Current Epidemiology Reports (2019) 6:476–485
https://doi.org/10.1007/s40471-019-00221-5

GENETIC EPIDEMIOLOGY (C AMOS, SECTION EDITOR)

Neighborhoods to Nucleotides—Advances and Gaps for an Obesity Disparities Systems Epidemiology Model

Marta M. Jankowska[1] · Kyle Gaulton[2] · Rob Knight[2,3,4,5] · Kevin Patrick[1,6] · Dorothy D. Sears[3,5,6,7,8]

Published online: 30 October 2019

Abstract

Purpose of Review Disparities in prevalence of obesity in the USA continue to increase. Here, we review progress and highlight gaps in understanding disparities in obesity with a focus on the Hispanic/Latino population from a systems epidemiology framework. We review seven domains: environment, behavior, biomarkers, nutrition, microbiome, genomics, and epigenomics/transcriptomics. We focus on recent advances that integrate at least two or more of these domains, and then provide a real-world example of data collection efforts that encompass these domains.

Recent Findings Research into discrimination-related DNA methylation patterns and how microbiome profiles are related to eating and physical activity behaviors is furthering understanding of why disparities in obesity persist. Environmental and neighborhood level research is uncovering the importance of exposures such as air and noise pollution and systematic or structural racism for obesity and related outcomes through behaviors such as sleep.

Summary Obesity disparities and the biological processes associated with them must be better contextualized within the social, economic, and political environments that contribute to them. One avenue for accomplishing this is by modeling relationships between within-body mechanisms and omics and beyond-body mechanisms and exposures. However, data integration across the various domains and data collection are significant challenges for generating a comprehensive systems model for obesity disparities.

Keywords Health disparities · Hispanic/Latino · Obesity · Systems epidemiology · Environmental exposure · Data integration

Introduction

The prevalence of obesity and associated outcomes are increasing worldwide, and the World Health Organization recognizes obesity as one of the greatest public health challenges of the twenty-first century [1]. In the USA, these trends have disproportionately affected underserved populations with low socioeconomic status (SES) and diverse race/ethnicity, including Latinos [2]. The 2015–2016 National Health and Nutrition Examination Survey estimated that 47% of Hispanics were obese, compared to 37.9% of non-Hispanic Whites. While the prevalence of obesity has remained steady over the past decade for some populations, it is increasing in other populations such as Mexican Americans [3]. The Hispanic Community Health Study/Study of Latinos (HCHS/SOL) reported a diabetes prevalence of 16.9% that varied among ancestry groups and was as high as 18.3% in Mexican Americans, who also had a metabolic syndrome prevalence of 35.0% [2, 4]. Comparatively, diabetes prevalence is estimated at 11.3% among US adults overall [5].

Obesity, insulin resistance, type 2 diabetes, and metabolic syndrome are complex diseases made more so when viewed

This article is part of the Topical Collection on *Genetic Epidemiology*

✉ Marta M. Jankowska
majankowska@ucsd.edu

[1] Qualcomm Institute/Calit2, 9500 Gilman Drive MC 0811, University of California San Diego, San Diego, CA, USA

[2] Department of Pediatrics, UC San Diego, San Diego, CA, USA

[3] Center for Microbiome Innovation, UC San Diego, La Jolla, San Diego, CA, USA

[4] Department of Computer Science and Engineering, UC San Diego, La Jolla, San Diego, CA, USA

[5] Center for Circadian Biology, UC San Diego, San Diego, CA, USA

[6] Department of Family Medicine and Public Health, UC San Diego, La Jolla, San Diego, CA, USA

[7] College of Health Solutions, Arizona State University, Phoenix, AZ, USA

[8] Department of Medicine, UC San Diego, La Jolla, San Diego, CA, USA

through the lens of health disparities. Recent developments in environmental monitoring, high-throughput omics technologies, behavioral and life course monitoring, biomarkers, etc. are expanding our understanding of how the human genome and human body fit into a larger system of factors that play a complex role in the development of obesity and associated outcomes. These efforts have identified novel therapeutic targets such as endocrine disruptors, effects of gut microbiota on obesity, epigenetic and transcriptomic regulation of obesity, and larger-scale sociological and built environment impacts on obesity. At the same time, increased specialization in each of these fields makes devising cumulative measures or systematic models that integrate and span these advances all the more challenging. Systems models for obesity are not new in the field (for classic examples, see the Foresight or Glass and McAtee models [6, 7]). However, where models stress integration across levels, published research has a large disconnect between sociological aspects of researched systems that focus on humans as decision-making agents within a broader socio-political and economic context and internal pathways that make up the biological obesity system. There is a fundamental gap in knowledge about how the various risk factors for obesity integrate and interact with one another [8]. When integration does occur, most prominently through gene × environment studies, focus is often placed on two factors in the system with little done to integrate the resulting relationship into the larger system.

Narrowing efforts to a single or two related risk factors is a significant problem for understanding health disparities. Race and/or ethnicity are not a single covariate to be controlled for, but rather a complex set of factors ranging from genetic variation to access to care to environmental and/or political injustice. In other words, it is too simplistic to consider race or ethnicity as a cause of obesity and must instead look at the number of regulating factors that influence risk in sub-populations. To study obesity and related diseases through the perspective of health disparities, we must adopt a systems approach that can accommodate and integrate many data dimensions and types, assess multi-level effects, incorporate spatial and temporal changes, consider synergistic or attenuating effects, and account for socio-political and environmental context [7, 9]. Multi-level and complex models are better suited to arriving at *why* race and ethnicity matter for obesity disparities.

Here, we review progress and highlight gaps in understanding disparities in obesity from a systems epidemiology framework drawing on numerous omics, behavioral, environmental, and sociological fields. Understanding obesity disparities with a systems approach is daunting for several reasons, one of the largest being choosing which of the many risk factors of obesity to include. Figure 1 illustrates the risk factors focused on in this review in larger bubbles (environment, behavior, biomarkers, nutrition, microbiome, genomics, epigenomics/transcriptomics). This is not an exhaustive list, but rather reflects areas where recent advances have been made in relationships *between* factors (e.g., linkages between microbiome and nutrition, or epigenomics and environment). We then provide a real-world example of using a systems epidemiology approach for data collection, processing, and analysis of a cohort of individuals in a study entitled Nucleotides to Neighborhoods, which focuses on obesity and disparities between Hispanic/Latino and non-Hispanic/Latino individuals.

