

RESEARCH NOTE

Open Access



Covariate dependent Markov chains constructed with gradient boost modeling can effectively generate long-term predictions of obesity trends

Alexander A. Huang^{1,2*} and Samuel Y. Huang¹

Abstract

Importance The prevalence of obesity among United States adults has increased from 30.5% in 1999 to 41.9% in 2020. However, despite the recognition of long-term weight gain as an important public health issue, there is a paucity of studies studying the long-term weight gain and building models for long-term projection.

Methods A retrospective, cross-sectional cohort study using the publicly available National Health and Nutrition Examination Survey (NHANES 2017–2020) was conducted in patients who completed the weight questionnaire and had accurate data for both weight at time of survey and weight ten years ago. Multistate gradient boost modeling classifiers were used to generate covariate dependent transition matrices and Markov chains were utilized for multistate modeling.

Results Of the 6146 patients that met the inclusion criteria, 3024 (49%) of patients were male and 3122 (51%) of patients were female. There were 2252 (37%) White patients, 1257 (20%) Hispanic patients, 1636 (37%) Black patients, and 739 (12%) Asian patients. The average BMI was 30.16 (SD = 7.15), the average weight was 83.67 kilos (SD = 22.04), and the average weight change was a 3.27 kg (SD = 14.97) increase in body weight (Fig. 1). A total of 2411 (39%) patients lost weight, and 3735 (61%) patients gained weight (Table 1). We observed that 87 (1%) of patients were underweight (BMI < 18.5), 2058 (33%) were normal weight ($18.5 \leq \text{BMI} < 25$), 1376 (22%) were overweight ($25 \leq \text{BMI} < 30$) and 2625 (43%) were obese (BMI > 30).

From analysis of the transitions between normal/underweight, overweight, and obese, we observed that after 10 years, of the patients who were underweight, 65% stayed underweight, 32% became normal weight, 2% became overweight, and 2% became obese. After 10 years, of the patients who were normal weight, 3% became underweight, 78% stayed normal weight, 17% became overweight, and 2% became obese. Of the patients who were overweight, 71% stayed overweight, 0% became underweight, 14% became normal weight, and 15% became obese. Of the patients who were obese, 84% stayed obese, 0% became underweight, 1% became normal weight, and 14% became overweight.

Conclusions United States adults are at risk of transitioning from normal weight to becoming overweight or obese. Covariate dependent Markov chains constructed with gradient boost modeling can effectively generate long-term predictions.

*Correspondence:

Alexander A. Huang
alexander.huang@northwestern.edu

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Machine-learning, Markov chains, Gradient boost modeling, Statistics, Mathematical modeling, Predictive modeling

Introduction

The prevalence of obesity among United States adults has increased from 30.5% in 1999 to 41.9% in 2020 [1–6]. Thus, new methods for modeling this increase in obesity overtime is necessary to best combat this public health emergency [7–10]. Multiple studies have identified major risk factors for obesity, including demographic factors (race, sex, age), lifestyle factors (exercise, sleep), and clinical comorbidities (hypertension, diabetes). Additionally, multiple studies have demonstrated the high prevalence of weight gain in the United States population and detailed the significant medical consequences of obesity related diseases such as coronary artery disease and type II diabetes [11–17]. Comorbidities such as cardiometabolic health and increased modeling of long-term diseases are of great focus in the building of models for obesity to guide medical care and limit the negative impacts of the obesity epidemic [18–21]. Despite the recognition of long-term weight gain as an important public health issue, there is a paucity of studies studying the long-term weight gain and building models for long-term projection [18, 20].

Markov chains are a very popular methodology to model change overtime, in which a transition probability is estimated, and the chain is executed overtime to generate long-term predictions [22–26]. Additionally, machine-learning techniques have been utilized to enhance predictive accuracy beyond that of parametric methods such as logistic and linear regression [27, 28]. Furthermore, very few studies utilize multiple states as their outcome, often choosing to rely on binary outcomes (obesity, not obesity), losing significant information in the process [29]. Development of a methodology to combine Markov chains, machine-learning techniques, and multi-state modeling will not only allow for accurate modeling of patient weight overtime, but also provide a new methodology that can be utilized to widely model any multi-state outcome, leading to broad scientific benefit [18, 20].

This study aims to provide a develop a methodology that utilizes multi-class machine learning classifiers with Markov chains to model obesity rates overtime in United States Adults within the National Health and Nutrition Examination Survey (NHANES) cohort [30–32].

Methods

A retrospective, cross-sectional cohort study using the publicly available National Health and Nutrition Examination Survey (NHANES 2017–2020) was conducted in patients who completed the weight questionnaire and

had accurate data for both weight at time of survey and weight ten years ago. The full dataset can be found on the cdc website (<https://wwwn.cdc.gov/nchs/nhanes/continuoushanes/default.aspx?Cycle=2017-2020>) and a full detailed description of the questionnaire can be found.

Ethics approval and consent to participate

The acquisition and analysis of the data within this study was approved by the National Center for Health Statistics Ethics Review Board.

