake	 Caloric intake	

*Some studies reported outcomes by subgroups and therefore are represented by more than one symbol (described below). We classified outcomes as favors or does not favor improved outcomes if a statistically significant difference was reported and no difference if no statistically significant difference was reported.

† 1 study reported inconsistent results for physical activity by gender: Women-unfavorable, Men- no effect (Chen, 2015)

‡ 1 study reported inconsistent results for fruit/vegetable intake: no effect on fruit intake, favorable effect on vegetable intake (Bolton, 2017)

Figure 3 Summary of results for secondary outcomes of dietary and physical activity behaviors, by intervention target. Each outcome symbol represents an individual study and their main result (favors, does not favor or no difference) for the particular outcome. Studies reporting inconsistent results are described in footnotes.*

compared with non-riders.¹⁵ Another study, rated as moderate risk of bias, found that bus-pass holders in England engaged in greater moderate or vigorous physical activity compared with non-bus pass holders (OR, 1.43, 95% CI, 1.12–1.84).²⁶ A large single study on the calorie labeling law in New York found no difference in exercise comparing counties that implemented the law with those that did not.³⁰ A study focused on multiple environments found that community programs aimed at healthy eating and physical activity had no effect on physical activity compared to control communities.²³

DISCUSSION

Because it is not always feasible to conduct controlled trials focused on adult obesity prevention and control, natural experiment studies offer a valuable opportunity to evaluate population-level programs, policies and built-environment changes. Of the 17 studies we identified, most evaluated programs or policies promoting physical activity or evaluating built-environment changes, followed by those which focused on efforts to improve the food and beverage environment. Overall, we found no evidence that policies promoting physical activity and healthy eating had beneficial effects on the weight of adults, and most studies performed had a high risk of bias. Few studies showed a reduction in weight or BMI, and those that did generally had small (≤ 0.5 kg/m² BMI) effect sizes, which would not be considered a clinically important difference. While the studies evaluating population-based policies and programs were generally of short duration, when there was no difference in weight/BMI reported, we did not identify early effects on diet or physical activity that would suggest that a longer follow-up would yield more robust changes in weight. In fact, overall, policies and programs targeting obesity resulted in small changes in dietary intake and physical activity, if any. Our results demonstrate a need for more and better evidence from natural experiment studies conducted with rigorous research methods.

Despite disappointing findings to date, novel approaches like natural experiment studies are needed to effectively assess the impact of population-level health policies and programs on obesity. The NIH has made it a priority to evaluate existing programs and policies via natural experiment studies. Recently funded NIH grants have included an evaluation of the impact on physical activity of a new light rail line in Houston,³⁵ the impact on consumer behavior of a New York City sugar-sweetened beverage policy limiting the sale of large drinks,³⁶ and the impact of a New York City park redesign and renovation initiative on physical activity.³⁷ These promising studies have added to the small but growing body of natural experiment studies examining adult obesity prevention and control through population-based solutions. A research database of effective programs, policies, and built-environment changes in obesity prevention and control does not currently exist, but would provide invaluable information to researchers, policy makers and funders.

We identified some limitations in natural experiment studies, including heterogeneity in the reporting of effect sizes. While all studies reported on weight outcomes using BMI, this data was analyzed and reported in different ways (e.g., odds of obesity, change in BMI). Physical activity was measured by different methods (e.g., accelerometers, questionnaires) using various measurement units (e.g., counts/minute, MET-hours/week) making it difficult to compare effect sizes between studies. Clinical trial data from intensive individual behavioral weight loss interventions often produce between-group differences of 5 kg.³ However, the expected effect size for the types of population-level programs, policies, or built-environment changes we reviewed is not clear, especially since these studies' primary focus is on preventing weight gain, not weight loss. Also, it is unclear what an adequate sample size for non-experimental studies is and such an assessment must account for numerous factors like clustering, different time points, and correlation of outcomes over time.

Another limitation was that most studies were 1–2 years in length, and longer studies may be needed to fully capture the impact on slow-to-appear obesity-related outcomes like population-level weights. Finally, most studies had a high risk of bias, due to high rates of losses to follow-up and lower rigor of study design (e.g., one-time surveys or simple pre-post designs) and high risk of confounding.