Advances and Gaps for an Obesity Disparities Systems Epidemiology Model

Genomics, Epigenomics, and Transcriptomics

Genetic variants play a large role in obesity and type 2 diabetes with studies finding Hispanic/Latino specific effects on variants on metabolic syndrome components and regulation of weight, sleep duration, and total energy expenditure [10, 11]. Assessment of genotypic interactions with nutrients [12], behaviors [13••], environment [14•], and the microbiome [15•] shows that these interactions have the capability to modify obesity and related outcomes. Research into the interplay between lifestyle, environment, microbiome, and genetic factors is limited in Hispanic/Latino populations with some focus on hepatic fat and diet [16]. One study has assessed the effects of physical activity and sedentary behavior on genetic variants on obesity in Hispanics/Latinos finding that increased physical activity and reduced sedentary time attenuates genetic association with obesity [17••]. A challenge for genomic research in Hispanic/Latinos is the admixture of the population, which sees a large genetic diversity with European, African, and Indigenous American ancestry [18].

Epigenomic mechanisms are of significant interest in obesity because they modify gene activity without changing the underlying DNA sequence. Research in this area has exploded in recent years, with a large focus on DNA methylation. DNA methylation generally silences gene activity, can be dynamic or static throughout adult life, and can be passed on to offspring [19]. Significant research has been conducted on changes in DNA methylation in obese individuals and type 2 diabetics [20, 21], with some research specifically in Hispanic/Latino cohorts [22–24]. In animals and humans, a number of behaviors influence DNA methylation including nutritional changes (high-fat diet, low-protein diet, high-nutrient-specific diet), physical activity, stress, smoking/alcohol consumption, and working habits [25]. Few of these studies specifically examine behaviors and methylation as related to racial disparities; however, racial discrimination has recently emerged as a measure of interest in this realm. In a study of Latina mothers, perceived everyday discrimination was significantly associated over time with DNA methylation of stress-related genes [26•]. This research could identify a functional mechanistic link between stress, inflammation,

Fig. 1 A systems epidemiology approach for understanding obesity disparities including seven risk factor domains (environment, behavior, biomarkers, nutrition, microbiome, genomics, and epigenomics/transcriptomics). Example measures for these domains are represented in smaller bubbles of the same color (e.g., biomarkers are measured through urine, blood, and clinical assessments). Additional domains and measures could be built into and/or swapped out of the model, as needed for a given study

and obesity in minority populations. Psychosocial factors may also play a role at the neighborhood level, with limited research into socioeconomic status (SES) showing that living in lower SES neighborhoods is associated with greater methylation of stress-response and inflammation-related genes even after accounting for individual SES [27••].

Other areas of promising research for relationships linking obesity, behaviors, and environment to genome function include epigenomic marks such as chromatin accessibility, histone modifications, and transcriptomic features such as microRNAs (miRNAs). External stimuli driven by environmental changes can affect histone modifications and chromatin accessibility both globally and at individual genomic loci, and these effects are largely mediated through the activity of chromatin modifiers, transcription factors, and DNA binding proteins. The downstream consequences of these effects are then reflected via the transcriptome in gene expression levels and the proteome.

Olden et al. dub this system a 'biosensor' that could potentially trace cumulative environmental exposures over the lifecourse and may be a key in better understanding of ethnic/racial health disparities [28]. Relative to DNA methylation, these epigenomic featrues can provide a more temporal link between environmental changes, genome function, and gene expression, representing a read-out for rapid changes in environment or behavioral change through the course of an intervention. Like DNA methylation, research connecting these epigenomic and transcriptomic factors with more than one domain of the model in Fig. 1 is currently sparse, especially with regard to neighborhood, socio-cultural, or higher level factors. Future research should focus on mechanistic pathways other than stress response such as obesity [29], include contexts or expand into neighborhood characteristics such as green space, social cohesion, and built food environments, as well as include ethnically diverse cohorts.

Microbiome and Nutrition

Research over the past decade has demonstrated that the human microbiome plays a key role in human health and disease [30]. Although many recent studies link microbial composition to specific phenotypes, we still lack sufficient understanding of how microbial diversity is reflected in various ethnic

populations [31] and the relative importance of lifestyle, health conditions, and diet in shaping this diversity. In relation to obesity, the clearest role the microbiome plays is through the gut and diet [32], while an emerging secondary role may be through a gut and physical activity connection [33••]. While some studies have shown dramatic changes in the gut microbiome due to dietary changes, these changes are short-term upon reversion to the original diet, which may explain why dietary intervention is not typically successful in treating obesity [34, 35]. Long-term diet, unlike short-term diet, has been shown to have an effect on the microbiome that is large compared to other factors even including antibiotic use [36], underscoring the importance of adequate nutritional environments to support long-term healthy dietary choices. Neighborhood SES was found to be associated with variability of microbiome diversity even when accounting for individual behaviors [37]; however, there has been no research of yet exploring relationships between built food environment, food related behaviors or diet, and gut microbiome. Recent research has also demonstrated an association between increased exposure to air pollution, gut microbial taxa, and fasting glucose levels in overweight and obese adolescents [38•]. Beyond the gut microbiome, the oral microbiome also holds a promising avenue for linking microbial communities to inflammatory conditions including obesity and diabetes, while also being easier to collect than gut microbial samples. Research into ethnic differences in oral microbiome are nascent, although one study has linked lower bacterial diversity in the oral microbiome to increased age as well as length of US residency and acculturation measures among recent Mexican American immigrants [39•].