Dataset and cohort selection

The National Health and Nutrition Examination Survey (NHANES) is a program designed by the National Center for Health Statistics (NCHS), which has been leveraged to assess the health and nutritional status of the United States population. The NHANES dataset is a series of cross-sectional, complex, multi-stage surveys conducted by the Centers for Disease Control and Prevention (CDC) on a nationally representative cohort of the United States population to provide health, nutritional, and physical activity data. In the present study, we analyzed adult (≥ 18 years old) patients in the NHANES dataset if they completed the weight questionnaire, leading to the inclusion of 6146 total patients.

Assessment of long-term weight change

From the questionnaire dataset in NHANES, the weight 10 years prior to the study and the current weight of the patient were extracted. The difference between the patient's weight and the weight 10 years ago was the metric used for Long-term Weight Change within this study. Weight categories defined by the Centers of Disease Control and Prevention (CDC) were utilized in this study: Patients who had a BMI less than 18.5 were considered underweight, patients with BMI between 18.5 and 25 were considered of normal weight, patients with BMI between 25 and 30 were overweight, and patients with BMI above 30 were considered obese. These are in accordance to CDC categories and definitions did not vary by age ranges.

Model construction and statistical analysis

Overall weight change was summarized within a histogram and the state transitions were summarized within in a transition-state figure. Additionally, patients were

classified as underweight, normal weight, overweight, or obese at both the time of survey and ten years prior. A transition matrix was built based upon this cohort and Markov Chains were used to project which states (underweight, normal weight, overweight, or obese) the patient was likely to be in for the next 10 decades. Covariates were taken from NHANES dataset and focused upon race and sex. These projections were plotted visually in line-graphs. Confidence intervals for these estimates were obtained through bootstrap simulation of the transition matrices. Additionally, multistate gradient boost modeling classifiers were used to generate covariate dependent transition matrices and Markov chains were utilized for multistate modeling.

Gradient boost modeling methodology: Patients were split into four groups (underweight, normal weight,

overweight and obese). A model was built for each of these groups, and the gradient boost modeling was utilized to predict the probability that any patient in the group would transition to the other states.

Results

Of the 6146 patients that met the inclusion criteria, 3024 (49%) of patients were male and 3122 (51%) of patients were female. There were 2252 (37%) White patients, 1257 (20%) Hispanic patients, 1636 (37%) Black patients, and 739 (12%) Asian patients (Table 1). The average BMI was 30.16 (SD=7.15), the average weight was 83.67 kilos (SD=22.04), and the average weight change was a 3.27 kg (SD=14.97) increase in body weight (Fig. 1). A total of 2411 (39%) patients lost weight, and 3735 (61%) patients gained weight. We observed that 87 (1%) of patients

Table 1 Summary of patient characteristics, stratified by patients with weight loss vs no weight loss

	All patients	Lost weight	Gained or maintained weight	P-value
Total patients (N)	6146 (1)	2411 (0.39)	3735 (0.61)	N/A
Age; mean (SD)	58.39 (12.94)	62.01 (12.85)	56.05 (12.45)	P<0.001
Gender male; count (%)	3024 (0.49)	1376 (0.57)	1648 (0.44)	P<0.001
Gender female; count (%)	3122 (0.51)	1035 (0.43)	2087 (0.56)	
Race white; count (%)	2252 (0.37)	943 (0.39)	1309 (0.35)	P<0.001
Race other; count (%)	262 (0.04)	113 (0.05)	149 (0.04)	
Race hispanic; count (%)	1257 (0.2)	462 (0.19)	795 (0.21)	
Race black; count (%)	1636 (0.27)	609 (0.25)	1027 (0.27)	
Race asian; count (%)	739 (0.12)	284 (0.12)	455 (0.12)	
Income_poverty_ratio; mean (sd)	2.7 (1.63)	2.59 (1.61)	2.77 (1.64)	p=0.53
BMXWT—weight (kg); mean (SD)	83.67 (22.04)	76.89 (18.94)	88.04 (22.79)	P<0.001
weight_change; mean (SD)	3.27 (14.97)	- 9.71 (10.39)	11.64 (10.96)	P<0.001
BMXBMI—body mass index (kg/m**2); Mean (SD)	30.16 (7.15)	27.58 (5.96)	31.82 (7.36)	P<0.001
Direct HDL-cholesterol (mg/dL); Mean (SD)	53.99 (16.34)	56.22 (17.36)	52.56 (15.48)	P<0.001
LDL-cholesterol, friedewald (mg/dL); Mean (SD)	110.89 (37.05)	105.63 (36.74)	114.41 (36.84)	P<0.001
Cholesterol, refrigerated serum (mg/dL); Mean (SD)	189.64 (41.88)	184.65 (43.57)	192.86 (40.44)	P<0.001
Triglyceride (mg/dL); mean (SD)	114.94 (97.1)	105.21 (76.42)	121.43 (108.26)	P<0.001
Albumin, urine (mg/L); mean (SD)	60.16 (367.89)	82.35 (490.35)	45.95 (259.97)	P<0.001
Creatinine, urine (mg/dL); mean (SD)	121.99 (80.62)	116.61 (77.26)	125.44 (82.52)	P<0.001
Albumin creatinine ratio (mg/g); mean (SD)	61.17 (396.2)	83.55 (517.76)	46.85 (292.24)	P<0.001
Total cholesterol (mg/dL); mean (SD)	189.33 (41.85)	184.25 (43.52)	192.6 (40.41)	P<0.001
Platelet count (1000 cells/uL); mean (SD)	241.76 (65.7)	234.72 (67.21)	246.32 (64.3)	P<0.001
Insulin (pmol/L); mean (SD)	90.79 (150.55)	75.21 (129.96)	101.13 (161.99)	P<0.001
Iron frozen, serum (ug/dL); mean (SD)	85.5 (34.49)	86.85 (35.02)	84.62 (34.12)	P<0.001
Fasting glucose (mg/dL); mean (SD)	117.65 (40.76)	121.14 (46.42)	115.32 (36.34)	P<0.001
Albumin, refrigerated serum (g/dL); mean (SD)	4.03 (0.33)	4.04 (0.34)	4.02 (0.32)	P<0.001
Glucose, refrigerated serum (mg/dL); mean (SD)	106.12 (40.39)	110 (48.29)	103.61 (34.14)	P<0.001
Phosphorus (mg/dL); mean (SD)	3.55 (0.53)	3.58 (0.55)	3.52 (0.51)	P<0.001
Uric acid (mg/dL); mean (SD)	5.48 (1.48)	5.39 (1.49)	5.54 (1.48)	P<0.001
SMDANY—used any tobacco product last 5 days?; Mean (SD)	1249 (0.2)	565 (0.23)	684 (0.18)	P<0.001

