3

THE IMPACT OF URBAN FORM ON WEIGHT LOSS

Combining a Spatial Agent-Based Model with a Transtheoretical Model of Health Behavior Change

Arika Ligmann-Zielinska, Sue C. Grady, and Jeremy McWhorter

Research Question: How might an intervention aimed at changing behavior impact the prevalence of obesity in San Diego?

System Science Method(s): Agent-based models

Things to Notice:
- Simulation of a real spatial environment using empirical data
- Testing the effectiveness of a hypothetical intervention

Since the early 1980s, obesity has become a major health issue in the United States. Although the prevalence of obesity is attributed to a myriad of biosocial, chemical, economic, and cultural factors that constitute a complex hierarchical system, published research suggests that the most important contributors to the obesity epidemic are individual behaviors (diet and physical activity), social interactions, and local environments. Three major improvements to existing obesity analytical approaches have recently been proposed. First, we should extend the methods to include a temporal dimension so that potential explanatory pathway(s) of obesity epidemics can be generated and evaluated. Second, we should make use of the constantly growing individual health databases. Third, we should include spatial heterogeneity and dependence of food and physical activity systems. This chapter reports on a research project that incorporates these three postulates in a spatial empirical agent-based model of obesity dynamics. The model provides a platform for computational experimentation, in which a synthetic population of heterogeneous human agents occupies a GIS-based urban environment. The model allows for the exploration of obesity prevalence by incorporating empirical health and geographic data collected for three selected neighborhoods in San Diego, California, USA, and integrated into a model that simulates weight change measured using the body mass index (BMI) due to the combined impact of health behavior and the built environment. The model is equipped with the transtheoretical sub-model of behavior change that mimics a public health intervention aiming at long-term lifestyle change. Based on the results we conclude that significant differences in obesity dynamics exist between the neighborhoods in San Diego. We also observe that a simple policy intervention can substantially impact population-specific weight loss over a period of five years. Using computational experimentation, we demonstrate how agent-based modeling can augment the conventional obesity analytical frameworks, by providing a means of studying the dynamics of obesity in spatially heterogeneous population and physical activity systems.
Worldwide, obesity claims more deaths annually than starvation and other underweight-related causes of death combined. Obesity is defined as a body mass index (BMI) of 30.0 or higher calculated by dividing an individual’s weight / height$^2$. Normal weight is defined as a BMI 18.5 to 24.9, with overweight BMI 25.0 to 29.9. Since 1980, the number of obese persons has doubled across the globe (2014). The United States holds the title of the fattest nation in the world—a dubious distinction—as evidenced by an obesity prevalence rate of 31.8% (FAO, 2013).

Although the prevalence of obesity is attributed to a myriad of biosocial, chemical, political, and cultural factors that constitute a complex system, published research suggests that the most important risk factors for obesity are individual behaviors (diet and physical activity), socioeconomic status, and local residential environments (Alvanides, Townshend, & Lake, 2010; Campbell & Dhand, 2000; Haslam & James, 2005; Hill, Wyatt, Reed, & Peters, 2003; Larkin, 2003; Townshend, Ells, Alvanides, & Lake, 2010). Consequently, there is an urgent need to address the problem of obesity from a variety of perspectives.

The well-known conceptual frameworks of obesity—namely, ANalysis Grid for Environments Linked to Obesity (Swinburn, Egger, & Raza, 1999), and International Obesity Taskforce (Kumanyika, Jeffery, Morabia, Ritenbaugh, & Antipatis, 2002)—are useful tools to represent the complex web of food-activity interrelationships, and theorize the potential ‘explanatory pathways’ of the obesity epidemic (Pearce & Witten, 2010). These models, however, have limited operational value due to their static, aspatial, and highly abstract nature. Conversely, real world applications to study obesity employ in-depth national surveys such as the Nutritional Health and Nutrition Examination Survey (CDC, 2014) or Behavioral Risk Factor Surveillance System – BRFSS (CDC, 2010), and geographic information systems – GIS (Curtis & Lee, 2010; Drewnowski, Rehm & Solet, 2007; Leslie et al., 2007; Plantinga & Bernell, 2007; Stewart et al., 2011). A major limitation of these studies, however, is the inadequate representation of causality due to the restraining nature of conventional (spatial) statistical methods. Pearce and Witten (2010) propose three major improvements to existing obesity modeling and analysis approaches: (1) introducing dynamics by adding a temporal dimension, (2) making use of additional health databases, and (3) incorporating geographic aspects of food and physical activity systems.