We searched PubMed for other relevant English-language systematic reviews within the past 10 years reporting on obesity-related outcomes in natural experiment studies. In one systematic review, Mayne et al. identified 37 studies that were natural- or quasi-experiments.³⁸ Only 3 of these studies reported on weight/BMI, and only 1 of the 3 showed a reduction in self-reported BMI (with use of a new light rail system).^{16, 38} In contrast, our review required all studies to report on BMI/weight outcomes, and we found a favorable reduction in less than half of the studies. In addition, Mayne et al. found that studies reporting positive impacts on physical activity or diet tended to have longer follow-up times than negative studies.³⁸ Our results for diet and physical activity outcomes were not consistent, based on sample size or follow-up time. Another systematic review examined policy interventions aimed at improving population nutrition,³⁹ and the included studies reported on outcomes not included in our review (e.g., change in calorie value of purchases, consumer knowledge).³⁹ A third systematic review examined the associations between physical activity and the built environment and assessed outcomes not included in our review: street and pedestrian connectivity, neighborhood parks, and green space.⁴⁰

Our review itself has several limitations. First, we excluded studies not reporting on adult BMI/weight outcomes, so we may have excluded studies showing a favorable effect on intermediate outcomes like dietary or physical activity behaviors which did not report weight or BMI. Second, the number of included studies was small, and BMI and weight data were reported heterogeneously across studies, making it difficult to

summarize the data quantitatively. Third, publication bias may have led to underreporting of policies, programs, or built-environment changes that showed no difference. Fourth, we described outcome results based on statistical significance, which may not constitute clinically-significant effects. As mentioned earlier, it is not clear what are clinically significant effect sizes for the types of population-level programs, policies, and built-environment changes we evaluated. Fifth, we used the AHRQ guide to evaluate the strength of evidence for weight/BMI outcomes and the EPHPP tool to assess risk of bias, but recognize that there is no standard method for assessing strength of evidence and risk of bias in natural experiment studies. A strength of natural experiment studies is enhancing generalizability and external validity, but the EPHPP tool was designed to evaluate internal validity, not external validity.¹⁸

In conclusion, we identified few natural experiment studies reporting on the effectiveness of policies/programs/built-environment changes on adult obesity, and overall, these studies demonstrated inconsistent effects on weight/BMI, with low or insufficient strength of evidence and high risk of bias.

In primary care, obesity is commonly managed through individual behavioral interventions, but population-level approaches can also play an important role. While natural experiment studies allow us to evaluate these approaches, this type of study has clear limitations, including the lack of rigorous research methods used and standards for data reporting. More high-quality research, including natural experiment studies, is needed to inform the population-level effectiveness of obesity prevention and control initiatives. Better understanding of these effects has the potential to augment the efforts of primary care providers in combatting the continuing global obesity epidemic.

Acknowledgments:

The authors thank Dr. Elizabeth Stuart and Ms. Emily Knapp for their input. They also thank Dr. Lionel Bañez, their task order officer at AHRQ.

Prior Presentation: This work was presented as a poster at the Society for General Internal Medicine Meeting, April 12, 2018.

Corresponding Author: Eva Tseng, MD MPH; Division of General Internal Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA (e-mail: etseng3@jhmi.edu).

Funding Information This report is based on research conducted by the Johns Hopkins University Evidence-based Practice Center under contract 290-2012-000071 to AHRQ, Rockville, Maryland.

Compliance with Ethical Standards:

Conflict of Interest: Dr. Lawrence Cheskin served on the scientific advisory board of Medifast during the conduct of the study. All other authors declare no conflicts of interest.

Disclaimer: The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore,

no statement in this report should be construed as an official position of AHRQ or of the US Department of Health and Human Services.

Reproducible Research Statement: Study protocol: Available at www.crd.york.ac.uk/PROSPERO/display_record.php?ID=CRD42017055750. Statistical code: Not applicable. Data set: Full technical report available at <https://www.effectivehealthcare.ahrq.gov/topics/obesity-research-methods/systematic-review>

Funding Source: National Institutes of Health Office of Disease Prevention and the Agency for Healthcare Research and Quality, PROSPERO registry #CRD42017055750.