were underweight (BMI < 18.5), 2058 (33%) were normal weight (18.5 ≤ BMI < 25), 1376 (22%) were overweight (25 ≤ BMI < 30) and 2625 (43%) were obese (BMI > 30).

From analysis of the transitions between normal/underweight, overweight, and obese, we observed that after 10 years, of the patients who were underweight, 65% stayed underweight, 32% became normal weight, 2% became overweight, and 2% became obese (Figs. 2, 3). Full 10-year transition trajectories present in Fig. 2. After 10 years, of the patients who were normal weight, 3% became underweight, 78% stayed normal weight, 17% became overweight, and 2% became obese. Of the patients who were overweight, 71% stayed overweight, 0% became underweight, 14% became normal weight, and 15% became obese. Of the patients who were obese, 84% stayed obese, 0% became underweight, 1% became normal weight, and 14% became overweight (Fig. 2).

Subgroup analyses allow for visualization of potential covariates. Use of Gradient boost modeling shows statistically identical distributions for the Markov projection overtime, thus validating this methodology. Furthermore, gradient boost modeling was able to be used for continuous covariates, an example was computed with a three covariate model: age, race, and gender. An example for Age=35, Race=Black, and Gender=Female was

computed, and an example for age=65, Race=Asian, and Gender=Male was computed to demonstrate the utility of the covariate dependent Markov chain in modeling change overtime.

Discussion

In this retrospective, cross sectional cohort of United States adults, we observed that patients on average gained 3.27 kg over the 10 year period. As a result of this age-related weight gain, we observed that a large number of patients who were normal weight transitioning to become overweight or obese within a decade. On subgroup analysis, we found that female patients were more likely to gain weight than male patients (Fig. 4). Furthermore, we found that compared to white patients, Black and Hispanic patients were more likely to gain weight and Asian patients were less likely to gain weight (Fig. 4). These observed differences highlights that there may be additional demographic, lifestyle, and biological risk factors for obesity that should be analyzed. Many of these risk factors have been examined in the literature, including increased intake of sugar, decreased exercise, increased sedentary activity, and decreased sleep [11, 13, 14, 16, 33].

