The Fast Food Phenomenon: Assessing Variations in Fast Food Consumption From 1996-2008

Sajan Patel & Max Tjen

Case Western Reserve University

PQHS 459

Dr. Yilun Sun

05 May 2023

Background

Literature Review

In the United States, individuals are surrounded by a plethora of food offerings when choosing what one wants to eat. One of these growing options is fast food, which is characterized by food that is to be ordered and paid by a customer before receiving the food.[1] Traditionally, dine-in and takeout options were the only two methods of obtaining this food. However, with the advent of mobile ordering, some fast food chains offer their own delivery service, or individuals can receive these food options through third-part food delivery platforms such as Grubhub and UberEats. This, coupled with the increased number of fast food outlets, advertising, and affordability, has led many Americans to choose to eat at fast food restaurants.[1] For these reasons, less Americans prepare food at home but rather dine out, a trend that has been documented since the 1970s.[2] According to data collected from the National Health and Nutrition Examination Survey, approximately 36.6% of adults consumed fast food in a given day from 2013-2016, with individuals age 20 to 39 years and non-Hispanic Blacks being the most frequent consumers.[3]

This pattern of food consumption may benefit the economy through the creation of jobs in the dining industry, but these decisions must be studied carefully due to the deleterious health effects that high consumption of fast food can cause. Fast food is high in sugar, salt, and fats, which can produce both short-term and long-term health effects to an individual. In the interim, fast food intake causes spikes in blood sugar, blood pressure, and inflammation.[4] However, the main concerns with eating fast food lie in the long-term harms that are associated with eating fast food with regularity. High consumption can produce irreparable effects, with risks such as obesity, insulin resistance, type 2 diabetes, and other cardiovascular conditions.[5] With low nutritional value, it appears that these food options may be doing more harm than good. Thus, understanding patterns of fast food consumption at both an individual and population level is imperative to intervene with these high-risk populations and make the necessary changes to their diet in hopes of mitigating their risks of future negative health outcomes.

Rationale

As previously mentioned, there has been a documented rise in the frequency of individuals eating fast food over the years. It is also known that a higher percentage of adult males consume fast food more often than their female counterparts (37.9% vs 35.4%).[3] However, there remains a gap in the literature regarding whether the frequency of fast consumption can be attributed to gender longitudinally. Additionally, due to the fact that patterns of adolescent behavior have strong associations with predicting their future habits, understanding adolescent fast food consumption can provide context as to whether high utilization at an earlier age predicts frequent fast food users when they become older. Previous approaches to address this gap have looked at males and females in their respective silos and fail to compare them. In one study that aimed to assess the longitudinal trends in fast food intake among adolescents, it was found that individually, males and females had an significant increase in their food consumption as they aged from early to middle adolescence (females: 15.8% to 27.3%, $p < 0.01$, males: 16.8% to 30.2%, $p < 0.01$).[6] However, this study does not focus on whether gender is attributable to this change, but rather changes in stages of adolescence. Other literature that tries to explain the differences in purchasing behaviors by gender demonstrated that habit, taste, and price were main influences that led both genders to consume fast food.[6] However, men were more likely to be influenced by offers and promotions, and did not respond as favorably to healthier options as females.[7] This study has a focus on cross-sectional data, which deviates from the aim of this study to characterize fast food consumption longitudinally.

Summary Statement

In this study, we aim to assess the variations in fast food intake in adolescents longitudinally by gender. Our objective can be summarized by the following research question: *Using The National Longitudinal Study of Adolescent to Adult Health data from 1996-2008, how have habits of eating fast food changed over time by gender? Are there any other factors associated with this change?*

Methods

Data Source

Data for this study comes from the National Longitudinal Study of Adolescent to Adult Health (Add Health). This publicly available database consists of a national representative sample of 20,000 adolescents in grades 7-12 during the 1994-1995 academic year.[8] These patients have been followed up in five points throughout the time, referred to as waves, with the most recent wave occurring in 2016-2018. Information from this database includes demographics, familial relationships, physical health, behavioral health, and an assortment of other domains. This wide capture of data from a variety of areas can allow for the analysis of a variety of phenomena longitudinally. For the purposes of this analysis, we will be looking at data from waves II-IV. Wave II data was collected from 1996, wave III in 2001-2002, and wave IV in 2008.