Biomarkers

A primary defect associated with obesity is insulin resistance, which is increasingly being shown to be an essential biomarker for modeling the systems pathways of obesity as well as a potential intervention point and metric for intervention efficacy. Recently, a greater focus is being placed on changes that are tissue specific: for example, measures of adipokines adiponectin and leptin can be used as biomarkers of adipose tissue insulin resistance [40]. Insulin resistance results in impaired insulin action in adipose tissue and is thus strongly correlated with decreased adiponectin and increased leptin. Leptin resistance is linked to increased appetite and highly correlated with obesity [41]. However, these associations can change with alterations of diet, nutrition, and exercise; findings suggest that DNA methylation around leptin-associated loci resulting from behavioral changes constitutes a significant determinant of leptin expression [42]. Environmental and pollution-related exposures have also been shown to influence insulin resistance, pro-inflammatory immune activation, hepatic endoplasmic reticulum stress, and other metabolic-related biomarkers, likely through epigenomic processes [43••]; however, behavior as both potentially mediating and moderating of these relationships has not been examined sufficiently. Additionally, how disparities in obesity-related biomarkers are attributable to genetic, behavior, environment, nutrition domain variation is understudied. Likely, these disparities are underestimated by the traditionally narrow focus of any given study population and the lack of concurrently measured, intertwining domains.

Behaviors and Environment

Behavior is the primary linking factor between environment and biological pathways, and increasingly specific health-related behaviors are being incorporated into models of obesity disparity that seek to understand genetic, epigenetic/transcriptomic, microbiome, and biomarker processes as seen in previously cited literature. In a recent systematic review, researchers found that behaviors including physical activity, smoking, alcohol consumption, and dietary patterns contributed to the socioeconomic gradient in cardiometabolic disorder inequities with some variation in geographies, gender, and age [44]. A newly emerging area for behavioral research and obesity disparities is in circadian rhythm, specifically sleep. Multiple social and environmental predictors have been found to affect sleep including discrimination, nighttime noise, and pollution—all factors that minorities are likely to be more exposed to [45, 46••]. Hispanics/Latinos have been found to have shorter sleep and more sleep disturbed breathing than non-Hispanic Whites [47], with associations found between poorer/less sleep, insulin resistance, hypertension, and obesity [48•, 49, 50]. These poorer sleep patterns have been linked to perceived social environment/neighborhood safety and higher levels of objective measures of traffic-related air pollution [51, 52•].

There is a large body of evidence showing that exposures and features of environments influence obesity and related morbidities [53–55]. At the same time, numerous studies have demonstrated major environmental disadvantages for Hispanic/Latino neighborhoods that can be linked to obesity including disparities in air pollution, water quality, walkability, green space, crime, traffic safety, pollution, isolation, and disorder [56–59]. Of recent interest is an increase in studies measuring neighborhood-level racial discrimination or structural racism and effects on health outcomes [60]. A study looking at racial inequalities in SES including poverty, unemployment, and homeownership found that inequality was associated with higher prevalence of obesity, while inequalities in median income, college graduates, and unemployment were associated with fewer fresh food stores and more fast food outlets [61•]. Bailey et al. map out a number of possible pathways between racism and ill health including economic injustice and social deprivation, environmental and

occupational health inequities, psychosocial trauma, targeted marketing of health-harming substances, inadequate health care, state-sanctioned violence and alienation from property and traditional lands, political exclusion, maladaptive coping behaviors, and stereotype threats [62•]. There is a critical need to advance the methods by which epidemiological research can consider obesity disparities as a system in order to incorporate such understandings and measures of systemic racism [63].

Data Collection and Integration in Systems Epidemiology

In developing an effective model for a systems approach to understanding obesity disparity, there must be consideration of how domains or pieces of the system will be operationalized through data collection. Often, epidemiological studies will generate measures of health and disease at the level of the individual looking at biological mechanisms in the body, behaviors the individual undertakes, and exposures the individual is subjected to. Problems with this approach stem from how environment is defined and how we conceptualize interactions with the environment, most problematically leading to trying to fit course level concepts into an individual-level model [64]. For example, instead of building a model that includes regional supplemental nutrition policy and number of times the participant uses food stamp benefits, only nutritional metabolomics for an individual are employed in an analysis. This process has been named molecularization, or "the social processes and transformations through which phenomena (diseases, identities, pollution, food, racial/ethnic classifications) are re-defined in terms of their molecular components and described in the language of molecular biology" [65]. This is a significant problem for health disparities research where complex social phenomena and population-level experiences are often fundamentally important for situating and understanding a biological pathway [66].

One way of being more representative of behaviors within their contexts is through better methods of measurement. Definitions of behavior or environmental exposure that lack specificity and variability have been tied to contradicting or counterintuitive results, have been shown to underestimate effects, and are questionable in terms of how effectively they can measure the target association [67]. Self-reported behavioral data has been repeatedly shown to be an inaccurate and unreliable measure of health related behaviors, resulting in an inability to evaluate current and changing behaviors, effects of interventions, and relationships between behaviors and health outcomes [68]. Ethnic differences in self-reported data have not always borne out when objective measures are used. For example, self-reported leisure time physical activity shows racial disparities [69]. Accelerometers, however, have detected fewer differences, not only potentially due to self-report biases, but also due to the pre-existing cultural biases in the questions themselves. New accelerometer processing methods are providing more informative patterns of behaviors beyond total amount exercised such as bouts, or specific behaviors from machine learned models [70]. Greater precision of objective measures also means we can discover associations in smaller samples [71].