REFERENCES

- World Health Organization. Global Status Report on Noncommunicable Diseases 2014. Geneva: World Health Organization, 2014.
- Kopelman P, Jebb S, Butland B. Executive summary: Foresight 'Tackling Obesity: Future Choices' project. *Obes Rev*. 2007;8(Suppl 1):vi-ix.
- Jensen MD, Ryan DH, Apovian CM, et al. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. *J Am Coll Cardiol*. 2014;63(25 Pt B):2985-3023.
- Swinburn BA, Sacks G, Hall et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet*. 2011;378(9793):804-14.
- Craig P, Cooper C, Gunnell D, et al. Using natural experiments to evaluate population health interventions: new Medical Research Council guidance. *J Epidemiol Community Health*. 2012;66(12):1182-6.
- Wood E, Lloyd-Smith E, Li K, et al. Frequent needle exchange use and HIV incidence in Vancouver, Canada. *Am J Med*. 2007;120(2):172-9.
- Ciaccio CE, Gurley-Calvez, Shireman TI. Indoor tobacco legislation is associated with fewer emergency department visits for asthma exacerbation in children. *Ann Allergy Asthma Immunol*. 2016;117(6):641-645.
- Frieden T, Bassett M, Thorpe L, Farley T. Public health in New York City, 2002-2007: confronting epidemics of the modern era. *Int J Epidemiol*. 2008;37(5):966-77.
- Bodker M, Pisinger C, Toft U, Jorgensen T. The rise and fall of the world's first fat tax. *Health Policy*. 2015;119(6):737-42.
- Smed S, Scarborough P, Rayner M, Jensen J. The effects of the Danish saturated fat tax on food and nutrient intake and modelled health outcomes: an econometric and comparative risk assessment evaluation. *Eur J Clin Nutr*. 2016;70(6):681-6.
- Taber D, Chiqui J, Powell L, Chaloupka F. Banning all sugar-sweetened beverages in middle schools: reduction of in-school access and purchasing but not overall consumption. *Arch Pediatr Adolesc Med*. 2012;166(3):256-62.
- Dubowitz T, Ghosh-Dastidar M, Cohen DA, et al. Diet and perceptions change with supermarket introduction in a food desert, but not because of supermarket use. *Health Aff (Millwood)*. 2015;34:1858-68.
- Cummins S, Flint E, Matthews SA. New neighborhood grocery store increased awareness of food access but did not alter dietary habits or obesity. *Health Aff (Millwood)*. 2014;33:283-91.
- Elbel B, Mijanovich T, Kiszko K, Abrams C, Cantor J, Dixon LB. The introduction of a supermarket via tax-credits in a low-income area. *Am J Health Promot*. 2017;31(1):50-66.
- Brown BB, Werner CM, Tribby CP, Miller HJ, Smith KR. Transit use, physical activity, and body mass index changes: objective measures associated with complete street light-rail construction. *Am J Public Health*. 2015;105:1468-74.
- MacDonald JM, Stokes RJ, Cohen DA, Kofner A, Ridgeway GK. The effect of light rail transit on body mass index and physical activity. *Am J Prev Med*. 2010;39:105-12.
- Panther J, Heinen E, Mackett R, Ogilvie D. Impact of new transport infrastructure on walking, cycling, and physical activity. *Am J Prev Med*. 2016;50(2):e45-53.
- Bennett WL, Cheskin LJ, Wilson RF, et al. Methods for Evaluating Natural Experiments in Obesity: Systematic Evidence Review. Comparative Effectiveness Review no. 204. (Prepared by the Johns Hopkins University Evidence-based Practice Center under contract no. 290-2012-000071-I.) AHRQ publication no. 18-EHC006-EF. Rockville: Agency for Healthcare Research and Quality; 2017.
- Bennett WL, Zhang A, Tseng E, et al. Methods for evaluating natural experiments in obesity: a systematic review. *Ann Intern Med*. 2018 1: 1-10.