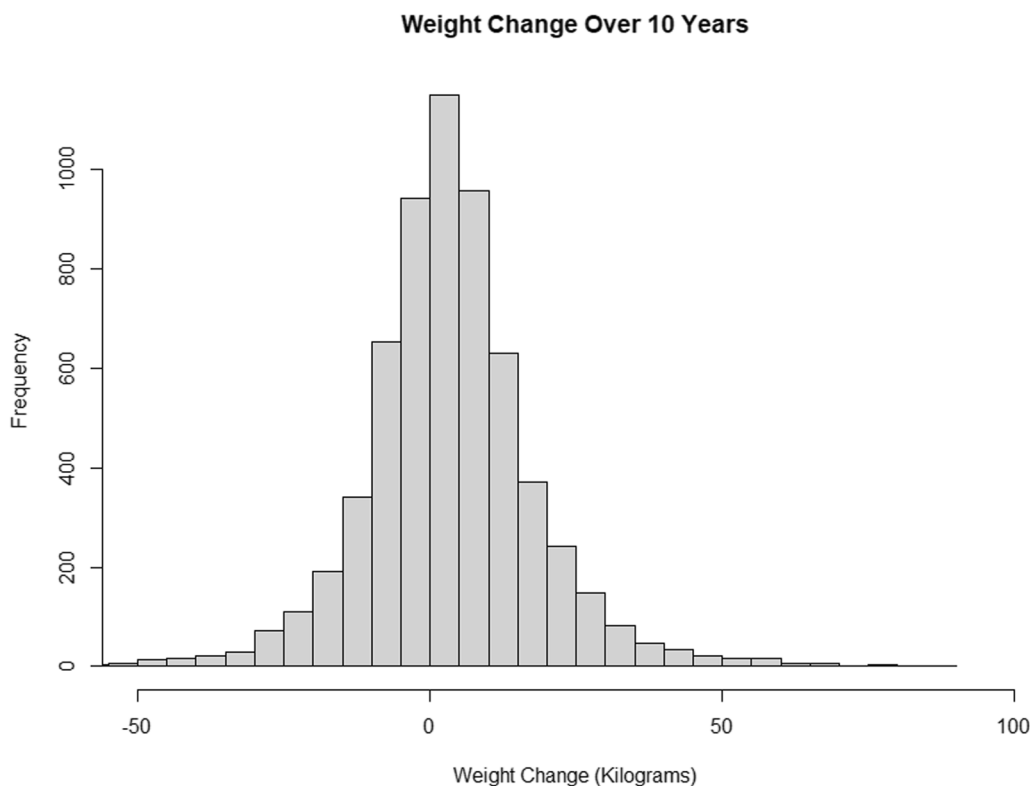


Fig. 1 Histogram of 10-year weight change

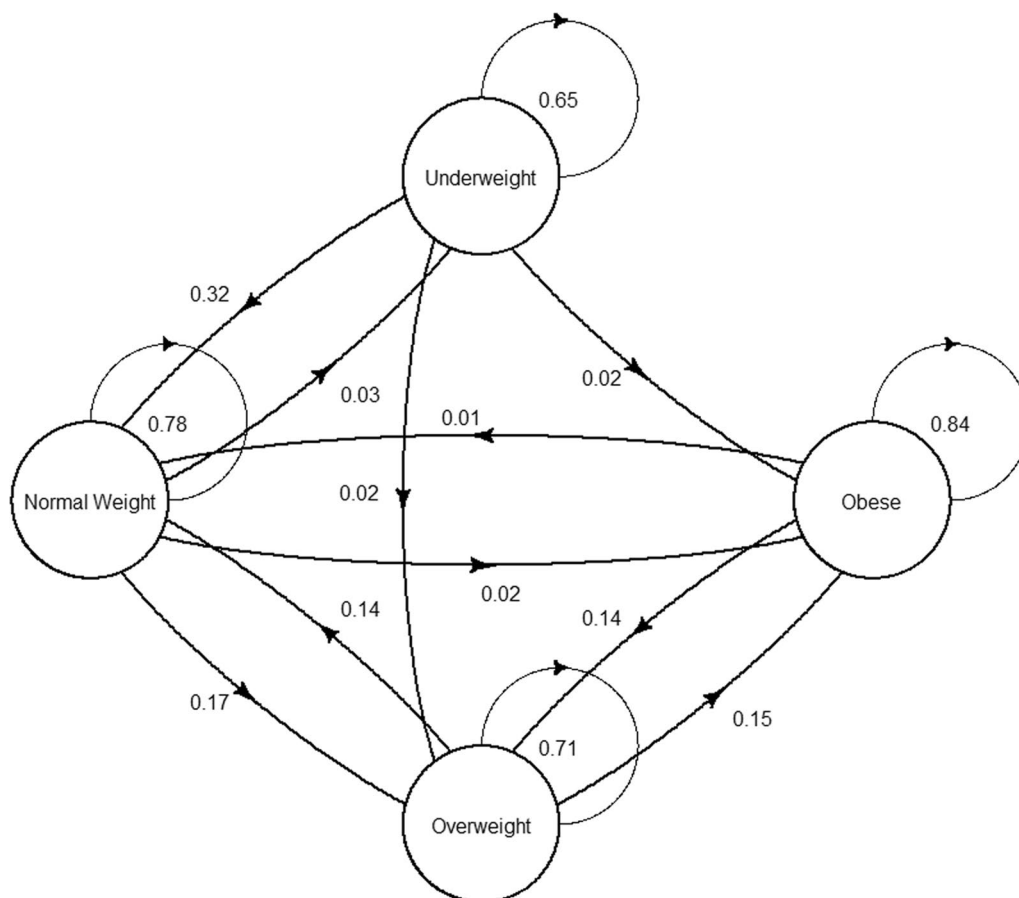


Fig. 2 Transitions between obese, overweight, normal weight, and underweight. Each circle represents a specific weight category, and arrows represent the proportion of one group transitioning to another group over the 10-year period