In response to general trends concerning obesity dynamics, agent-based models (ABMs) have received increasing attention as tools to investigate how dynamic processes shape the distribution of health outcomes (Auchincloss & Diez Roux, 2008; Galea, Riddle, & Kaplan, 2010). In their pioneering work Yang et al. (Yang, Diez Roux, Auchincloss, Rodriguez, & Brown, 2011, 2012) demonstrated how ABMs might help to better understand the determinants of walking and identified the most promising interventions to increase physical activity by walking. Evidence-based studies on health have concluded that “who you are” (e.g. age, gender, race, income, social status, routines, and lifestyle habits) is the main predictor of your overall health, but that “where you live” also matters (Brewis, 2010; Pickett & Pearl, 2001). Agent-based modeling provides the opportunity to systematically model how people’s interactions within their local environment(s) offer protection from, or contributes to the obesity epidemic.

The objective of this chapter is to address a critical barrier to understanding the micro-level complexities of obesogenic systems (systems that contribute to obesity) by developing and evaluating a prototype of a spatial empirical ABM, which we call oABM (agent-based model of obesity). Spatial agent-based modeling is a method of computational experimentation aimed at modeling dynamic systems, in which individual entities (like humans) operate in a common heterogeneous geographic environment (Brown, Riolo, Robinson, North, & Rand, 2005; Ligmann-Zielinska, 2010). ABM is well suited to address the complex issue of obesity for a number of reasons. First, it directly simulates the interactions between people and those urban
structures that affect weight loss and gain. Second, it represents obesity as a system-level property resulting from unbalanced and unhealthy diet reinforced by the lack of physical activity. By operating at an individual-level, the oABM accounts for the variability in sociodemographic, economic, health, and locational characteristics of people, their interactions, feedbacks, and effects on the emergence of obesity. Third, oABM simulates temporal variability and, compared to other modeling approaches, provides a relatively easy procedure to track causality. Tracking obesity is a longitudinal process composed of different stages, including transitional and maintenance lifestyles. Consequently, explicit representation of temporal variability is critical in studies that assess the long-term implications of public health policy interventions.

The oABM presented here allows for exploration of obesity prevalence by incorporating individual, geographic and empirical health data integrated into a model that simulates weight change measured using the BMI. Our oABM is designed to examine different levels of energy expenditure though two distinct forms of physical activity: resistance training and aerobics. The overarching objective is to simulate the joint impact of health behavior and the built environment on BMI dynamics. We choose three distinct spatial factors previously identified as potential drivers of physical activity: accessibility to exercise amenities (Roux et al., 2007; Sallis et al., 1990), neighborhood walkability (Frank, Andresen, & Schmid, 2004; Leslie et al., 2007; Yang et al., 2012), and neighborhood safety (Wilson, Kirtland, Ainsworth, & Addy, 2004). Assuming that a portion of overweight and obese individuals (represented as agents in the model) in a neighborhood decides to lose weight, and the lifestyle change intervention is framed in a selected model of health behavior augmented by social support, we focus on the following research questions: “To what extent does accessibility to physical activity facilities influence BMI change?” “Do different configurations of the built environment affect walkability?” “How does neighborhood safety affect exercise and, as a consequence, BMI dynamics?”

The remainder of this chapter is organized as follows. First we introduce the case study of San Diego, California. We then describe two variants of our oABM, before outlining six computational experiments conducted using the oABM. What follows is a detailed description of datasets and data processing necessary to parameterize the model. Analysis of the results is reported, after which we reflect on the limitations of the study, the challenges faced when developing agent-based models, and the opportunities for improvement. We conclude the chapter by providing insight into future directions of research.

**Study Site**

For our study, we selected a region characterized by high obesity prevalence, albeit with a diverse population to better model the country as a whole. In 2010, San Diego, California was the 10th most obese (26 per 100 people) large city in the United States (Figure 3.1), and of the other top obese cities, it was one of the most demographically heterogeneous (CDC, 2012). San Diego is also an attractive location for modeling obesity because of its limited weather variability throughout the year, which minimizes confounding in activity due to weather and allows for maintaining the same physical activity routines over longer periods of time.

Given the large size of San Diego and the inability to sample all corridors of the city, we chose three distinct districts to represent the city as a whole. These include Emerald Hills, a low- to medium-income area with a large presence of African Americans; La Jolla, a predominately white and high-income area; and Logan Heights, a predominately Latino and low-income area (Figure 3.1, Table 3.1) (U.S. Census Bureau, 2014). The physical environments of the districts were evaluated using a number of spatial metrics, including compactness, the density of physical activity centers (referred to as gyms), and walkability (Figure 3.2, Table 3.1) (Galster et al., 2001).
Overall, these three areas are characterized by population and environment heterogeneity, serving as a microcosm of spatial diversity of the San Diego metropolitan area. Finally, previous studies in San Diego suggest that community design and access to exercise amenities have a considerable influence on individual physical activity in the city (Frank et al., 2004; Sallis et al., 1990) making this an ideal city to study the role of physical activity on obesity prevalence in districts with different populations and environmental characteristics.