Global Positioning System (GPS) devices can obtain accurate representations of a person's movements and trajectories by recording latitude and longitude at varying time intervals (down to the second) while a person engages in their daily routine, and ascertain if an individual is indoors or outdoors [72]. Coupled with Geographic Information Systems (GIS), which represent layers of neighborhood and environmental data such as air quality, sidewalk density, poverty, and food stores, GPS is able to create representations of individuals' daily exposures to environments based on where somebody is and how much time they spend there, as compared to home or neighborhood measures [73, 74]. Accelerometry coupled with GPS and GIS data results in dynamic exposure measures that can assess where individuals are exposed, for how long, and during what behaviors [73]. Wearing such sensors can be burdensome for participants, so it is promising that large studies of daily mobility patterns have demonstrated that people are largely habitual, and 1 to 2 weeks of sensor data can account for the vast portion of types of environments that an individual is typically exposed to [75]. Furthermore, as smartphones continue to penetrate the population, eventually we will move toward monthly, yearly, and life course metrics of total exposures. Dynamic exposures that track an individual's actual movement and exposure will enhance our ability to accurately understand the associations between environments, behaviors, and epigenome/microbiome by accurately classifying people's engagement with environments rather than relying on static home/neighborhood associations. Integration of these sensors into epidemiology studies is an important step for accurately quantifying behavior and environment.

With the collection of heterogenous data types, studies that include multiple domains must decide when data integration should occur: before the modeling process, during intermediate steps, or late in the process after modeling each individual component [76]. Data integration before a modeling process will inherently need to reduce, simplify, or flatten data to achieve a uniform format across all data types. For example, physical activity data might be averaged for an individual and combined with a single alpha diversity metric of microbiome. The drawback of this approach is that the inherent richness of each data type is significantly reduced. Conversely, late-stage integration requires expert knowledge across multiple domains. There has been progress in methods for integration of heterogenous omics data [77•, 78]. For example, LUCID

Table 1 Outcomes measured in the neighborhoods to nucleotides sample ($n = 209$) across seven domains

Domain	N2N measures	Outcomes
Genomic	SNPs	Illumina Infinium CoreExome-24 BeadChip Kits (> 500,000 SNPs at single-nucleotide resolution), further SNP imputation using 1000 Genomes Database (UofMichigan). Due to small sample size of cohort, we used published GWAS studies to identify smaller subsets of SNPs for index-based risk scores (e.g., obesity risk).
Epigenomic/transcriptomic	DNA methylation	Illumina Infinium MethylationEpic BeadChip Kits (> 850,000 methylation sites at single-nucleotide resolution)
	mRNA and miRNA	Illumina Next Generation Sequencing HiSeq4000
Microbiome	Diversity	16S rRNA gene amplicon sequencing for calculating alpha diversity (diversity within each sample) such as the Shannon index
	Dissimilarity	Metrics of beta diversity such as UniFrac calculated from the 16S rRNA data
	Taxonomy	Assignment of each 16S rRNA read to a bacterial taxon, at multiple levels (typically from the phylum to the genus)
Nutrition	24 h recall	Nutrition Data Systems for Research 24 h food recalls (total calories, fats, fiber, number of food types, e.g., vegetables or fruit)
	Food frequency	VioScreen food frequency questionnaire (instances of food types eaten over past weeks/months)
Biomarker	Clinical	BMI, waist-hip circumference, medications
	Urine	Metabolomics
	Blood	Metabolomics, hormone, adipokines, biomarkers of inflammation, glycemic regulation, lipid metabolism, liver health
Behavior	Accelerometer	Physical activity (bouts, total duration, daily patterns), sedentary behavior (bouts, total duration, daily patterns), sleep time, sleep quality, machine learned behaviors (biking, running, walking, in vehicle)
	GPS	Time spent indoor/outdoor, dynamic exposure measures to all environmental features (e.g., total air pollution participant is exposed to measured by movement)
	Self-report	Questionnaires on health, physical functioning, sleep, etc.
Environment	Pollution and hazards	Air pollution, noise pollution, light pollution, water quality
	Socio-demographic	Diversity, language, crime, poverty, advantage
	Built environment	Walkability, recreation, built food environment, transit, road safety, green space
	Self-report	Feelings of safety, neighborhood cohesion, access to food environment, perceptions of walkability

estimates latent unknown clusters from diverse omics data, accounting for differential patterns across data types while jointly estimating subgroups relevant to the outcome of interest [79]. However, these methods cannot currently include behavioral data generated from sensors, daily recalls, or most biomarkers due to mismatches between data dimensions. An area of inquiry for this problem may be in computer science with statistical relational learning and other forms of artificial intelligence that deal with heterogenous data types.

The Nucleotides to Neighborhoods Study

The Nucleotides to Neighborhoods study (N2N) was a pilot study of 209 individuals living in San Diego County. Participants were a subset of a larger cohort study examining environmental effects on cancer-related biomarkers [80]. All N2N participants provided informed consent, including separate consent for DNA analyses. N2N participants were selected based on obesity status and were asked to complete additional data collection to fulfill all seven domains of the systems model outlined in Fig. 1. Participants from the larger cohort were evenly recruited from four different types of census tracts (highly walkable, low access to fast food restaurants; high walkable, high fast food; low walkable, low fast food; and low walkable, high fast food) to ensure environmental variability in home locations of the cohort. The N2N sample follows the distribution of 48, 55, 59, and 47 participants in each neighborhood type, respectively. N2N participants

are on average 60 years of age, 49% female, 34% Hispanic, 30% lean/normal, 41% overweight, and 29% obese. Figure 1 shows measures that participants were asked to contribute, which are briefly described here.

Once enrolled in the study, participants were asked to wear an accelerometer on the hip and wrist, and a GPS device on a belt with the hip accelerometer for 14 days with a minimum of 10 h of wear time per day. They were asked to complete a sleep log daily during their 2-week participation. Participants came in for a clinic visit after 1 week, before which they were asked to fast for 12 h in preparation for a 40-mL blood draw and urine sample collection. At the visit, blood pressure, height, weight, hip and waist circumference were recorded. Plasma and buffy coat from blood drawn into EDTA tubes was isolated by centrifugation at 4 °C, then aliquoted and stored at − 80 °C. A medical history form including current medication was completed, and demographic characteristics were collected via self-report survey. Additional self-report surveys assessed health conditions, depressive symptoms, quality of life, sleep quality, and neighborhood perceptions. Participants completed two 24-h dietary assessments (one weekend, one weekday). Participants separately consented to participate in the American Gut Project (AGP) [81] and were given stool collection kits with detailed instructions. Kits were returned to the AGP and participants took an online food frequency questionnaire associated with the AGP.