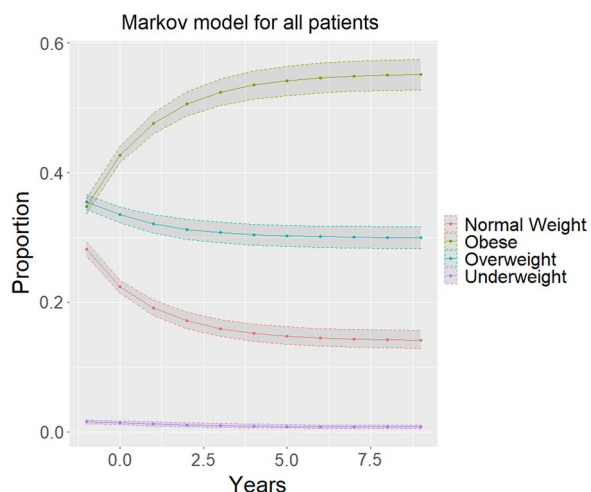


Fig. 3 Markov model of weight overtime—each new dot represents a decade. X-axis years in units of 10 years. Y-axis is the prevalence of each weight category. Time prior to zero is modeling backwards for increased context for long-term projection

Recognizing that there are numerous clinical and demographic covariates that may help model weight change overtime, a clear methodology that is able to help model weight change over time accurately is needed [3, 12]. Our methodology combines multistate modeling with Markov chains with machine-learning techniques. Multistate modeling allows us to model obesity as a set of four outcomes: underweight, normal weight, overweight, and obese. Other methods used to model obesity use logistic regression, in which the model is limited to a binary outcome, often obesity vs no obesity. By utilizing multistate modeling, a more accurate depiction of a patient’s weight can be done, and no loss of information is present. Furthermore, using multistate modeling through Markov chains allows for extrapolation into the future. Another innovation is the combination of machine learning with Markov chains for multistate modeling. Most methods that use machine learning to model outcomes often model binary outcomes (obesity vs no obesity). However, we were able to utilize multi-class classification

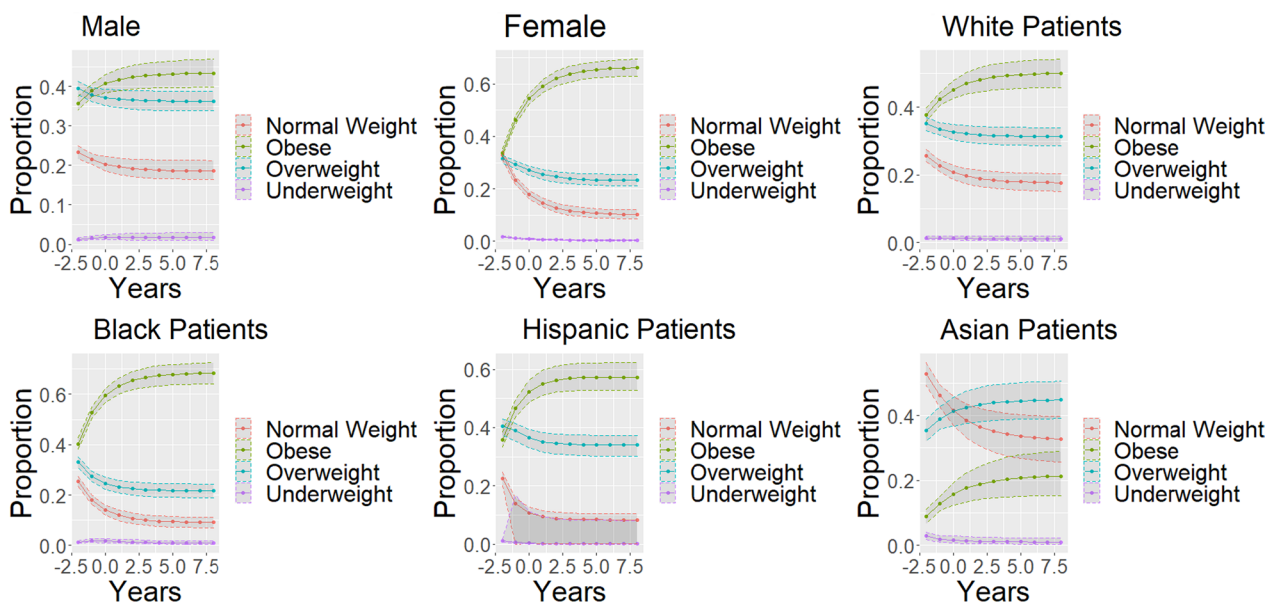


Fig. 4 Sex and weight differences in weight overtime—each new dot represents a decade. X-axis years in units of 10 years. Y-axis is the prevalence of each weight category

instead of binary classifiers and be able to have the machine learning have output of underweight, normal weight, overweight, and obese instead of traditional binary outcomes of obese vs not obese. By being able to combine the multi-class classification of machine learning models, (in our case gradient boost modeling) combined with Markov chains, we are able to have the utility for long-term projection Markov chains offer as

well as the predictive accuracy that machine learning algorithms are able to offer (Fig. 5).