Model Description

Our agent-based model of obesity (oABM) is a simulation method in which adult individuals (adult ≥ 18 years) are represented by heterogeneous decision-making entities (agents), who follow their daily diets and perform physical activities in a shared spatial environment. The synthetic population of human agents occupies an artificial space composed of parcel lots (places of agents’ residence), fitness centers (gyms), walkable roads and crime zones. The fundamental agent behavioral mechanism is the energy balance model (EBM) of energy intake and energy expenditure to imitate weight dynamics (Figure 3.3). Each agent is equipped with weight in kg (w), height in m (h), sex (s), and age in years (y) derived from empirical data. These attributes

![Figure 3.1 Three districts in San Diego, CA, used in this study. BMI: body mass index. Initial BMI represents the start value of all ABM simulations, scaled down from BRFSS data collected in 2009 for the whole of San Diego.](image-url)
are used to estimate agent’s BMI, equation 1 (WHO, 2006), and its basal metabolic rate, BMR, equation 2 (in calories), which is the energy expended daily at rest (Mifflin et al., 1990). Agents also have a lifestyle attribute called a physical activity level (PAL, unitless) which varies from sedentary to vigorous. PAL and BMR are then used to calculate the total energy expended by agents in a day (TEE, in calories) (FAO, 2004):

\[ \text{BMI} = \frac{w}{h^2} \]  

\[ \text{BMR} = 10w + 6.25h + 5y + c \]  

\[ \text{TEE} = \text{BMR} \times \text{PAL} \]

where \( c = 5 \) for males and \( c = -161 \) for females (Mifflin et al., 1990).

The EBM operates as follows (Figure 3.3). For a given day, an agent retrieves its BMR and PAL to calculate TEE. An agent carries a value of calories – \( \text{agcal} \) (estimated based on secondary data), which are consumed daily. An agent is also assigned empirically derived workout time (in minutes) and intensity (calories burned per minute) to calculate total calories burned.

---

Table 3.1 Selected spatial metrics of the tree districts

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Description</th>
<th>Emerald</th>
<th>Logan Hills</th>
<th>Lajolla Heights</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>Total number of adults</td>
<td>53213</td>
<td>64299</td>
<td>34471</td>
</tr>
<tr>
<td>Female</td>
<td>Female population (% of adults)</td>
<td>50.5</td>
<td>41.1</td>
<td>51.7</td>
</tr>
<tr>
<td>White</td>
<td>White population (% of total adults)</td>
<td>28.0</td>
<td>20.0</td>
<td>90.0</td>
</tr>
<tr>
<td>Black</td>
<td>Black population (% of total adults)</td>
<td>30.0</td>
<td>10.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Latino</td>
<td>Latino population (% of total adults)</td>
<td>42.0</td>
<td>70.0</td>
<td>9.0</td>
</tr>
<tr>
<td>Low income</td>
<td>Population with income below poverty level (%)</td>
<td>29.0</td>
<td>43.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Agents</td>
<td>Number of agents</td>
<td>263</td>
<td>319</td>
<td>171</td>
</tr>
<tr>
<td>Area</td>
<td>District area in square miles</td>
<td>10.5</td>
<td>7.2</td>
<td>10.1</td>
</tr>
<tr>
<td>Perimeter</td>
<td>District perimeter in miles</td>
<td>16.6</td>
<td>15.8</td>
<td>15.8</td>
</tr>
<tr>
<td>Population density</td>
<td>Average number of adults per square mile density</td>
<td>5068</td>
<td>8930</td>
<td>3413</td>
</tr>
<tr>
<td>Gym density</td>
<td>Average number of gyms per square mile</td>
<td>0.4</td>
<td>1.4</td>
<td>4.0</td>
</tr>
<tr>
<td>Gym proximity</td>
<td>Average number of gyms in 5 mile radius from agent</td>
<td>0.3</td>
<td>0.5</td>
<td>3.3</td>
</tr>
<tr>
<td>Crime</td>
<td>Neighborhood crime index</td>
<td>0.2</td>
<td>0.2</td>
<td>0.7</td>
</tr>
<tr>
<td>Walkability</td>
<td>Walkable roads in neighborhood</td>
<td>4.2</td>
<td>5.6</td>
<td>4.5</td>
</tr>
<tr>
<td>Agent clustering</td>
<td>Variation coefficient of nearest neighbor distance (%)</td>
<td>56.0</td>
<td>67.0</td>
<td>62.0</td>
</tr>
<tr>
<td>Gym accessibility</td>
<td>Agents within 0.5 mile walking distance from gyms (%)</td>
<td>10.0</td>
<td>8.9</td>
<td>29.0</td>
</tr>
<tr>
<td>Shape index</td>
<td>Compactness (most compact is square with value 1)</td>
<td>1.3</td>
<td>1.5</td>
<td>1.2</td>
</tr>
</tbody>
</table>