Table 1 illustrates the outcomes obtained from each of the domain measures for the N2N study. A significant challenge for the study was to come up with hypotheses that could leverage at least two or more domains of data. Examples of the types of questions that we are asking with the data set are (1) does GPS-based environmental exposure to air pollution correlate with alpha diversity in the gut microbiome and, if so, are there significant differences between Hispanic and non-Hispanics that can be explained by differences in exposure levels? (2) does objectively measured physical activity play a mediating role in the effect of genetic variants on insulin sensitivity? (3) do circadian sleep and eating cycles differ between Hispanic and non-Hispanic participants, and if so, are these differences associated with biomarkers like inflammation? Results from this study are forthcoming with several publications in preparation.

Several lessons were learned from the pilot N2N study. As expected, data collection for a study with this many domains is a significant challenge. Participants that were willing to undergo one aspect of data collection were sometimes less open to others. For example, some participants who gave blood during the clinical visit were not willing to have DNA analysis completed. Another issue was that data processing required six different lab teams, thus proper instrument or bio-sample handling and delivery from the collection team to each lab required significant logistical planning and quality control checks. This is an essential factor to consider when planning for budget and funding for such projects. These challenges all limited the sample size of the study, which for a traditional epidemiological study is considered small. In that regard, being a part of an already ongoing and successful cohort study was an essential aspect for successful recruitment. This model has been successfully utilized in large ethnic cohort studies like HCSC SOL and the Multi-Ethnic Study of Atherosclerosis (MESA), both of which have added domains such as objectively measured behavior, environment, and epigenetics.

Conclusions

The links between the domains discussed in this review are extensive, and research into their role in obesity, diabetes, and other complex disease has only begun to scratch the surface. Of note is the current dearth of studies that assess these relationships for specific ethnic groups, particularly when environmental disparities are so closely tied to ethnic and racial disparities. Furthermore, it is not enough to simply put race or ethnicity into modeling frameworks. A deeper understanding of why these factors are playing a role in biological mechanisms is essential, which can only be accomplished with data collection on a systems level that can account for interactions with other biological systems as well as behaviors occurring within environmental and political contexts. One of the largest gaps in current research is linkages between environmental disparities/associated behaviors (e.g., nutritional decisions made in context of food environments) and epigenomic and microbiome pathways (e.g., fiber content as important for microbiome diversity). However, progress is being made for incorporating measures that model the real-world experience of racial and ethnic minorities into epidemiological studies such as individual discrimination and structural racism. Developing research cohorts and studies that have the necessary data domains for studying these complex pathways will remain a challenge from both a funding and logistical organizational standpoint. A probable best path forward will be expanding data domains in existing cohort studies.

The specific epigenomic and microbiome factors that affect obesity in Hispanics/Latinos are still largely uncharacterized, and the potential reversibility of these changes is largely unknown. By conducting research to identify these changes, and more specifically changes impacted by behaviors and environments, we can begin to design clinical trials to assess how effectively lifestyle or environmental interventions may decrease or reverse obesity. Importantly, epigenomic and microbiome pathways that include environments and behaviors can provide compelling evidence for supporting both

individual and environmental-level strategies for reducing disparities through nutrition, sleep, physical activity, circadian rhythm alignment, land use, emissions, transportation, and other environmental public policy. A more comprehensive approach may provide evidence that environmental changes could have far-reaching and lasting impact on entire populations. Finally, a systems epidemiology approach should also aid progress in precision medicine efforts to identify individuals at heightened risk who can be targeted for intervention. Being able to identify and model these factors would give us a more accurate appraisal of individual disease risk and how to more effectively treat obesity and diabetes.

Funding Information Funding for this research was provided by a grant from the National Institutes of Health, National Cancer Institute (R01 CA179977). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The Nucleotides to Neighborhoods study was a Demonstration Project in Systems Biomedicine supported by a grant from the University of California San Diego Center for Computational Biology and Bioinformatics and San Diego Center for Systems Biology.