What our study contributes to the literature is a novel method for modeling obesity trends over time. We are able to effectively combine two well validated techniques for predictive modeling: Markov chains and supervised machine learning (gradient boost modeling). Additionally, the novel use of bootstrap simulations allowed for

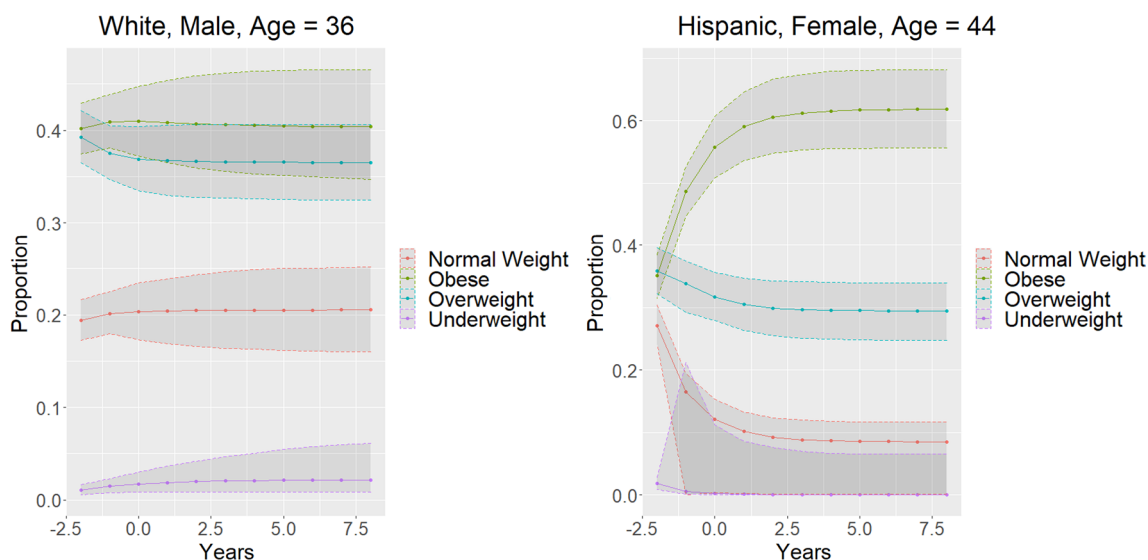


Fig. 5 Modeling of two particular patients based upon race, sex, and age—each new dot represents a decade. X-axis years in units of 10 years. Y-axis is the prevalence of each weight category

accurate quantification of the uncertainty of each of these observations that were constructed by the Markov chains [34–36]. This novel method of applying covariate-dependent transition states within Markov chains has utility beyond obesity metrics. Any multistate outcome of interest can be modeled, and the transition states being dependent on any characteristics of the patient can be completed using this methodology. Furthermore, we were able to validate the reliability of this methodology by comparing univariable models with subgroups. The near identical match of the predictions from our algorithm and the subgroups gives confidence that this methodology is able to generate unbiased estimations.

Limitations

This study has several strengths and weaknesses. We utilized the NHANES dataset, which is a retrospective cohort, carrying the limitations of retrospective studies. However, this study allows for the selection of a large cohort, evaluation of data quality, and due to the publicly available nature of the cohort, allows for increased replication and follow-up studies based upon the same cohort. Furthermore, the cohort relied on surveys to obtain the outcome of interest (weight 10 years ago) as well as the dietary and lifestyle information. More accurate measurements may have been achieved with prospective studies with lab measurements of weight, but these may interfere with natural patient habits since they know they are part of a nutrition study and may have significantly different behaviors than the general population. Additionally, self-reported survey information allows for the volume of participants to be included within this study. Another weakness was the voluntary nature of this cohort, with participants choosing to opt into the study instead of being randomly selected. This may artificially select a different cohort that may significantly differ from the population. However, our analysis found a demographically diverse population, so these results may still be generalizable to other cohorts.

Conclusion

United States adults are at risk of transitioning from normal weight to becoming overweight or obese. Covariate dependent Markov chains constructed with gradient boost modeling can effectively generate long-term predictions.

Acknowledgements

The authors acknowledge the NHANES dataset, which was used within this paper.

Author contributions

AH took part in study design, formal analysis, and writing. SH took part in study design and writing.

Funding

Authors have no funding to disclose.

Availability of data and materials

The data from this cohort can be found on the NHANES section of the CDC website. <https://www.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?cycle=2017-2020>.

Declarations

Ethics approval and consent to participate

The acquisition and analysis of the data within this study was approved by the National Center for Health Statistics Ethics Review Board.

Consent for publication

Not applicable.

Competing interests

Authors have no competing interests.

Author details

¹Cornell University, Ithaca, NY, USA. ²Northwestern University Feinberg School of Medicine, Chicago, IL, USA.