Number of agents in the oABM is proportional to adult population where one agent represents approximately 200 people. Data processing: Network Analyst in ArcGIS (www.esri.com/software/arcgis/extensions/networkanalyst), GeoDa (http://geodacenter.asu.edu/), and FRAGSTATS (www.umass.edu/landeco/research/fragstats/fragstats.html). Census data for 2010. The study concerns only adults due to obesity data availability.
banced) during physical activity like muscle strength-training and aerobic. At the end of the day, agent’s weight is updated as follows:

\[
wt_{t+1} = wt_t + \frac{(agcal - TEE - burned)}{bf}
\]

where \(wt_t\) is weight in kg at the beginning of the day, \(wt_{t+1}\) is weight in kg at the end of the day, and \(bf = 7716\) calories/kg is an approximated amount of energy needed to burn 1 kg of body fat (FAO, 2004). What follows is a daily update of BMI, BMR, and TEE. Importantly, after each year, the agent updates its BMR and TEE to account for its metabolism decrease with increasing age (Mifflin et al., 1990).

The EBM reflects a simple projection of agent’s current lifestyle into the future. In order to simulate an increase in exercise and, consequently, weight loss, we introduced a public health policy intervention by adopting the transtheoretical model of behavior change (TTM), Figure 3.4. The TTM, first proposed by Prochaska and colleagues in the 1970s (Prochaska & DiClemente, 1992), provides a temporal framework leading to healthier diet behavior. TTM has been used in a number of interventions including smoking cessation, stress management, and weight change (Dray & Wade, 2012; Johnson et al., 2008). The model is composed of five stages: precontemplation, contemplation, preparation, action, and maintenance. TTM is implemented in our oABM as follows (Figure 3.4). The agent starts from its base lifestyle according to its weight and adjusts its lifestyle accordingly.
Figure 3.3 The energy balance behavioral rule applied daily to agents. BMR: basal metabolic rate; TEE: total energy expenditure based on agent’s physical activity level.

Figure 3.4 Transtheoretical model of health behavior change applied to agents as policy intervention.
to EBM (precontemplation). If its BMI > 25.0 (i.e., exceeds the normal range) the agent considers lifestyle change (contemplation), provided that it can be influenced by its physician, dietitian, or other specialist. What follows is the preparation stage during which the agent maintains its base lifestyle. Within one day to six months (readiness), the agent makes gradual adjustments to behavior by increasing its physical activity (action). For simplicity, we assume that its diet remains unchanged and focus solely on fitness. Moreover, during the transitional lifestyle, the agent may get support from its social network. The agent evaluates its BMI on a daily basis, and if it is dissatisfied with the results, it continues the transitional stage, provided that its persistent. If a relapse occurs, the agent goes back to its initial lifestyle. If the agent’s BMI falls within the normal range, it moves to maintenance, during which its energy input is balanced by energy output. As before, the agent may relapse into its base lifestyle, after which the whole TTM cycle is repeated.

We introduced the spatial component of the model into the TTM by adjusting the agent’s burned calories. This modification is dependent on the choice of exercise (a workout in a gym, a walk or a combination of both). Walking can be further reduced due to high crime in agent’s neighborhood.

In summary, the EBM requires five attributes: age, weight, height, workout, and calories burned, which are populated based on data obtained from the 2009 BRFSS (CDC, 2010). The TTM needs additional four variables. Due to the lack of empirical observations, influenceability, persistence, and readiness are all populated with simulated values whereas social support is approximated based on previous studies (Kiernan et al., 2012; Leahey, Kumar, Weinberg, & Wing, 2012; Wing & Jeffery, 1999).