Study Population

The study population consists of all individuals that were followed in the Add Health database during waves II, III, and IV. Additionally, individuals must have at least one wave of information on all of our variables of interest — fast food consumption in the past seven days, gender, body weight, and smoking status — to be included in the study. There are 2,377 unique individuals that are included in the study population. With respect to the individual waves in which the data came from, there were 1,457 participants in wave II, 1,972 participants in wave III, and 2,082 participants in wave IV. This can pose a challenge when analyzing the data due to this unbalanced study design in which not all individuals participated in all three waves of interest.

Variables of Interest

Our outcome of interest is the number of times that an individual has eaten fast food in the past 7 days, found by the response to the question "How many times in the past seven days did you eat food from a fast food restaurant, such as McDonald's, Burger King, Wendy's, Arby's, Pizza Hut, Taco Bell, or Kentucky Fried Chicken or a local fast food restaurant?." In this analysis, fast food intake is treated as a binary variable. Individuals who have eaten fast food in the past 7 days 7 or more times are classified as daily consumers, while those who ate fast food less than 7 times will be categorized as non-daily consumers. The main independent variable in this study is gender, which is a binary variable — male and female. The time variable used to measure the trends in fast food consumption over time are the three waves of data — wave II, wave III, and wave IV. Other variables that are considered to see if there are any improvements in our modeling include an individual's body weight, a quantitative predictor, and a binary response to the question “Have you ever smoked cigarettes regularly — that is, at least one cigarette every day for 30 days?” — yes and no. Challenges of analyzing this data reside in the processes of data cleaning that must occur. The codebook has indicated several values that refer to missing and invalid responses that must be taken into account before analyzing the results.

Statistical Assumptions

Using a general linear mixed effects model (GLMM), the logit link function will be used to predict daily fast food consumption, as opposed to non-daily intake. As part of this modeling approach, several assumptions will be necessary. Firstly, the outcome of daily fast food consumption will be assumed to follow a Bernouli distribution. Additionally, a random intercept model was used and this random intercept is assumed to be normally distributed. All random effects are assumed to follow a multivariate normal distribution with the mean being a zero vector and a (qxq)-covariance matrix, D, that characterizes among-individual variance and correlation. As part of this modeling approach, it was also assumed that all residuals observed normality and homoscedasticity.

Analytic Approach

First, we wanted to assess the study population by gathering summary statistics by each wave of data, and then running a logistic regression model to predict daily fast food consumption using main effects from our four covariates — wave, gender, weight, and history of daily smoking. Afterwards, modeling was completed to assess our main interest of whether the patterns of change in fast food consumption are the same by gender. Before running our models, the 2,377 unique subjects were randomly split into training (70%) and testing (30%) datasets. As a result, 1,644 unique subjects were included in the training set, which all models were run on.

The first GLMM used will be a main effects model of gender and wave. This model will be compared against another GLMM that has an interaction term between gender and wave. A likelihood ratio test (LRT) was run to see if there is significance in the gender-by-time interaction term. Then, in order to understand the impact of other factors on fast food consumption, body weight and history of daily smoking were added as predictors to the better model. Then, another LRT test was conducted to see if there were improvements in predicting daily fast food intake with the inclusion of these extra covariates. Note that for all statistical analyses, conclusions were drawn at a significance level of 0.05. The best model was validated using the testing set of 713 unique patients, to understand the replicability of the model results.

Results

From the study population of 2,377 adolescents that were followed between the years 1996 to 2008 by Add Health, there was a larger participation of individuals in wave IV (Table 1). Over time, there appears to be an increase in the frequency of males surveyed. There also is a documented increase in daily fast food consumption from wave II to wave III (5.1% to 9.3%), followed by a slight decrease in the frequency by wave IV (7.7%). The percentage of participants that had smoked at least one cigarette every day for 30 days during the course of their life and the median weight of individuals increased over time.