20. **McGuire S.** Institute of Medicine. 2012. Accelerating progress in obesity prevention: solving the weight of the nation. Washington, DC: the National Academies Press. Adv Nutr. 2012;3(5):708–9.
21. **Berkman ND, Lohr KN, Ansari M, et al.** Grading the Strength of a Body of Evidence When Assessing Health Care Interventions for the Effective Health Care Program of the Agency for Healthcare Research and Quality: An Update. Methods Guide for Effectiveness and Comparative Effectiveness Reviews [Internet]. 2013.
22. **Hu Y, van Lenthe FJ, Judge K, et al.** Did the English strategy reduce inequalities in health? A difference-in-difference analysis comparing England with three other European countries BMC Public Health. 2016;16:865.
23. **Bolton KA, Kremer P, Gibbs L, Waters E, Swinburn B, de Silva A.** The outcomes of health-promoting communities: being active eating well initiative—a community-based obesity prevention intervention in Victoria, Australia. Int J Obes (Lond). 2017;41(7):1080–1090.
24. **Cleland V, Dwyer T, Blizzard L, Venn A.** The provision of compulsory school physical activity: associations with physical activity, fitness and overweight in childhood and twenty years later. Int J Behav Nutr Phys Act. 2008;5:14.
25. **Webb E, Netuveli G, Millett C.** Free bus passes, use of public transport and obesity among older people in England. J Epidemiol Community Health. 2012;66:176–80.
26. **Webb E, Laverty A, Mindell J, Millett C.** Free bus travel and physical activity, gait speed, and adiposity in the english longitudinal study of ageing. Am J Public Health. 2016;106:136–42.
27. **Camacho-Rivera M, Rosenbaum E, Yama C, Chambers E.** Low-income housing rental assistance, perceptions of neighborhood food environment, and dietary patterns among latino adults: the AHOME Study. J Racial Ethn Health Disparities. 2017;4:346–53.
28. **Sadler RC, Gilliland JA, Arku G.** A food retail-based intervention on food security and consumption. Int J Environ Res Public Health. 2013;10:3325–46.
29. **Sturm R, Hattori A.** Diet and obesity in Los Angeles County 2007–2012: is there a measurable effect of the 2008 “fast-food ban”? Soc Sci Med. 2015;133:205–11.
30. **Restrepo BJ.** Calorie labeling in chain restaurants and body weight: evidence from New York. Health Econ. 2017;26(10):1191–209.
31. **Bekemeier B, Pui-Yan Yip M, Flaxman AD, Barrington W.** Five community-wide approaches to physical activity promotion: a cluster analysis of these activities in local health jurisdictions in 6 states. J Public Health Manag Pract. 2018;24(2):112–20.
32. **Chen C, Chou S-Y, Thornton RJ.** The effect of household technology on weight and health outcomes among chinese adults: evidence from China's 'home appliances going to the countryside' policy. J Hum Cap. 2015;9:364–401.
33. **Brown AL, Khattak AJ, Rodriguez DA.** Neighbourhood types, travel and body mass: a study of new urbanist and suburban neighbourhoods in the US. Urban Stud. 2008;45:963–88.
34. **Rigdon J, Berkowitz SA, Seligman HK, Basu S.** Re-evaluating associations between the Supplemental Nutrition Assistance Program participation and body mass index in the context of unmeasured confounders. Soc Sci Med. 2017;192:112–24.
35. **Durand CP, Oluyomi AO, Gabriel KP, et al.** The effect of light rail transit on physical activity: design and methods of the travel-related activity in neighborhoods study. Front Public Health. 2016;4:103.
36. **Taksler GB, Kisko K, Abrams C, Elbel B.** Adults who order sugar-sweetened beverages: sociodemographics and meal patterns at fast food chains. Am J Prev Med. 2016;51(6):890–897.
37. **Huang TT, Wyka KE, Ferris EB, et al.** The Physical Activity and Redesigned Community Spaces (PARCS) Study: Protocol of a natural experiment to investigate the impact of citywide park redesign and renovation. BMC Public Health. 2016;16(1):1160.
38. **Mayne SL, Auchincloss AH, Michael YL.** Impact of policy and built environment changes on obesity-related outcomes: a systematic review of naturally occurring experiments. Obes Rev. 2015;16(5):362–75.
39. **Sisnowski J, Street JMA-Ohoo, Merlin T.** Improving food environments and tackling obesity: A realist systematic review of the policy success of regulatory interventions targeting population nutrition. PLoS One. 2017;12(8):e0182581.
40. **McCormack GR, Shiell A.** In search of causality: a systematic review of the relationship between the built environment and physical activity among adults. Int J Behav Nutr Phys Act. 2011;8:125.

APPENDIX

Table 4 PubMed Search Strategy

#	Term/string
1	obesity[mh]
2	obese[tiab]
3	obesity[tiab]
4	overweight[tiab]
5	“over weight”[tiab]
6	BMI[tiab]
7	“body mass index”[tiab]
8	“Body mass index”[mh]
9	1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7 OR 8
10	“Policy”[Mesh]
11	policy[tiab]
12	policies[tiab]
13	law[tiab]
14	Purchasing[tiab]
15	Purchases[tiab]
16	“food labeling”[mh]
17	“calorie information”[tiab]
18	“calorie labeling”[tiab]
19	Environment[tiab]
20	10 OR 11 OR 12 OR 13 OR 14 OR 15 OR 16 OR 17 OR 18 OR 19
21	9 AND 20
	Date limited: 2000 to present