Computational Experiments

To account for different combinations of the three spatial factors evaluated in the model, namely, accessibility to gyms, walkability, and neighborhood safety, designed six different computational experiments. Each experiment was executed 500 times, using simple random sampling. The oABM output (called avgBMI in the following sections) is the value of BMI averaged over all agents in a given district and reported for each model run at the end of model execution. The oABM runs for five years with one day increments, which amounts to 1825 execution loops. We selected five years as the simulation time period to assure that the experiments run long enough to render a trend in BMI rather than short-term fluctuations (i.e. the model reached a stable equilibrium after 1825 steps), and short enough so that permanent and major changes in the built environment can be ignored.

The first model (EXP0, Table 3.2) is a lifestyle maintenance model with EBM as the only mechanism driving BMI dynamics. In this version of the oABM, the agents, who emulate individuals recorded in BRFSS, do not change their eating, PAL, and workout habits over the course of model execution. They follow the same lifestyle from day to day, they do not interact, and they do not evaluate their environment in order to select physical activity sites. In this sense, EXP0 produces a benchmark output distribution of avgBMI which is a simple projection of recorded observations into the future. This oABM is therefore conceptually closer to microsimulation than a fully fledged ABM, and is often referred to as a protomodel or a model with protoagents (North & Macal, 2007). Models two through six (EXP1 to EXP5, Table 3.2) are all variants of the lifestyle change TTM, in which access to physical activity is represented using different combinations of selected aspects of the workout-relevant built environment: gym access, walkability, and neighborhood safety. In EXP1 (labeled GYM), the only type of exercise
allowed is workout in gyms, which predominantly represents muscle strength (resistance) training. The factor influencing calories burned is distance to the nearest gym. The further the gym, the less likely that the agent will do its strength training in a given day. EXP2 (WALK) is the other extreme case, where agents can only walk/run in their neighborhood. It represents mainly aerobic or cardio training. In this case, the lower the surrounding walkability, the less likely it is for the agent to walk or run in a given day. EXP3 is a modification of EXP2, in which walkability is further modified by neighborhood safety (SAFE WALK) – the higher the crime index, the less likely that the agent takes a walk in a given day.

EXP4 and EXP5 represent combinations of resistance training (GYM) and aerobics (SAFE WALK). Specifically, in EXP4 agents can select either gym training or walking/running (chosen randomly with equal probability), whereas EXP5 allows for a combination of GYM and SAFE WALK. The specific rule driving EXP5 is as follows. An agent randomly selects what fraction of workout in a given day is allocated to GYM. SAFE WALK is then calculated as the remainder of the daily workout. What follows is that the GYM fraction is further influenced by accessibility to activity amenities (as in EXP1) whereas the SAFE WALK fraction is influenced by neighborhood walkability and safety (as in EXP3). Note that EXP5 was designed to reflect physical activity guidelines for adults issued by the U.S. Department of Health and Human Services (HHS, 2008). These physical activity guidelines introduce variation to exercise that encourages lifestyle change endurance by combining muscle strength-training and aerobic. This combination of activity has been shown to burn body fat rather than muscle, leading to better physique, much more effective weight reduction, and ensuring such health benefits as decreasing blood pressure and/or glucose levels (Church, Blair, Cocreham, & et al., 2010; HHS, 2008; Sigal et al., 2007; Thorogood et al., 2011). In all five experiments, the decision to actively burn excess calories is independently made every day. The model was implemented in Python (www.python.org/) and executed using the computing resources in the High Performance Computer Center at Michigan State University (http://icer.msu.edu/).
Data Acquisition and Processing

For the oABM, we utilize publicly available data. Except for the TTM parameters, all inputs are derived from empirical observations. Model inputs comprise agent data and spatial data. Below we describe data sources and data processing required to fine-tune the oABM.

Agent Data

The EBM requires demographic and body measurement data collected at an individual level. We used a subset of the 2009 California BRFSS survey data collected from adults only, and aggregated for selected zip codes (n = 408) in the San Diego metropolitan area. The data was pre-processed by Survey Research Group (http://s-r-g.org/).

Research suggests that significant differences in BMI exist among different demographic groups, and variables like sex, race, or income can be used as parameters to differentiate population groups (Sallis et al., 2009). Consequently, based on the sample data, we subdivided the BRFSS variables into four empirical agent groups: male (M); female white (FW); female black/Latino and low income (FBL); and female black/Latino and other income (FBO). For each agent grouping, we generated probability density functions for the following model inputs used in EBM: age, weight, height, and workout time. These probability density functions were later

![Figure 3.5 Probability density functions for age, weight, height, and exercise duration generated for the four distinct groups of agents from the BRFSS sample](image-url)
sampled in order to build diverse parameter sets used in model runs, i.e., when a simulation was populated, each agent’s value for every input was randomly drawn from these functions. Figure 3.5 summarizes the distributions.