Compliance with Ethical Standards

Conflict of Interest Marta M. Jankowska, Kyle Gaulton, Rob Knight, Kevin Patrick, and Dorothy D. Sears each declare no potential conflicts of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. WHO (World Health Organization) (2013) WHO obesity and overweight fact sheet no 311. Obes Oveweight Fact Sheet.
2. Daviglus ML, Talavera GA, Avilés-Santa ML, et al. Prevalence of major cardiovascular risk factors and cardiovascular diseases among Hispanic/Latino individuals of diverse backgrounds in the United States. JAMA. 2012;308:1775. https://doi.org/10.1001/jama.2012.14517.
3. Hales CM, Carroll MD, Fryar CD, Ogden CL (2017) prevalence of obesity among adults and youth: United States, 2015–2016.
4. Schneiderman N, Llabre M, Cowie CC, et al. Prevalence of diabetes among Hispanics/Latinos from diverse backgrounds: the Hispanic community health study/study of Latinos (HCHS/SOL). Diabetes Care. 2014;37:2233–9. https://doi.org/10.2337/dc13-2939.
5. U.S. Centers for Disease Control and Prevention (2011) National diabetes fact sheet: national estimates and general information on diabetes and prediabetes in the United States, 2011. US Dep Heal Hum Serv Centers Dis Control Prev 3:1–12. https://doi.org/201
6. Butland B, Jebb S. Kopelman P, et al. Foresight Tackling Obesities: Future Choices. Project Report. London; 2007.
7. Glass TA, McAtee MJ. Behavioral science at the crossroads in public health: extending horizons, envisioning the future. Soc Sci Med. 2006. https://doi.org/10.1016/j.socscimed.2005.08.044.
8. Adela Hruby, PhD M, Frank B. Hu, MD, PhD M (2015) The epidemiology of obesity: a big picture. Pharmacoeconomics 33:673–689. https://doi.org/10.1007/s40273-014-0243-x.
9. Mabry PL, Kaplan RM. Systems science: a good investment for the public's health. Health Educ Behav. 2013;40:9S–12S. https://doi.org/10.1177/1090198113503469.
10. Fowler SP, Puppala S, Arya R, et al. Genetic epidemiology of cardiometabolic risk factors and their clustering patterns in Mexican American children and adolescents: the SAFARI study. Hum Genet. 2013;132:1059–71. https://doi.org/10.1007/s00439-013-1315-2.
11. Comuzzie AG, Cole SA, Laston SL, et al. Novel genetic loci identified for the pathophysiology of childhood obesity in the Hispanic population. PLoS One. 2012. https://doi.org/10.1371/journal.pone.0051954.
12. Mathers JC. Nutrigenomics in the modern era. In: Proceedings of the Nutrition Society; 2017.
13.•• Klimentidis YC, Raichlen DA, Bea J, et al. Genome-wide association study of habitual physical activity in over 377,000 UK biobank participants identifies multiple variants including CADM2 and APOE. Int J Obes. 2018;42:1161–76. https://doi.org/10.1038/s41366-018-0120-3. **First GWAS study to examine genetic heritablility of habitual exercise (measured with both self report and actigraphy).**
14.• Robinette JW, Boardman JD, Crimmins EM (2019) Differential vulnerability to neighbourhood disorder: A gene×environment interaction study. J Epidemiol Community Health 73:. https://doi.org/10.1136/jech-2018-211373. **Examines effects of genetic markers of type 2 diabetes and self-reported perceptions of environmental disorder on type 2 diabetes outcomes finding positive associations.**
15.• Le Roy CI, Beaumont M, Jackson MA, et al. Heritable components of the human fecal microbiome are associated with visceral fat. Gut Microbes. 2018;9:61–7. https://doi.org/10.1080/19490976.2017.1356556. **Builds on previous research in the TwinsUK cohort demonstrating that heritable micorbial OTUs are associated with accumulation of visceral fat phenotype.**
16. Davis JN, Lê KA, Walker RW, et al. Increased hepatic fat in overweight Hispanic youth influenced by interaction between genetic variation in PNPLA3 and high dietary carbohydrate and sugar consumption. Am J Clin Nutr. 2010. https://doi.org/10.3945/ajcn.2010.30185.
17.•• Moon JY, Wang T, Sofer T, et al. Objectively measured physical activity, sedentary behavior, and genetic predisposition to obesity in U.S. Hispanics/Latinos: results from the hispanic community health study/study of Latinos (HCHS/SOL). Diabetes. 2017. https://doi.org/10.2337/db17-0573. **First study to examine interactions between accelerometer measured physical activity/sednetary behavior and genetic variants on obesity in a large Hispanic/Latino cohort.**
18. Conomos MP, Laurie CA, Stilp AM, et al. Genetic diversity and association studies in US Hispanic/Latino. Populations: Applications in the Hispanic Community Health Study/Study of Latinos. Am J Hum Genet; 2016. https://doi.org/10.1016/j.ajhg.2015.12.001.
19. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16:6–21. https://doi.org/10.1101/gad.947102.
20. Muka T, Nano J, Voortman T, et al. The role of global and regional DNA methylation and histone modifications in glycemic traits and type 2 diabetes: a systematic review. Nutr Metab Cardiovasc Dis. 2016;26:553–66. https://doi.org/10.1016/j.numecd.2016.04.002.
21. Van Dijk SJ, Molloy PL, Varinli H, et al. Epigenetics and human obesity. Int J Obes. 2015.