Received: 27 November 2022 Accepted: 31 October 2023

Published online: 24 November 2023

References

- Shah NS, Luncheon C, Kandula NR. Summary for patients: prevalence of obesity in Asian American Adults. *Ann Intern Med*. 2022. <https://doi.org/10.7326/P22-0018>.
- Al-Hazzaa HM, Alrasheedi AA, Alsulaimani RA, et al. Prevalence of overweight and obesity among Saudi children: a comparison of two widely used international standards and the national growth references. *Front Endocrinol (Lausanne)*. 2022;13: 954755. <https://doi.org/10.3389/fendo.2022.954755>.
- Ali N, Mahmud F, Akter SA, et al. The prevalence of general obesity, abdominal obesity, and hypertension and its related risk factors among young adult students in Bangladesh. *J Clin Hypertens (Greenwich)*. 2022. <https://doi.org/10.1111/jch.14560>.
- Ayele BA, Tiruneh SA, Ayele AA, Zemene MA, Chanie ES, Hailemeskel HS. Prevalence and determinants of overweight/obesity among under-five children in sub-Saharan Africa: a multilevel analysis. *BMC Pediatr*. 2022;22(1):585. <https://doi.org/10.1186/s12887-022-03645-z>.
- Bevan G, De Poli C, Keng MJ, Raine R. How valid are projections of the future prevalence of diabetes? Rapid reviews of prevalence-based and Markov chain models and comparisons of different models' projections for England. *BMJ Open*. 2020;10(3): e033483. <https://doi.org/10.1136/bmjopen-2019-033483>.
- Brown A, Ravi B. As the prevalence of obesity increases, rationing arthritis care is not the answer. *Osteoarthritis Cartilage*. 2022;30(9):1157–8. <https://doi.org/10.1016/j.joca.2022.06.007>.
- Canfell OJ, Littlewood R, Wright ORL, Walker JL. i-PATHWAY: development and validation of a prediction model for childhood obesity in an Australian prospective birth cohort. *J Paediatr Child Health*. 2021;57(8):1250–8. <https://doi.org/10.1111/jpc.15436>.
- Ivanescu AE, Li P, George B, et al. The importance of prediction model validation and assessment in obesity and nutrition research. *Int J Obes (Lond)*. 2016;40(6):887–94. <https://doi.org/10.1038/ijo.2015.214>.
- Seo H, Kwon CO, Park JH, et al. Dietary efficacy evaluation by applying a prediction model using clinical fecal microbiome data of colorectal disease to a controlled animal model from an obesity perspective. *Microorganisms*. 2022. <https://doi.org/10.3390/microorganisms10091833>.
- Zhou X, Lin X, Chen J, et al. Clinical spectrum transition and prediction model of nonalcoholic fatty liver disease in children with obesity. *Front Endocrinol (Lausanne)*. 2022;13: 986841. <https://doi.org/10.3389/fendo.2022.986841>.