EBM also requires estimates for PAL, calories consumed, and calories burned to calculate the total energy expenditure and daily weight change (Equations 3 and 4). Since such measurements are not recorded in BRFSS, we used secondary resources to approximate the probability density functions for these variables. Table 3.3 shows the estimates for calories burned per agent group. PAL data by sex, age, and BMI for adults comes from National Academies Press (NAP, 2005) and Roberts and Dallal (2005). Calories consumed were derived from data on mean energy intake among adults by sex and age reported by CDC (Wright & Wang, 2010).

We did not find any data that could be directly utilized to set the TTM parameters (Figure 3.4). Consequently, we had to resort to stylized data. After a number of test simulations, we set the probability density functions of influenceability to a uniform distribution with minimum = 0.001, and maximum = 0.05, and the probability density functions of persistence to a uniform distribution with minimum = 0.95, and maximum = 0.999. Since, by definition, both influenceability and persistence can vary from zero to one, the assumed distributions denote that agents are not easily influenced but, if they decide on lifestyle change, they are fairly committed. We also assumed that, to lose weight, agents have to perform additional workout burning from 250 to 760 calories daily, so that they lose from 0.5 to 1.5 pounds per week.

### Spatial Data

The BRFSS data is de-identified and, as a result, it is not geocoded. To allocate agents, we used ancillary information on the number of adults, their sex, race, and income per census tract (U.S. Census Bureau, 2014). These variables were used to subdivide adult population into FBL, FBO, FW, and M. We narrowed down the acceptable locations only to residential parcels (Table 3.4) and then randomly spread a number of points proportional to the adult population in a given tract (Figure 3.2). After attribute initialization, agents were assigned to the points in a manner that reflects the percentage of FBL, FBO, FW, and M in a given tract. AvgBMI at time zero in each of the three study districts in San Diego is shown in Figure 3.1.

Data on recreation and athletic centers (physical activity centers in Figure 3.6) and calculated distance to the nearest workout facility was retrieved for each agent point. Walkability was operationalized based on previous conceptual and empirical studies (Saelens, Sallis, & Frank, 2003). We used walkable roads (roads with pedestrian-oriented design, street connectivity, and sidewalk) as the base spatial dataset (Table 3.4). We first derived a road density surface which was then reclassified into nine equal-interval categories, where high walkability is associated with high density of walkable roads (Table 3.4, Figure 3.6 left). Our rationale for linking

---

### Table 3.3 Parameters for the triangular probability distribution of calories burned per minute for the four distinct groups of agents

<table>
<thead>
<tr>
<th>Group</th>
<th>Calories per minute</th>
</tr>
</thead>
<tbody>
<tr>
<td>FW</td>
<td>a = 2.8, b = 9.1, c = 5.7</td>
</tr>
<tr>
<td>FBL</td>
<td>a = 3.3, b = 6.6, c = 10.5</td>
</tr>
<tr>
<td>FBO</td>
<td>a = 2.7, b = 8.7, c = 5.4</td>
</tr>
<tr>
<td>M</td>
<td>a = 3.6, b = 11.5, c = 7.2</td>
</tr>
</tbody>
</table>

Data source: [www.nutristrategy.com/](http://www.nutristrategy.com/)
high road density areas with longer exercise is based on a San Diego study by Saelens, Sallis, Black, & Chen (2003). Data from that study showed that, in the high walkability neighborhoods, individuals walk/run on average about 210 minutes per week, as opposed to low walkability communities where the duration of walking/running amounts to about 140 minutes weekly. Finally, agent neighborhood safety was evaluated using total crimes per police neighborhood, normalized by the maximum number of crimes recorded in San Diego (Table 3.4, Figure 3.6 right).

Results and Discussion

The eighteen distributions of output avgBMI for all five experiments (EXP0–EXP5) are rendered in two ways. First, we chart the distributions by experiment (a total of six diagrams with three box plots each) to demonstrate the consequences of spatial heterogeneity on avgBMI with all other processes unchanged (Figure 3.7). Second, we display the results of all experiments per district (three diagrams with six box plots each) to facilitate the comparison between interventions and their influence on avgBMI within a given study site (Figure 3.8).