22. Mamtani M, Kulkarni H, Dyer TD, et al. Genome- and epigenome-wide association study of hypertriglyceridemic waist in Mexican American families. Clin Epigenetics. 2016. https://doi.org/10.1186/s13148-016-0173-x.
23. Kulkarni H, Kos MZ, Neary J, et al. Novel epigenetic determinants of type 2 diabetes in Mexican-American families. Hum Mol Genet. 2015. https://doi.org/10.1093/hmg/ddv232.
24. Carless MA, Kulkarni H, Kos MZ, et al. Genetic effects on DNA methylation and its potential relevance for obesity in Mexican Americans. PLoS One. 2013. https://doi.org/10.1371/journal.pone.0073950.
25. Alegría-Torres JA, Baccarelli A, Bollati V. Epigenetics and lifestyle. Epigenomics. 2011;3:267–77. https://doi.org/10.2217/epi.11.22.
26.• Santos HP, Nephew BC, Bhattacharya A, et al. Discrimination exposure and DNA methylation of stress-related genes in Latina mothers. Psychoneuroendocrinology. 2018;98:131–8. https://doi.org/10.1016/j.psyneuen.2018.08.014. **Study considers percieved descrimination and its association with DNA methylation over time in a Hispanic/Latino cohort.**
27.•• Smith JA, Zhao W, Wang X, et al. Neighborhood characteristics influence DNA methylation of genes involved in stress response and inflammation: The Multi-Ethnic Study of Atherosclerosis. Epigenetics. 2017. https://doi.org/10.1080/15592294.2017.1341026. **An excellent example of a study that integrates neighborhood, epigenomics, and biomarker outcomes to understand health disparities. The study considers several components of neighborhood context and finds several influence DNA methylatoin on stress and inflammation-related genes after accounting for individual covariates.**
28. Olden K, Lin YS, Gruber D, Sonawane B. Epigenome: biosensor of cumulative exposure to chemical and nonchemical stressors related to environmental justice. Am. J: Public Health; 2014.
29. Giurgescu C, Nowak AL. Gillespie S, et al. Neighborhood Environment and DNA Methylation: Implications for Cardiovascular Disease Risk. J. Urban Heal; 2019.
30. Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. Nat Rev Genet. 2012. https://doi.org/10.1038/nrg3182.
31. Fortenberry JD. The uses of race and ethnicity in human microbiome research. Trends Microbiol. 2013;21:165–6.
32. Castaner O, Goday A, Park YM, et al. The gut microbiome profile in obesity: a systematic review. Int J Endocrinol. 2018. https://doi.org/10.1155/2018/4095789.
33.•• Mitchell CM, Davy BM, Hulver MW, et al. Does exercise Alter gut microbial composition? A systematic review. Med Sci Sports Exerc. 2019. **A first review of interplay between gut microbiome and physical activity finding that results are currently mixed partially due to lack of consistency in physical activity measurement methods.**
34. David LA, Maurice CF, Carmody RN, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature. 2013;505:559–63. https://doi.org/10.1038/nature12820.
35. Voreades N, Kozil A, Weir TL. Diet and the development of the human intestinal microbiome. Front Microbiol. 2014;5. https://doi.org/10.3389/fmicb.2014.00494.
36. Xu Z, Knight R. Dietary effects on human gut microbiome diversity. Br J Nutr. 2014;113(Suppl):1–5. https://doi.org/10.1017/S0007114514004127.
37. Miller GE, Engen PA, Gillevet PM, et al. Lower neighborhood socioeconomic status associated with reduced diversity of the colonic microbiota in healthy adults. PLoS One. 2016. https://doi.org/10.1371/journal.pone.0148952.
38.• Alderete TL, Jones RB, Chen Z, et al. Exposure to traffic-related air pollution and the composition of the gut microbiota in overweight and obese adolescents. Environ Res. 2018. https://doi.org/10.1016/j.envres.2017.11.046. **First paper to show how air pollution may be influencing obesity in adolescents through a gut microbiome mechanism.**
39.• Hoffman KL, Hutchinson DS, Fowler J, et al. Oral microbiota reveals signs of acculturation in Mexican American women. PLoS One. 2018. https://doi.org/10.1371/journal.pone.0194100. **Novel approach for understanding how acculturation may be influencing health by assessing oral microbial diversity.**
40. Chen M-W, Ye S, Zhao L-L, et al. Association of plasma total and high-molecular-weight adiponectin with risk of colorectal cancer: an observational study in Chinese male. Med Oncol. 2012;29:1–7. https://doi.org/10.1007/s12032-012-0280-2.
41. Dash S. Causes of severe obesity: genes to environment. In: Sockalingam S, Hawa R, editors. Psychiatric Care in Severe Obesity. Cham: Springer; 2017. p. 21–36.
42. Martinez JA, Milagro FI, Claycombe KJ, Schalinske KL. Epigenetics in adipose tissue, obesity, weight loss, and diabetes. Adv Nutr An Int Rev J. 2014;5:71–81. https://doi.org/10.3945/an.113.004705.
43.•• Dang J, Yang M, Zhang X, et al (2018) Associations of Exposure to Air Pollution with Insulin Resistance: A Systematic Review and Meta-Analysis. Int J Environ Res Public Health 15:. https://doi.org/10.3390/ijerph15112593. **Excellent review of current research linking air pollution to insulin resistance.**
44. Petrovic D, de Mestral C, Bochud M, et al. The contribution of health behaviors to socioeconomic inequalities in health: a systematic review. Prev Med (Baltim). 2018;113:15–31. https://doi.org/10.1016/j.ypmed.2018.05.003.
45. Slopen N, Lewis TT, Williams DR. Discrimination and sleep: a systematic review. Sleep Med. 2016.
46.•• Jackson CL (2017) Determinants of racial/ethnic disparities in disordered sleep and obesity. Sleep heal. https://doi.org/10.1016/j.sleh.2017.08.001. **Thoughtful review and framework for undersatnding how racial and ethnic disparities in sleep are influencing obesity, associated mechanisms, and enironmental causes.**
47. Chen X, Wang R, Zee P, et al. Racial/ethnic differences in sleep disturbances: the multi-ethnic study of atherosclerosis (MESA). Sleep. 2015;38:877–88. https://doi.org/10.5665/sleep.4732.
48.• Knutson KL, Wu D, Patel SR, et al. Association between sleep timing, obesity, diabetes: the hispanic community health study/study of latinos (hchs/sol) cohort study. Sleep. 2017. https://doi.org/10.1093/sleep/zsx014. **One of the first larger studies to utilize accelerometer measured sleep and relate both sleep disturbances and length of sleep to obesity and diabetes in Hispanic/Latinos.**
49. Ramos AR, Weng J, Wallace DM, et al. Sleep patterns and hypertension using Actigraphy in the Hispanic community health study/study of Latinos. Chest. 2018. https://doi.org/10.1016/j.chest.2017.09.028.
50. Loredo SJ, Weng HJ, Ramos AR, et al. Sleep patterns and obesity: Hispanic community health study/study of Latinos Sueño Ancillar study. Chest. 2019;156:348–56. https://doi.org/10.1016/j.chest.2018.12.004.
51. Billings ME, Gold DR, Leary PJ, et al. Relationship of air pollution to sleep disruption: the multi-ethnic study of atherosclerosis (MESA) sleep and MESA-air studies. Am J Respir Crit Care Med. 2017;195:A2930.
52.• Simonelli G, Dudley KA, Weng J, et al. Neighborhood Factors as Predictors of Poor Sleep in the Sueño Ancillary Study of the Hispanic Community Health Study/Study of Latinos. Sleep. 2017;40. https://doi.org/10.1093/sleep/zsw025. **This study extends literature showing negative health effects of adverse neighborhood factors and finds that percieved safety, violence and noise had impacts on length and quality of sleep in a cohort of Hispanic/Latinos.**