11. Ali N, Mohanto NC, Nurunnabi SM, Haque T, Islam F. Prevalence and risk factors of general and abdominal obesity and hypertension in rural and urban residents in Bangladesh: a cross-sectional study. *BMC Public Health*. 2022;22(1):1707. <https://doi.org/10.1186/s12889-022-14087-8>.
12. Alreshidi FF, Alreshidi FS, Ahmed HG. An insight into the prevalence rates of obesity and its related risk factors among the Saudi community in the Hai'l region. *Eur Rev Med Pharmacol Sci*. 2022;26(16):5755–62. https://doi.org/10.26355/eurrev_202208_29512.
13. Barreto Silva MI, Picard K, Klein M. Sarcopenia and sarcopenic obesity in chronic kidney disease: update on prevalence, outcomes, risk factors and nutrition treatment. *Curr Opin Clin Nutr Metab Care*. 2022;25(6):371–7. <https://doi.org/10.1097/MCO.0000000000000871>.
14. Saha J, Chouhan P, Ahmed F, et al. Overweight/obesity prevalence among under-five children and risk factors in India: a cross-sectional study using the national family health survey (2015–2016). *Nutrients*. 2022. <https://doi.org/10.3390/nu14173621>.
15. Sat-Munoz D, Martinez-Herrera BE, Quiroga-Morales LA, et al. Adipocytokines and insulin resistance: their role as benign breast disease and breast cancer risk factors in a high-prevalence overweight-obesity group of women over 40 years old. *Int J Environ Res Public Health*. 2022. <https://doi.org/10.3390/ijerph19106093>.
16. Tekalegn Y, Solomon D, Sahiledengle B, et al. Prevalence of central obesity and its associated risk factors among adults in Southeast Ethiopia: a community-based cross-sectional study. *PLoS ONE*. 2022;17(8):e0265107. <https://doi.org/10.1371/journal.pone.0265107>.
17. Wan Mahmud Sabri WMN, Mohamed RZ, Yaacob NM, Hussain S. Prevalence of metabolic syndrome and its associated risk factors in pediatric obesity. *J ASEAN Fed Endocr Soc*. 2022;37(1):24–30. <https://doi.org/10.15605/jafes.037.01.05>.
18. De Pauw R, Claessens M, Gorasso V, Drieskens S, Faes C, Devleeschauwer B. Past, present, and future trends of overweight and obesity in Belgium using Bayesian age-period-cohort models. *BMC Public Health*. 2022;22(1):1309. <https://doi.org/10.1186/s12889-022-13685-w>.
19. Ward ZJ, Bleich SN, Cradock AL, et al. Projected U.S. State-level prevalence of adult obesity and severe obesity. *N Engl J Med*. 2019;381(25):2440–50. <https://doi.org/10.1056/NEJMsa1909301>.
20. Janssen F, Bardoutsos A, Vidra N. Obesity prevalence in the long-term future in 18 European countries and in the USA. *Obes Facts*. 2020;13(5):514–27. <https://doi.org/10.1159/000511023>.
21. Richardson AS, Zutshi R, Nguyen P, Tysinger B, Sturm R. Microsimulation projections of obesity interventions on cardiometabolic health disparities in the United States. *Obesity (Silver Spring)*. 2022;30(1):62–74. <https://doi.org/10.1002/oby.23297>.
22. Albers JL, Steibel JP, Klingler RH, et al. Altered larval yellow perch swimming behavior due to methylmercury and PCB126 detected using hidden markov chain models. *Environ Sci Technol*. 2022;56(6):3514–23. <https://doi.org/10.1021/acs.est.1c07505>.
23. Alberts I, Mingels C, Zacho HD, et al. Comparing the clinical performance and cost efficacy of [(68)Ga]Ga-PSMA-11 and [(18)F]PSMA-1007 in the diagnosis of recurrent prostate cancer: a Markov chain decision analysis. *Eur J Nucl Med Mol Imaging*. 2022;49(12):4252–61. <https://doi.org/10.1007/s00259-021-05620-9>.
24. Alexeev N, Isomurodov J, Sukhov V, Korotkevich G, Sergushichev A. Markov chain Monte Carlo for active module identification problem. *BMC Bioinformatics*. 2020;21(Suppl 6):261. <https://doi.org/10.1186/s12859-020-03572-9>.
25. Alyousifi Y, Ibrahim K, Kang W, Zin WZW. Modeling the spatio-temporal dynamics of air pollution index based on spatial Markov chain model. *Environ Monit Assess*. 2020;192(11):719. <https://doi.org/10.1007/s10661-020-08666-8>.
26. Du ZC, Hao YT, Wei YY, et al. Using Markov chain Monte Carlo methods to estimate the age-specific case fatality rate of COVID-19. *Zhonghua Liu Xing Bing Xue Za Zhi*. 2020;41(11):1777–81. <https://doi.org/10.3760/cma.j.cn112338-20200609-00823>.
27. Bai Y, Lu W, Li J, Chang Z, Wang H. Groundwater contamination source identification using improved differential evolution Markov chain algorithm. *Environ Sci Pollut Res Int*. 2022;29(13):19679–92. <https://doi.org/10.1007/s11356-021-17120-2>.
28. Er JC. Longitudinal projection of herd prevalence of influenza A(H1N1) pdm09 virus infection in the Norwegian pig population by discrete-time markov chain modelling. *Infect Dis Rep*. 2021;13(3):748–56. <https://doi.org/10.3390/idr13030070>.
29. Bresson G, Chaturvedi A, Rahman MA. Seemingly unrelated regression with measurement error: estimation via Markov Chain Monte Carlo and mean field variational Bayes approximation. *Int J Biostat*. 2020;17(1):75–97. <https://doi.org/10.1515/ijb-2019-0120>.
30. Odeyemi YE, Lewis O, Ngwa J, Dodd K, Gillum RF, Mehari A. Does low FEV1 in addition to fixed ratio and/or lower limit of normal of FEV1/FVC improve prediction of mortality in COPD? The NHANES-III-linked-mortality cohort. *J Natl Med Assoc*. 2019;111(1):94–100. <https://doi.org/10.1016/j.jnma.2018.06.006>.
31. Veeranna V, Zalawadiya SK, Panaich S, Patel KV, Afonso L. Comparative analysis of red cell distribution width and high sensitivity C-reactive protein for coronary heart disease mortality prediction in multi-ethnic population: findings from the 1999–2004 NHANES. *Int J Cardiol*. 2013;168(6):5156–61. <https://doi.org/10.1016/j.ijcard.2013.07.109>.
32. Williams KM, Hnatiuk O, Mitchell SA, et al. NHANES III equations enhance early detection and mortality prediction of bronchiolitis obliterans syndrome after hematopoietic SCT. *Bone Marrow Transplant*. 2014;49(4):561–6. <https://doi.org/10.1038/bmt.2013.222>.
33. Seyednasrollah F, Makela J, Pitkanen N, et al. Prediction of adulthood obesity using genetic and childhood clinical risk factors in the cardiovascular risk in Young Finns study. *Circ Cardiovasc Genet*. 2017. <https://doi.org/10.1161/CIRCGENETICS.116.001554>.
34. Austin PC. Bootstrap vs asymptotic variance estimation when using propensity score weighting with continuous and binary outcomes. *Stat Med*. 2022;41(22):4426–43. <https://doi.org/10.1002/sim.9519>.
35. Tibbe TD, Montoya AK. Correcting the bias correction for the bootstrap confidence interval in mediation analysis. *Front Psychol*. 2022;13:810258. <https://doi.org/10.3389/fpsyg.2022.810258>.
36. Xie H, Tang Q, Zhu Q. A multiplier bootstrap approach to designing robust algorithms for contextual bandits. *IEEE Trans Neural Netw Learn Syst*. 2022. <https://doi.org/10.1109/TNNLS.2022.3161806>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