EXP0 (Figure 3.7) serves as a baseline scenario. At a base lifestyle, after five simulated years, avgBMI is elevated (BMI > 30) demonstrating high obesity prevalence in all three districts. Although the mechanisms driving the null model are spatially independent, significant differences...
in avgBMI between districts are observed (F-test (2, 1497) = 45.96, p-value = 0.00) (Figure 3.7, top left). These differences appear to be due to the dominating agent groups within districts (Table 3.1). For example, output avgBMI for La Jolla = 30.00 (95% CI, 29.96–30.04) compared to Emerald Hills = 30.79 (95% CI 30.76–30.83). This difference is not surprising, since Emerald Hills has a relatively high fraction of adult population falling into the FBL group with the highest initial BMI and, at the same time, La Jolla has a relatively large FW group with

Figure 3.7 Box plots (distributions) of average BMI per all agents in a district at the end of model execution for n = 500 model runs shown by experiment
Figure 3.8 Box plots (distributions) of average BMI per all agents in a district at the end of model execution for \( n = 500 \) model runs shown by district.
the lowest initial BMI. Observe, however, that this association is neither univariate nor linear. To validate the results, we compared the outcomes from the baseline scenario to an independently derived projection. We used adult BMI from the Nutritional Health and Nutrition Examination Survey database (CDC, 2014), recorded every other year from 1999 to 2009, generated a trend line, and extrapolated the trend to 2014 ($R^2 = 0.95$). The resulting $\text{avgBMI} = 30.0$, which is similar to the $\text{avgBMI}$ in EXP0. We can therefore conclude that the base oABM exhibits satisfactory performance.

EXP1 through EXP5 are all equipped with the TTM mechanism. As anticipated, the behavior change model leads to a decrease in $\text{avgBMI}$, even though only a small fraction of overweight and obese agents follows the intervention (due to low agent influenceability). Since we tried to minimize the impact of agent locational distribution on model behavior (the shape index and agent clustering are relatively similar across the districts—Table 3.1) we hypothesize that between-district differences in the TTM experiments are caused solely by the spatial differentiation of gym location, walkability, and crime. There were significant differences between the three districts for all five experiments (data not shown, $p$-value 0.000 for EXP0–EXP5). The influence of accessibility to fitness centers (EXP1) on $\text{avgBMI}$ reduction is moderate. The results indicate that the impact of GYM on $\text{avgBMI}$ is more complex than expected. Based on the three gym metrics in Table 3.1, we expected that the highest drop in $\text{avgBMI}$ would be observed in La Jolla. While all three districts show $\text{avgBMI}$ reduction, it is Logan Heights that exhibits the greatest reduction in $\text{avgBMI}$ ($–1.36$), most probably due to the fact that, although La Jolla has the greatest number of gyms that are relatively accessible, it also has the smallest population of agents that require lifestyle change to reduce their BMI. EXP2 (WALK only) produces a considerably greater drop in $\text{avgBMI}$ from baseline than EXP1 (across all three districts $–2.08$).

Again, the outcomes suggest that the relationship between the fraction of population that employs TTM and the spatial configuration of walkable roads is more complex. While Logan Heights has the highest average walkability (Figure 3.6), it is Emerald Hills and La Jolla that record the greatest BMI reduction ($–2.66$, $–2.25$), possibly due to the relatively high population of FBL.

Neighborhood safety (EXP3) significantly undermined weight loss due to walkability. Logan Heights and Emerald Hills showed only a slight reduction in $\text{avgBMI}$ ($–0.38$, $–0.12$) when compared to the base scenario. The reduction of $\text{avgBMI}$ was greatest for La Jolla ($–1.51$), despite it having the highest crime index of the three districts (Table 3.1), suggestive that this population is walking in locations other than the city (e.g., perhaps walking along the shoreline).

In the last two experiments, we allow agents to choose GYM or WALK (but not both – EXP4) or they can combine walking with resistance training (EXP5). In both cases, we observe that the between-district decrease resembles the output distribution of the EBM model, where change in $\text{avgBMI}$ from base lifestyle is the greatest in La Jolla ($–1.13$), followed by Emerald Hills ($–0.84$) and Logan Heights ($–0.81$). Importantly, the downward shift in $\text{avgBMI}$ is the most pronounced in the final, most flexible model (GYM and WALK), in which the likelihood of exercise in any given day is the highest (overall reduction in $\text{avgBMI} = –2.12$; Emerald Hills $= –2.11$, Logan Heights $= –2.05$, La Jolla $= –2.20$).

The differences in $\text{avgBMI}$ from base for each activity presented were significantly different (data not shown, $p$-value $= 0.00$). Recall that the last scenario has the highest potential for long-term maintenance of BMI reduction. These results show the most promise in achieving the goal of 5% reduction in average BMI over the next ten years, which could draw the health care costs down by about $28$ billion in the state of California alone (Levi, Segal, Laurent, Lang, & Rayburn, 2012; Robert Wood Johnson Foundation, 2012).