Table 1. Summary of variables of interest for the study population (n=2377)

	Wave II (n = 1457)	**Wave III (n =1972)**	**Wave IV (n = 2082)**
Variables	**N (col %)**	**N (col %)**	**N (col %)**
Gender			
Male	656 (45.0)	942 (47.8)	1018 (48.9)
Female	801 (55.0)	1030 (52.2)	1064 (51.1)
Fast Food Consumption			
Daily	74 (5.1)	184 (9.3)	160 (7.7)
Non-Daily	1383 (94.9)	1788 (90.7)	1922 (92.3)
History of Daily Smoking			
Yes	658 (45.2)	1315 (66.7)	1429 (68.6)
No	799 (54.8)	657 (33.3)	653 (31.4)
Median Weight (in pounds) [IQR]			
	140 [123, 168]	162 [139, 190]	180 [150, 211]

The first model that we ran was a baseline normal logistic regression, where we predicted a subject being classified as a daily fast food consumer using all four covariates. The model coefficients are exponentiated in order to provide adjusted odds ratios (aOR) for the covariates (Figure 1).

Term	Estimate	95% CI	p-value
Wave III	1.721	(1.297, 2.305)	0.000
Wave IV	1.331	(0.987, 1.809)	0.063
Male	1.759	(1.421, 2.184)	0.000
Weight	1.002	(1.000, 1.004)	0.114
SmokerHx	1.332	(1.071, 1.664)	0.011

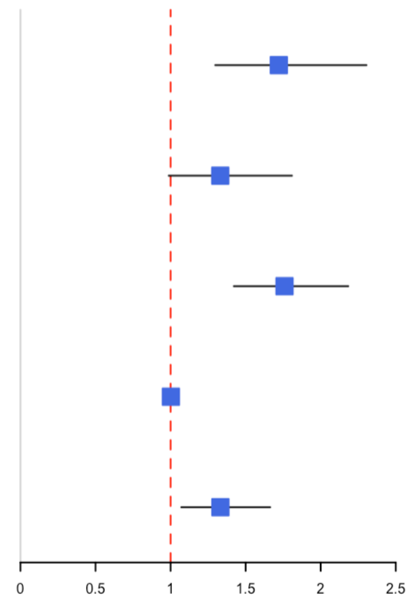

Figure 1. Adjusted-odds ratios and associated forest plot for baseline logistic regression

From this logistic regression, the odds of daily fast food consumption for an individual in wave III was 1.721 (95% CI: 1.297, 2.305), times the odds of an individual in Wave II when controlling for all other variables (p = 0.000). However, no meaningful difference in fast food consumption (aOR [95% CI]: 1.721 [0.987, 1.809]) was gathered when comparing between wave IV to wave II when holding all other variables constant. The model did predict that males had higher odds of being daily fast food users than their female counterparts (aOR [95% CI]: 1.759 [1.421, 2.184], p < 0.001) and individuals that had a history of smoking had higher odds had elevated odds (aOR [95% CI]: 1.332 [1.071, 1.664], p = 0.011). For a one pound increase in individual body weight, the model predicted there to be a 0.2% (95% CI: 0%, 0.4%) increase in the odds of being a daily fast food consumer when holding all other predictors constant (p = 0.114).

Next, two GLMMs were created, with both using wave and gender as predictor variables and one including an interaction term between them. This must be done in order to run an LRT and see if an interaction term between time and gender is significant and must be added to our models (Figure 2). These models were both run on our training dataset, which contains 1,664 unique subjects. Both of these models are random intercept models with a link logit function to use logistic regression to predict the odds of being a daily fast food consumer. The LRT is valid in this case due to the nested structure of these two models.

```
Data: train_data
Models:
glmm_1: food ~ wave + gender + (1 | id)
glmm_0: food ~ wave * gender + (1 | id)
       npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
glmm_1    5 1888.8 1920.0 -939.38   1878.8
glmm_0    7 1892.6 1936.4 -939.32   1878.6 0.1238  2       0.94
```

Figure 2. LRT to test for interaction between gender and wave

At a significance level of 0.05, we fail to reject the null hypothesis ($p = 0.94$) and conclude that there isn't a statistically significant interaction between time and gender, so we will not include it in our future models. This means that the longitudinal patterns of change in fast food consumption were not different when assessing by gender differences. We proceeded using the model without an interaction term.