53. Leal C, Chaix B. The influence of geographic life environments on cardiometabolic risk factors: a systematic review, a methodological assessment and a research agenda. Obes Rev. 2011;12:217–30. https://doi.org/10.1111/j.1467-789X.2010.00726.x.
54. Sallis JF, Floyd MF, Rodriguez DA, Saelens BE. The role of built environments in physical activity, obesity, and CVD. Circulation. 2012;125:729–37. https://doi.org/10.1161/CIRCULATIONAHA.110.969022.
55. Feng J, Glass TA, Curriero FC, et al. The built environment and obesity: a systematic review of the epidemiologic evidence. Health Place. 2010;16:175–90. https://doi.org/10.1016/j.healthplace.2009.09.008.
56. Lovasi GS, Hutson MA, Guerra M, Neckerman KM. Built environments and obesity in disadvantaged populations. Epidemiol Rev. 2009;31:7–20. https://doi.org/10.1093/epirev/mxp005.
57. Piccolo RS, Duncan DT, Pearce N, McKinlay JB. The role of neighborhood characteristics in racial/ethnic disparities in type 2 diabetes: results from the Boston area community health (BACH) survey. Soc Sci Med. 2015;130:79–90. https://doi.org/10.1016/j.socscimed.2015.01.041.
58. Wen M, Maloney TN. Latino residential isolation and the risk of obesity in Utah: the role of neighborhood socioeconomic, built-environmental, and subcultural context. J Immigr Minor Health. 2011;13:1134–41. https://doi.org/10.1007/s10903-011-9439-8.
59. Fields R, Kaczynski A, Bopp M, Fallon E. Built environment associations with health behaviors among Hispanics. J Phys Act Health. 2013;10:355–42.
60. Paradies Y, Ben J, Denson N, et al. Racism as a determinant of health: a systematic review and meta-analysis. PLoS One. 2015. https://doi.org/10.1371/journal.pone.0138511.
61.• Bell CN, Kerr J, Young JL. Associations between obesity, obesogenic environments, and structural racism vary by county-level racial composition. Int J Environ Res Public Health. 2019. https://doi.org/10.3390/ijerph16050861. **One of the first studies to implement a county level measure of racial inequality by SES level across the United States to find that inequality was associated with obesity and obesogenic environments.**
62.• Bailey ZD, Krieger N, Agénor M, et al. Structural racism and health inequities in the USA: evidence and interventions. Lancet. 2017. **An important piece that lays out various ways that structural racism impacts health inequalities, but also ways to assess and measure strucutral racism in epidemiological studies and interventions.**
63. Castle B, Wendel M, Kerr J, et al. Public Health's approach to systemic racism: a systematic literature review. Disparities: J. Racial Ethn. Heal; 2019.
64. Müller R, Hanson C, Hanson M, et al. The biosocial genome? EMBO Rep. 2017;18. https://doi.org/10.15252/embr.201744953.
65. Darling KW, Ackerman SL, Hiatt RH, et al. Enacting the molecular imperative: how gene-environment interaction research links bodies and environments in the post-genomic age. Soc Sci Med. 2016;155:51–60. https://doi.org/10.1016/j.socscimed.2016.03.007.
66. Senier L, Brown P, Shostak S, Hanna B. The socio-exposome: advancing exposure science and environmental justice in a postgenomic era. Environ Sociol. 2017;3. https://doi.org/10.1080/23251042.2016.1220848.
67. Liu C, Maity A, Lin X, et al. Design and analysis issues in gene and environment studies. Environ Health. 2012;11:93. https://doi.org/10.1186/1476-069X-11-93.
68. Kerr J, Patterson RE, Ellis K, et al. Objective assessment of physical activity: classifiers for public health. Med Sci Sports Exerc. 2016. https://doi.org/10.1249/MSS.0000000000000841.
69. Troiano RP, Berrigan D, Dodd KW, et al. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc. 2008;40:181–8. https://doi.org/10.1249/mss.0b013e31815a51b3.
70. Ellis K, Kerr J, Godbole S, et al. Hip and wrist accelerometer algorithms for free-living behavior classification objective measurement of physical activity. Med Sci Sports Exerc. 2016;48:933–40. https://doi.org/10.1249/MSS.0000000000000840.
71. Dodge HH, Zhu J, Mattek NC, et al. Use of high-frequency in-home monitoring data may reduce sample sizes needed in clinical trials. PLoS One. 2015:10. https://doi.org/10.1371/journal.pone.0138095.
72. Krenn PJ, Titze S, Oja P, et al. Use of global positioning systems to study physical activity and the environment: a systematic review. Am J Prev Med. 2011;41:508–15. https://doi.org/10.1016/j.amepre.2011.06.046.
73. Jankowska MM, Schipperijn J, Kerr J. A framework for using GPS data in physical activity and sedentary behavior studies. Exerc Sport Sci Rev. 2015;43:48–56.
74. Berrigan D, Hipp A, Hurvitz PM, et al. Geospatial and contextual approaches to energy balance and health. Ann GIS. 2015;21:157–68. https://doi.org/10.1080/19475683.2015.1019925.
75. Rainham D, McDowell I, Krewski D, Sawada M. Conceptualizing the healthscape: contributions of time geography, location technologies and spatial ecology to place and health research. Soc Sci Med. 2010;70:668–76. https://doi.org/10.1016/j.socscimed.2009.10.035.
76. Kim D, Joung JG, Sohn KA, et al. Knowledge boosting: a graph-based integration approach with multi-omics data and genomic knowledge for cancer clinical outcome prediction. J Am Med Inform Assoc. 2015. https://doi.org/10.1136/amiajnl-2013-002481.
77.• Huang S, Chaudhary K, Garmire LX. More is better: recent progress in multi-omics data integration methods. Front Genet. 2017. **A good review of varoius methods for heterogenous data integration methods in the omics sciences.**
78. Pastrello C, Pasini E, Kotlyar M, et al. Integration, visualization and analysis of human interactome. Biochem Biophys Res Commun. 2014.
79. Peng C, Wang J, Asante I, et al. A latent unknown clustering integrating multi-Omics data (LUCID) with phenotypic traits. Bioinformatics. 2019. https://doi.org/10.1093/bioinformatics/btz667.
80. Jankowska MM, Sears DD, Natarajan L, et al. Protocol for a cross sectional study of cancer risk, environmental exposures and lifestyle behaviors in a diverse community sample: the Community of Mine study. BMC Public Health. 2019;19. https://doi.org/10.1186/s12889-019-6501-2.
81. McDonald D, Hyde E, Debelius JW, et al American Gut: an Open Platform for Citizen-Science Microbiome Research. Science (80-).