The results from the five experiments by district are reported in Figure 3.8. Beginning with Emerald Hills, which has the highest $\text{avgBMI}$ of the three districts ($30.79$) WALK (EXP2) and
GYM and WALK (EXP5) show the greatest reduction in avgBMI (−2.66, −2.20) from the base scenario. BMI was least reduced in the SAFE WALK experiment (−0.38) despite the crime index being lower in Emerald Hills compared to Logan Heights and La Jolla (0.7 vs. 0.2, 0.2) suggestive that the perception of crime is more important in influencing outdoor activity than actual crime. In Logan Heights utilizing GYM and WALK (EXP5) reduced avgBMI twice (−2.05) as much as participating in these activities independently (GYM only = −1.39; WALK only = −1.35) again, demonstrating the value of multiple forms of activity. Finally, La Jolla, which had the lowest avgBMI of the three districts, showed WALK (−2.25) and GYM and WALK (−2.11) to be the most important activities for reducing BMI. The high reduction by WALK may be associated with walking in locations other then the city, such as the shoreline. All of these avgBMI differences in activities from the base lifestyle are statistically significant (p-value = 0.00).

The Role of Agent-Based Modeling in Studying Obesogenic Systems

The major benefit of utilizing agent-based modeling in obesity research is the ability to utilize cross-sectional datasets to model longitudinal intervention scenarios, and identify place-specific obesity interventions for public health policy and practice. Cross-sectional health datasets however, are limited by their lack of georeferencing individual activity spaces and potential environmental exposures and detailed representation of social networks. This study attempted to address these data limitations by implementing geographic methods to best allocate surveyed individuals (BRFSS) within three districts, define population (sociodemographic) characteristics most at-risk of obesity to represent the “agents” of study and to utilize available information on lifestyle practices to improve our understanding of behavioral-activity interventions. Another advantage of agent-based models over other commonly used models in obesity research is the ability to actively apply theoretical applications of behavior change and develop assumption scenarios to model and assess the complex interaction between obesity, behavior and environment interactions. Behavioral qualitative variables, like the type of lifestyle, persistence during the transitional period, or the underlying factors of the decision to participate in an intervention, are critical in developing rules that guide the actions of agents. Agent-based modeling allows for tracing the results back to consequences. It demonstrates how modeled behavior, rather than probabilities, affects weight gain and weight loss. It provides a means to evaluate the effectiveness of interventions. Finally, it allows for coupling the dominant drivers of weight change with other indirect causes, which together shape the dynamics of obesity.

The ABM presented here, however, has some distinct limitations that are worth mentioning. For simplicity, this study concentrates solely on calories expended during exercise, when in fact individual weight change is a product of diet and exercise, with such entangled factors as social influence, norms affecting food choices, cultural conventions, or conformism. A comprehensive analysis of the identified sociodemographic groups (FBL, FBO, FW, and M) could provide more valuable information used to design a suite of interventions tailored to specific high-risk cohorts. We excluded the influence of proximity to beaches on physical activity – a potentially significant factor in coastal areas (Bauman, Bellew, & Wales, 1996).

Regardless of these drawbacks, ABM is an immensely useful method for policy-relevant analysis of obesity prevalence. This study found that, across districts of distinct population and environmental characteristics, multiple forms of activity, specifically utilizing the gym and walking, are significantly associated with obesity reduction. In La Jolla, which had a high index of crime, walking was still an important contributor to obesity reduction possibly because of the availability of safe local amenities such as the waterfront. Importantly, the perception of crime
limits walking activity more than the actual crime of an area. Thus crime and physical activity need to be studied in relation to the availability of resources and other environmental characteristics. Black women of low income remain the primary population group to experience obesity and future research should continue to focus on interventions that specifically address this population.

Summary

In this chapter we demonstrate how spatial ABM can be employed to study the prevalence of obesity in selected populations and districts of a major metropolitan area. We develop a prototype model, where individual total energy expenditure is modified by physical activity. We extend the ABM with the transtheoretical model of health behavior change that serves as an intervention aimed at obesity reduction. To parameterize the model we use publicly available (BRFSS) individual and spatial data. Based on the computational experiments, we conclude that the selected geographic factors: accessibility to physical activity centers, walkability, and neighborhood safety, play a substantial role in the spatiotemporal variability of individual weight measured with the body mass index. To a lesser extent, the factors also influence the magnitude of weight loss in the overweight and obese agents.

References

A. Ligmann-Zielinska et al.

FAO. (2013). The State of Food and Agriculture (pp. 100). Rome: Food and Agriculture Organization.