A third GLMM was made using all the covariates as predictor variables. This model is our saturated model consisting of the main effects of time, gender, body weight and smoking history, which will be compared with the reduced model that only has time and gender main effects.

Term	Estimate	95% CI	p-value
Wave III	4.30	(2.58, 7.16)	0.000
Wave IV	2.53	(1.52, 4.21)	0.000
Male	2.06	(1.12, 3.77)	0.002

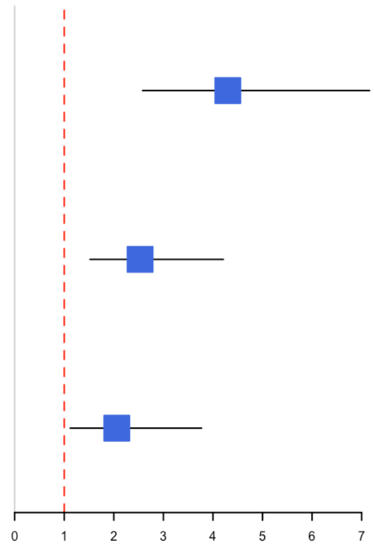

Figure 3. Adjusted-odds ratios and associated forest plot for GLMM with wave and gender main effects

For the reduced GLMM model using only wave and gender as predictors, all variables were statistically significant (Figure 3). Looking at the adjusted odds ratios when comparing wave, both wave III (aOR [95% CI]: 4.30 [2.58, 7.16], $p < 0.001$) and wave IV (aOR [95% CI: 2.53 [1.52, 4.21], $p < 0.001$) showed statistically significant differences in the odds of daily fast food intake, compared to wave II, for a typical subject in the study population. Additionally, the odds of daily fast food consumption, given that the individual is male, is elevated (aOR [95% CI]: 2.06 [1.12, 3.77], $p = 0.002$) compared to their female counterparts.

Term	Estimate	95% CI	p-value
Wave III	4.27	(2.47, 7.40)	0.000
Wave IV	2.59	(1.44, 4.66)	0.001
Male	2.22	(1.17, 4.21)	0.015
Weight	0.997	(0.991, 1.00)	0.460
SmokerHx	1.30	(0.784, 2.16)	0.308

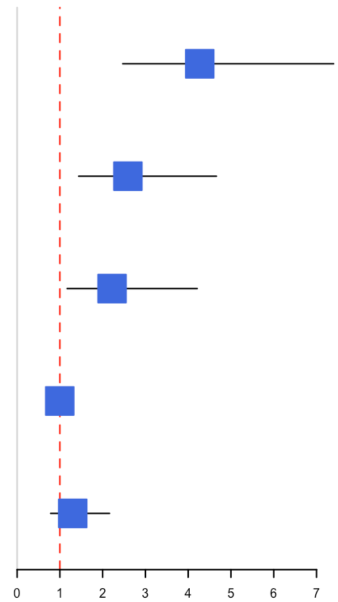

Figure 4. Adjusted-odds ratios and associated forest plot for GLMM with wave, gender, weight, and smoking history main effects

Similarly for the full GLMM model with all predictors of interest, each wave and gender was statistically significant (Figure 4). However, there were no statistically significant differences in the odds of being a daily fast food consumer when comparing by weight (aOR [95% CI]: 0.997 [0.991, 1.00], p = 0.460) and by history of daily smoking (aOR [95% CI]: 1.30 [0.784, 2.16], p = 0.308). This demonstrates that the addition of these two predictors did not meaningfully improve our modeling approach.

```
Data: train_data
Models:
glmm_1: food ~ wave + gender + (1 | id)
glmm_2: food ~ wave + gender + weight + smoke + (1 | id)
       npar    AIC    BIC  logLik deviance  Chisq Df Pr(>Chisq)
glmm_1    5 1888.8 1920.0 -939.38   1878.8
glmm_2    7 1891.1 1934.9 -938.57   1877.1 1.6365  2     0.4412
```

Figure 5. LRT to assess improvements in saturated model

To confirm this, we ran an ANOVA test to compare the models, which returned a p-value of 0.441. This demonstrates that there is not a statistically significant improvement by adding the two predictors of body weight and history of daily smoking. Thus, the reduced model including the main effects of wave and gender seems to best predict our outcome of interest.

By setting aside a testing subset of individuals in this study population, we performed validation to see if the results from our winning model are observed in new data. We ran the same reduced GLMM, only changing the subjects included in the modeling. From this, we were able to obtain a validated R^2 and validated C-statistic, and compare this to the same statistics from our initial model (Table 2).

Table 2. Comparing performance of initial model to validated model

	R^2	**C-statistic**
Initial Model	0.526	0.976
Validated Model	0.488	0.972

The model performance when applied to the testing dataset was marginally worse, as there was a small decrease in the R^2 value (0.526 vs. 0.488). However, the high C-statistic (0.976 vs. 0.972) demonstrates that both models perform very well when correctly classifying individuals as daily and non-daily fast food consumers.

Discussion

Summary of Findings

In this study, we used data from Add Health to understand if patterns of fast food consumption differed by gender over time, and if factoring in other variables such as body weight and smoking history improve the ability to predict an individual to be a daily fast food consumer. Using GLMMs with a logit link function, we conducted a likelihood ratio test to see if the gender by wave interaction was significant. From this test, we failed to reject our null hypothesis that the groups did not differ ($p = 0.94$), indicating that there is not enough evidence to say that the trends in fast food consumption differed by gender longitudinally. In an effort to see the impact of body weight and smoking status on predicting daily fast food intake, we conducted another LRT to see if these covariates improved the GLMM with only gender and wave main effects. This test also demonstrated that the addition of complexity to the model with these added variables did not meaningfully improve the model ($p = 0.44$). Further validation was run in order to assess the performance of our winning model, the one that included only the main effects of gender and wave. The C-statistic of 0.972 when cross-validating to the testing set indicates that the model does an excellent job at predicting whether an individual consumes fast food daily, on average.

Strengths

One key strength of this study includes the database selection. Add Health not only provides a wide capture of demographic and clinical information on adolescents, but they do a great job of limiting loss to follow up. This is indicated by the fact that more subjects participated in wave IV than in waves II and III, respectively. Loss to follow up is one of the primary concerns in longitudinal studies, and Add Health succeeds in mitigating the risk of this bias infiltrating the data. Additionally, the study fulfills a unique gap in existing research, as most studies assess differences in gender fast food habits cross-sectionally. This study allows for a temporal aspect in order to ascertain whether gender differences are attributable to the trends in fast food uptake.

Limitations

Some of the caveats in our statistical approach include the usage of a binary outcome for fast food consumption. Perhaps if we kept the number of times an individual has eaten fast food in the past 7 days as a count variable and used a poisson link function, more exact results could have been obtained pertaining to exactly how many times they are eating fast food. By using a logistic approach and choosing a cut point of 7 times to classify fast food consumption, there was a small percentage of patients that we categorized in the daily intake group. 5.1%, 9.3%, and 7.7% of participants in waves II, III, and IV, respectively, were classified as daily fast food consumers. This limited number of individuals that had our outcome of interest that we predicted in our logit link GLMMs forced us to only use a small amount of predictors due to the lack of degrees of freedom we could spend in our modeling approaches. Additionally, the variable selection for the added predictors may not have been appropriate. Given that this is a longitudinal study that starts their assessment in adolescence, there will naturally be an increase in one's body weight as their body matures. Body weight was included as an extra covariate in the third GLMM. From the initial summary run on each wave of data, there was a noticeable increase in the weight for each successive wave of data collected. However, this documented increase is attributable to developmental reasons and should not be used to predict fast food intake. Also, another limitation of this study is that the data is old as it uses information gathered from 1996-2008, so these results will not be able to be generalized to the current patterns and engagement with fast food intake.

Future Work

To improve the GLMM used in this study in the future, it would be interesting to see whether using a random slope and random intercept model, as opposed to solely random intercept, would improve the model's ability to predict daily fast food consumption. By incorporating different random slope structures and paying closer attention to random effects, we could better understand subject-specific variations in fast food consumption. Additionally, a sensitivity analysis could be conducted to better understand the predictive value of our modeling approach.

We would also like to improve this model by taking into account other predictors that may have influence on fast food eating behaviors. Specifically, understanding the role that social determinants of health play on fast food consumption over time could be interesting. A related research question could be to ask whether the patterns of fast food consumption have changed longitudinally with respect to socio-demographic factors such as race, geographic area, and area poverty level. Some of these factors, such as race, are already recorded by Add Health, and would not require further efforts to collect this data. Factors such as geographic area and area poverty level can be found by using participant zip codes. This information can be linked to databases that possess this zip code specific information in order to assign classifications to these individuals. For example, geographic area can be used by linking subject responses to the U.S. Department of Agriculture's Economic Research Service, which has defined areas as metropolitan and non-metropolitan using a classification scheme based on the Rural-Urban Continuum Codes (RUCC).[9] With this information, we can predict the number of times an individual ate fast food in the past 7 days using a GLMM with a poisson link function. Both main effects models with wave, race, geographic area, and area poverty level and interaction models can be compared using the same techniques used in this study with LRTs, assuming a random intercept model in order to have nested models. Despite these areas for future work, our modeling approaches utilized in this study allowed us to successfully answer our research question, in which we aimed to assess longitudinal trends in fast food intake by gender.

References

1. U.S. Department of Agriculture. (2017, December 4). *ERS's updated food environment atlas shows an increase in fast food restaurants between 2009 and 2014*. USDA ERS - ERS's Updated Food Environment Atlas Shows an Increase in Fast Food Restaurants Between 2009 and 2014. Retrieved April 29, 2023, from https://www.ers.usda.gov/amber-waves/2017/december/ers-s-updated-food-environment-atlas-shows-an-increase-in-fast-food-restaurants-between-2009-and-2014/
2. US Department of Agriculture. (2022, February 4). *Food away from home*. USDA ERS - Food Away from Home. Retrieved April 29, 2023, from https://www.ers.usda.gov/topics/food-choices-health/food-consumption-demand/food-away-from-home.aspx
3. Centers for Disease Control and Prevention. (2018, October 30). *Products - data briefs - number 320 - September 2018*. Centers for Disease Control and Prevention. Retrieved April 29, 2023, from https://www.cdc.gov/nchs/products/databriefs/db322.htm
4. MediLexicon International. (n.d.). *Fast food effects: Short-term, long-term, physical, mental, and more*. Medical News Today. Retrieved April 29, 2023, from https://www.medicalnewstoday.com/articles/324847
5. Bahadoran, Z., Mirmiran, P., & Azizi, F. (2015). Fast Food Pattern and cardiometabolic disorders: A review of current studies. *Health Promotion Perspectives*, *5*(4), 231–240. https://doi.org/10.15171/hpp.2015.028

6. Bauer, K. W., Larson, N. I., Nelson, M. C., Story, M., & Neumark-Sztainer, D. (2009). Fast food intake among adolescents: Secular and longitudinal trends from 1999 to 2004. *Preventive Medicine*, *48*(3), 284–287. https://doi.org/10.1016/j.ypmed.2008.12.021
7. Lassen, A. D., Lehmann, C., Andersen, E. W., Werther, M. N., Thorsen, A. V., Trolle, E., Gross, G., & Tetens, I. (2016). Gender differences in purchase intentions and reasons for meal selection among fast food customers – opportunities for healthier and more sustainable fast food. *Food Quality and Preference*, *47*, 123–129. https://doi.org/10.1016/j.foodqual.2015.06.011
8. *Add health (the National Longitudinal Study of adolescent to adult health)*. Carolina Population Center. (n.d.). Retrieved April 29, 2023, from https://www.cpc.unc.edu/research-themes/projects/add-health-the-national-longitudinal-study-of-adolescent-to-adult-health/
9. *Rural-urban continuum codes*. USDA ERS - Rural-Urban Continuum Codes. (n.d.). Retrieved April 30, 2023, from https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx