M³ Challenge Third Place Team
Cum Laude Team Prize: $10,000

NC School of Science and Mathematics - Team #3379,
Durham, North Carolina
Coach: Daniel Teague
Students: Allan Jiang, Pranay Orugunta, Anshul Subramanya, Margaret Tian, Christopher Yuan

Problem: Lunch Crunch: Can Nutritious Be Affordable and Delicious?

****Note: This cover sheet has been added by SIAM to identify the winning team after judging was completed. Any identifying information other than team # on an M^3 Challenge submission is a rules violation.*

Lunch Crunch: Can Nutritious Be Affordable and Delicious?

Team #3379

March 8, 2014

Contents

1 Executive Summary

Childhood nutrition is crucial to the maintenance of good health and the development of healthy lifelong habits. Failing to eat healthy meals can lead to obesity, diabetes, cancer, and psychological disorders [13, 12]. This federal government has created Acts such as the Healthy, Hunger-Free Kids Act of 2010 in an attempt to increase the nutritional quality of school lunches across the United States. However, this has led to many financial problems within many school districts. Healthier foods are generally more expensive, and they drive down participation in lunch programs nationwide because students claim the new meals are not tasty enough. No wonder so many school districts' lunch programs are financially insecure. Our consulting team has been charged with mathematically analyzing the complex situation of school lunches across the United States. We will look into reconciling the preferences of the three major stakeholders in the school lunch programs: students, who are concerned about taste and quantity; school districts, who are concerned about costs; and the federal government, who is concerned about lifelong healthy eating habits.

We tackle this problem by creating a new and improved mathematical model to predict the required caloric intake of a student, computing the percentage of students that are currently nutritionally satisfied by school lunches, and creating a new lunch meal plan that is nutritionally balanced, filling, and cost-effective enough for school districts all over the country to implement.

Our team first set out to find the best way to calculate a student's caloric requirements based on a number of factors. We analyzed several well-established equations and hypothesized that these equations would actually have a great deal of error, so we set out to create a new prediction equation. With raw data from the Institute of Medicine, we used a machine-learning function to optimize energy intake based on our factors (both physical and behavioral). Our equation ran with significantly less error than the well-established equations.

Next, we aimed to calculate the percentage of students that would be satisfied with a standard school lunch (850 calories). Using statistical methods on a distribution of high school students' physical and behavioral factors, we determined that, nationwide, 60% of students would have their caloric needs met by a standard school lunch. Interestingly enough, this number varied greatly with geographic location. We hypothesize that this difference in satisfaction stems from a difference in activity and obesity levels. Furthermore, we saw an inverse correlation between socioeconomic status and obesity levels. This suggests that perhaps an increase in socioeconomic status would increase the percent of satisfied students. The USDA should note that socioeconomic factors may also need to be controlled in order for children across the nation to eat healthy.

Finally, our team created looked to create a tasty, nutritionally balanced, and inexpensive lunch meal plan. By compiling data for per ounce costs of various food groups, we determined that a \$6 per student per week budget was just sufficient to provide fully nutritional lunches based on guidelines set by the USDA. These lunches also met the caloric needs of the students. If our budget were increased to \$7, we would increase the fruit, unprocessed meat, and portion sizes of our meal plan to better appeal to students' tastes. This meal plan successfully satisfied the interests of our three stakeholders.

In this solution, we created a new model to determine required intake for a series of factors. We used this model to determine the percentage of students that have their needs currently met, and then created a new meal plan that would satisfy nutritional requirements while staying within cost. We propose that the USDA consider the new model we have created, and for them to incorporate the meal plan that our organization has created into schools on a pilot basis.

2 Introduction

2.1 Background

Nutrition and eating habits during childhood are key for the development of healthy bodies. Our world is currently facing a nutritional crisis. In the last thirty years, childhood obesity has doubled in children (7% obese in 1980 to 18% obese in 2012) and quadrupled in adolescents (5% obese in 1980 to 21% obese in 2012) [10, 20].

Childhood obesity is a problem that has significant implications in both the short-term and the long-term. 70% of obese children have risk factors for heart disease [13], and these children are more likely to develop prediabetes, a condition that leads directly to Type II diabetes [16]. Obese children also suffer psychological and emotional problems due to predatory social interactions with other children [12]. In the long run, obese children will likely become obese adults with higher risk for stroke, cancer (breast, colon, endometrium, esophagus, kidney, etc.), and arthritis [15].

Scientists have determined that childhood obesity is the result of "caloric imbalance," which generally means that the number of calories taken in is far greater than the number of calories expended during the day [11]. In addition, not all calories are made the same; calories from fresh vegetables and lean meats are nutritionally different than fast foods. Children should consume the correct number of calories while maintaining balanced consumption of food categories in order to have the least risk of becoming overweight or obese.

2.2 Problem Restatement

Since students K–12 may eat up to one third of their meals at school, it is critical for school meals to be nutritionally balanced. The Healthy, Hunger-Free Kids Act of 2010 aims to make school lunch meals healthier for children across the United States. However, there exists conflict between the three major stakeholders of this program. Students would like their meals to be tasty and filling. School districts would like to minimize cost so that schools can be more financially sustainable. The federal government would like to promote long-term healthy eating habits in all children across the United States.

The USDA has asked our consulting firm to produce a report containing our recommendations for how these three stakeholders can be reconciled. The agency would like us to consider three main concerns:

1. How many calories does a child need for lunch? We would like to take into account many different factors such as age, weight, and physical activity level to determine the adequate consumption level for any individual student.

2. What percentage of students have their needs met by the average school lunch? The guidelines given by the Healthy Hunger-Free Kids Act of 2010 are formulated for the average student. We want to determine the percent of students whose caloric needs are met if all students are fed a standard school lunch.

3. How can a healthy meal be financially feasible? We aim to optimize a nutritious and balanced lunch plan in terms of food groups for a school district that has a weekly per-student allocation of $6–7.

3 Planned Approach

Below is an overview of our planned approach toward creating solutions for the three problems given to us by the USDA.

1. A model will be developed to calculate calories required for lunch based on existing equations made by the Institute of Medicine (IOM), which takes into account a person's age, gender, size, and physical activity. We will verify these equations by creating our own using raw data and a multivariate linear regression. These equations will then be modified to incorporate additional factors of ethnicity, sleep, and breakfast.

2. We will develop a model that simulates the satisfaction of the entire US and regional populations. Satisfaction is determined by observing whether their lunch caloric intake requirement is met based on seven variables. We plan to do this by using a large sample size using distributions extracted from provided data and calculating their required caloric intake.

3. We will create a tasty and nutritionally balanced \$6 meal plan by first examining the USDA recommended portions of the different food groups. An average cost per ounce of each group will be calculated using the per unit prices of representative foods within the group, and the recommended portions will be adjusted to meet the budget. We will then examine what changes can be made to taste or nutrition with a \$7 budget.

4 You Are What You Eat

Every student requires a different daily caloric intake during lunch to maintain good health. This caloric intake depends on several physical attributes (gender, age, weight, height, and ethnicity) as well as lifestyle choices (activity level, sleep patterns, and whether they ate breakfast). We studied existing models developed by the Institute of Medicine [18] and then developed two multivariate regression models using machine-learning and the factors listed below to determine the caloric intake for lunch. We also looked to incorporate varying amounts of sleep, as well as ethnicity and whether the student ate breakfast that particular day.

4.1 Assumptions

We make the following assumptions in creating our model.

- **Children eat 3 meals per day**. This traditional meal structure makes sense given the size of a school meal. Students would have to eat at least one more meal at home in order to meet their caloric requirements. In addition, most students eat breakfast at home before school [5].

- **Each meal (breakfast, lunch, dinner) contains one third of a child's daily calorie intake**. Each person has different eating habits; some eat a large breakfast while some consume the most at dinner. Over the entire population, we assume that these patterns average out to an even distribution for all three meals. Moreover, after online research, we found no evidence to the contrary of this assumption.

- Eating breakfast has a significant impact on metabolic rate. A study done at Northumbria University shows that **people who eat breakfast consume 17% fewer calories throughout the day** [9].

- **There is no appreciable difference in metabolic rate between different ethnicities** (African Americans have a 0.97% lower metabolic rate on average according to a study in the *American Journal of Chemical Nutrition*) [21].
- The average K–12 student **requires 8 hours of sleep to be considered "well-rested."** It has been shown that children who are not well-rested require 5% more calories throughout the day [17].
- Obesity is defined as a body mass index (BMI) of **30 or higher**, which is the 95th percentile. Overweight is defined as a BMI of **25 to 30**, which correlates to 85th–95th percentile [6].

4.2 Model Design

There are many existing equations that calculate recommended caloric intake (e.g., Harris–Benedict, IOM, etc.). However, these equations only account for age, gender, weight, height, and activity level. It is clear that there are many factors besides those listed above that affect how many calories we consume. For example, it has been shown that sleeping more leads to fewer calories consumed. Our model incorporates several additional factors (amount of sleep, breakfast) to more accurately calculate calories needed in a school lunch.

4.2.1 Estimated Energy Requirements

We first observed caloric intake recommendations made by the Institute of Medicine in a 2005 study on dietary reference intakes [18]. This study created four equations, separated by gender and obesity, that take into account age, height, weight, and physical activity. The Institute of Medicine equations are more applicable than others (like Harris–Benedict) because they make specific recommendations of the number of calories required per day, while older equations calculate the basal metabolic rate and total calories per day without specific recommendation of calorie count. Furthermore, the Institute of Medicine equations are the most recent and reflect new understanding of nutritional balance as defined by the 2005 Dietary Guidelines for Americans.

The physical activity level (PAL) is the ratio of total energy expenditure to basal energy expenditure (TEE/BEE), where TEE is total energy expended per day and BEE is energy expended during rest. The physical activity level categories are defined as sedentary (PAL 1.0–1.39), low active (PAL 1.4–1.59), active (PAL 1.6–1.89), and very active (PAL 1.9–2.5).

The estimated energy requirement (EER) is given in the units of kcal/day and is equal to the sum of total energy expenditure and energy deposition. Below are the IOM's EER equations:

Boys 3–8 years old:	$EER = 88.5 - (61.9 \times Age) + PA \times (26.7 \times Wt + 903 \times Ht) + 20$
Girls 3–8 years old:	$EER = 135.3 - (30.8 \times Age) + PA \times (10.0 \times Wt + 934 \times Ht) + 20$
Boys 9–18 years old:	$EER = 88.5 - (61.9 \times Age) + PA \times (26.7 \times Wt + 903 \times Ht) + 25$
Girls 9–18 years old:	$EER = 135.3 - (30.8 \times Age) + PA \times (10.0 \times Wt + 934 \times Ht) + 25$
Adults > 19 – Men:	$EER = 662 - (9.53 \times Age) + PA \times (15.91 \times Wt + 539.6 \times Ht)$
Adults > 19 – Women:	$EER = 354 - (6.91 \times Age) + PA \times (9.36 \times Wt + 726 \times Ht)$

Overweight/Obese 13–18	
Male:	$TEE = -114 - (50.9 \times Age) + PA \times (19.5 \times Wt + 1161.4 \times Ht)$
Female:	$TEE = 389 - (41.2 \times Age) + PA \times (15 \times Wt + 701.6 \times Ht)$

- where Age is in years,

- physical activity, PA, is a unit-less coefficient defined as the ratio of total energy expenditure to basal energy expenditure,
- weight, Wt, is in kilograms, and
- height, Ht, is in meters.

Notice that in general, $EER = c_1 - (c_2 \times Age) + PA \times (c_3 \times Wt + c_4 \times Ht) + c_5$**.**

4.2.2 Improving Preexisting IOM Equations

We were interested in how the Institute of Medicine determined their equations for EER, so we examined the methodology detailed in their paper *Dietary Reference Intakes For Energy, Carbohydrate, Fiber, Fat, Fatty Acids, Cholesterol, Protein,and Amino Acids* [18]. Their study used a population of 525 subjects between the ages of 3 and 18 and collected data on sex, height, weight, BMI, BEE, TEE, and PAL of those individuals.

Looking at the equations the IOM devised to fit the data, we see that their fit hypothesis has four regression variables. Because this fit is relatively simple and thus susceptible to high bias (underfitting), we predicted that the existing function would not be able to fit the observed data with high precision. Thus, our team used a multivariate linear regression algorithm to find a degree-two polynomial that could output caloric requirements that better explained the experimental data.

Our function fits the age, weight, height, physical activity, sleep, and breakfast consumption of an individual to his/her estimated energy requirement (EER). The EER is defined as the average caloric intake needed to maintain energy balance and good health for an individual. It is essentially how many calories a person should eat in a day [8].

We created our model with raw data provided by a 2005 study conducted by the Institute of Medicine [18]. The study compared age, height, weight, BMI, and PAL to the observed BEE of an individual. The study had $n = 525$ and we used all values.

Let us put the problem in a machine-learning context. Suppose we have the data set below:

Age	$Height$	$Weight$	BMI	PAL	$TEE = EER$	
3	1.13	20.3	15.9	1.56	1684	
4	1.07	18.4	16.1	1.39	1334	
4	1.13	20	15.7	1.53	1483	
⋮	⋮	⋮	⋮	⋮	⋮	⋮
18	1.62	58.5	22.3	1.56	2304	
18	1.77	58.1	18.5	1.83	2713	
18	1.62	60.3	23	1.95	2634	

Table 1: Example rows of IOM data. We split into two data sets, one for males and one for females, as sex is a common biological covariate.

In the data given above, x is the matrix of all data to the left of the double vertical line. Each data point input $x^{(i)}$ is a 6-dimensional vector in $\mathbb{R}^5$, and its output $y^{(i)}$ is its corresponding EER. So for example, the first data point input, $x^{(1)} = [3, 1.13, 20.3, 15.9, 1.56, 1684]$ and the first output, $y^{(1)} = 1684$. To perform linear regression on this data, we may let our hypothesis fit function be

$$h_\theta(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_5 x_5, \tag{1}$$

where x is a new vector of input features, and θ is a parameter vector of weights $[\theta_0, \ldots, \theta_5]$ in $\mathbb{R}^5$.

However, we see that this hypothesis will only fit a linear combination of our input variables. In order to better fit our hypothesis to the data, let us introduce new features $x_7, x_8, \ldots, x_{14}$ such that $x_7 = x_1^2, x_8 = x_2^2, \ldots, x_{10} = x_5^2$. Let us further add the features $x_{11} = PAL \times Weight = x_3 x_5$ and $x_{12} = PAL \times Height = x_2 x_5$, as they seem to have biological significance according to the IOM.

Now, notice that we may rewrite $h_\theta(x) = \sum_{i=0}^{12} \theta_i x_i$ if we automatically set $x_0 = 1$ for each training example $x^{(i)}$. We may further simplify our notation by writing

$$h_\theta(x) = \theta^T x. \tag{2}$$

Now, given our data points, (or training data), we measure how closely our hypothesis $h_\theta(x^{(i)})$ matches its corresponding $y^{(i)}$'s using the **cost function**. In this case, we use the least squares regression, given below:

$$J(\theta) = \frac{1}{2m} \sum_{i=1}^{m} \left(h_\theta(x^{(i)}) - y^{(i)} \right)^2. \tag{3}$$

Next, we use the batch gradient descent method to find the best θ to minimize the cost function J. The algorithm runs as follows: we start with an randomly initialized value for θ. Then, in each iteration of the gradient descent algorithm, we perform the update

$$\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta). \tag{4}$$

Working out the partial derivative, we see that $\frac{\partial}{\partial \theta_j} J(\theta) = (h_\theta(x^{(i)}) - y^{(i)}) x_j$, so we repeatedly perform the update rule above until the value of J converges to a minimum. Gradient descent works rather well in the case of multivariate least-squares linear regression, as the local minimum is also the global minimum.

We wrote our algorithm to implement this approach in Octave. Figures 1 and 2 are graphs generated by Octave that show decreasing cost J over the number of updates on θ.

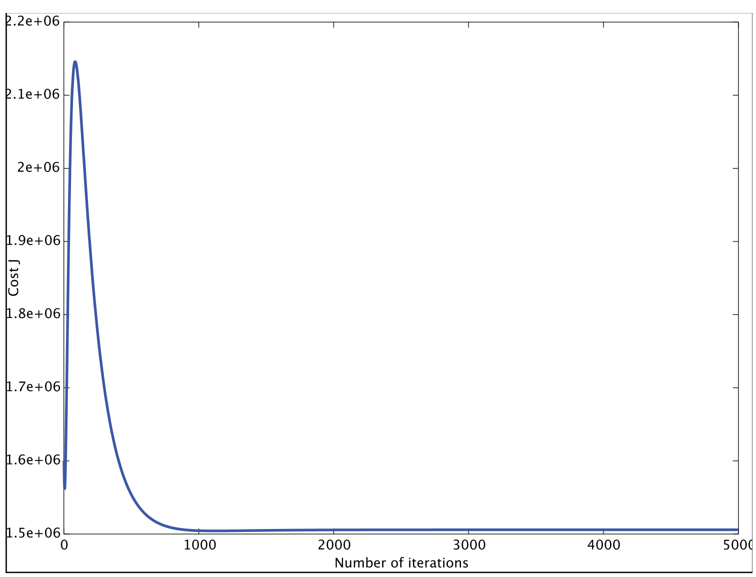

Figure 1: Optimization of regularized cost function $J_{\text{reg}}(\theta)$ over the number of iterations for males

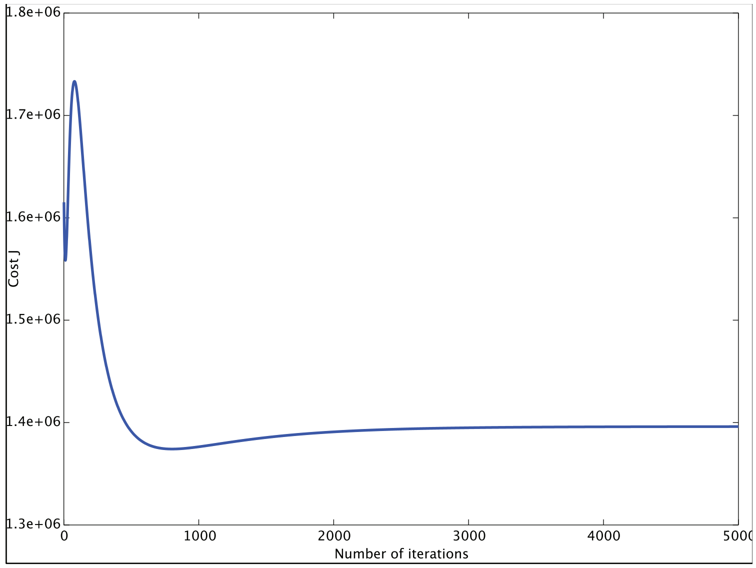

Figure 2: Optimization of regularized cost function $J_{\text{reg}}(\theta)$ over the number of iterations for females

Our machine-learning classifier's outputs for vector θ for females is

$$\theta_{\text{female}} = [2188.775, -3.157, 28.330, 29.778, 19.333, 104.274, -3.748, 29.203, 30.968, 18.994, 102.959],$$

and for males, it is

$$\theta_{\text{male}} = [2651.236, 27.161, 49.374, 50.439, 35.854, 50.892, 23.244, 49.486, 49.942, 34.732, 49.080].$$

4.2.3 Additional Factors

Now, we modify our above equations to **incorporate additional factors of sleep, ethnicity and whether the student ate breakfast.** If a student sleeps 8 or more hours, they will need 5% fewer calories the next day [17]. In addition, if a student does not eat breakfast, they will, on average, eat 17% more calories for lunch [5]. As discussed in the assumptions (Section 4.1), incorporating ethnicity would not greatly change the output of calories needed. Our model will therefore recommend the same lunch caloric intake for all ethnicities if they share the same physical and lifestyle attributes. We can therefore modify our model with a simple multiplication operation.

Thus, our predicted estimated energy requirement for a given male or female student is equal to

$$EER = \theta^T_{\text{male/female}} x \times .95^{\text{S}-1} \times 1.17^{\text{B}-1},$$

where S is 1 if the student is well rested and 0 otherwise, and B is 1 if the student had breakfast and 0 otherwise.

4.3 Model Verification

To assess the strength of our model, we compare the caloric intake equations we determined to those created by the IOM [18]. We test both our model and their model using an unregularized cost function calculation of how well the input parameter θ fits the given data x. Our results are shown in the code in Figure 3.

```
>> computeCost(X, y, theta_male_all) \%Our classifier's values of \theta for males
ans =  1.3570e+04
>> computeCost(X, y, IOM_male_9-18) \%The IOM's values of \theta for males
ans =  1.3027e+06

>> computeCost(X, y, theta_female_all) \%Our classifier's values of \theta for females
ans =  1.1311e+04
>> computeCost(X, y, IOM_female_9-18) \%The IOM's values of \theta for females
ans =  1.3568e+06
```

Figure 3: Octave code output of cost function of our hypothesis and the IOM's EER equation hypothesis.

Comparing the values of the cost function given our values of θ to the IOM's, we see that our regression is able to fit the given training data in the study with higher precision. Furthermore, our hypothesis works for all age groups from ages 3 to 18 (the ages in the data set), whereas the IOM provides different equations for different ages. Our cost is significantly lower than theirs in all cases.

4.4 Overview of Results

Our final model to determine calories needed in a school lunch takes into account a host of factors: age, gender, ethnicity, weight, height, physical activity, hours of sleep, and breakfast intake. It is shown below:

$$EER_{\text{male}} = 1691.864 \cdot Age + 32.212 \cdot Ht + 56.211 \cdot +Wt + 54.507 \cdot BMI37.117 \cdot PALo54.168 \cdot Age^2$$
$$+\ 27.695 \cdot Ht^2 + 57.508 \cdot Wt^2 + 55.046 \cdot BMI^2 + 35.508 \cdot PALo + 52.375 \cdot PA \cdot Wt + 58.604 \cdot PA \cdot Ht,$$

$$EER_{\text{female}} = 1742.444 \cdot Age + 17.232 \cdot Ht + 43.563 \cdot +Wt + 31.558 \cdot BMI16.958 \cdot PALo80.131 \cdot Age^2$$
$$+\ 1.408 \cdot Ht^2 + 40.101 \cdot Wt^2 + 23.524 \cdot BMI^2 + 14.751 \cdot PALo + 80.059 \cdot PA \cdot Wt + 56.399 \cdot PA \cdot Ht.$$

5 One Size Doesn't Necessarily Fit All

5.1 Assumptions

- Students in 9th grade are fourteen years old, in 10th grade are fifteen, and so on. We make this assumption because the data we use is given by grade level instead of age. This number represents the age a student will be when he/she enters a grade.
- Distributions within specific attributes are even. Our data is given to us in discrete levels (e.g., sedentary, somewhat active, etc.). Since we do not know the distribution within each level, we assume it to be even. We found no literature disproving this assumption.
- The standard school lunch contains around 850 calories [3]. Each day's lunch for each school district will vary, so keeping the value constant at 850 calories greatly simplifies our calculations.
- We assume that the grade distributions for high schools are even, with 25% of the student body being in each grade.

5.2 Model Design

We set out to develop a model that determines the percentage of individuals whose needs are met by the current standard lunch offered by schools. Using our team's equation that was developed in Section 4.2.2, we were able to calculate caloric requirements at lunch for any given student. We combined this equation with the distribution of attributes determined by a Centers for Disease Control study in order to determine every student's requirement [19]. Knowing every student's caloric requirement, we were able to calculate the percentage of students that would be nutritionally satisfied by the caloric content of a standard school meal.

This model was created in C++. The program iterated through several lists that each described an attribute (age, weight, etc.) with a nested if-then loop structure. The loop structure guaranteed that all combinations were accounted for, and the lists took relative frequencies into consideration (e.g., in a 100-person sample, 20 are obese and 15 are overweight). The algorithm of the program is summarized below:

- Create global variables of total individuals and calorically satisfied individuals. These variables are not yet initialized but will eventually store final values.
- Define the function from problem 1 to determine the necessary caloric intake at lunch based on all attributes in Section 4.
- Define arrays of attribute distributions, making sure that the relative frequency of attributes from the data set is preserved. Based on the percentages found within the dataset, there can be up to 100 elements in each of the attribute arrays.
- The main program runs through a eight-level nested for loop and an if-then loop for each attribute array and adds one to the calorically satisfied variable if the number of calories required for a certain individual is less than 850.
- Display the success percentage.

After performing this analysis on a national scale, we wanted to separate the data into distinct geographic regions and examine what differences, if any, would arise. We split the country into four regions and used regional data from the CDC data set as input to our program. It is important to note that data was not available for many states. We therefore did not consider these areas within our four distinct regions. In total, we were able to find data for 42 different states.

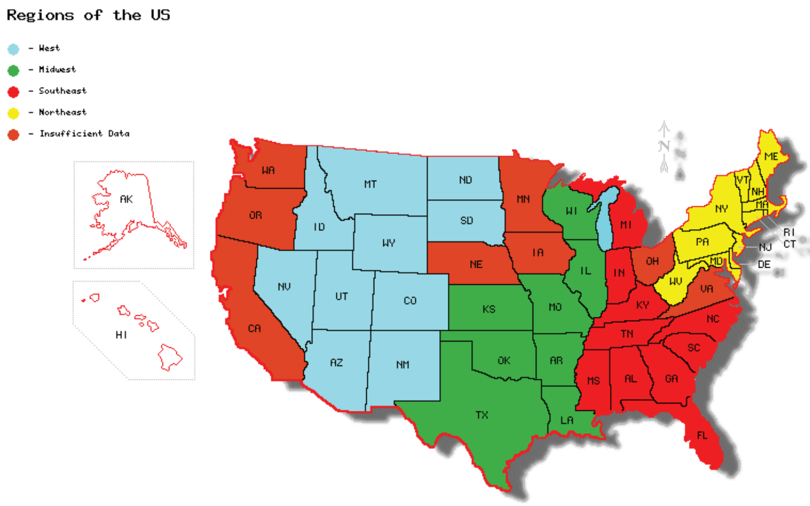

Figure 4: Our defined regions of the United States

5.3 Data Collection

Region	Percentage Satisfied
Entire US	**60.01%**
West	75.69%
Midwest	72.95%
Southeast	41.48%
Northeast	45.47%

Table 2: Individuals with satisfied caloric needs

5.4 Model Justification

This model's strength is that it analytically finds the percentage of satisfaction over a large population size. Due to the different combinations of variables involved, we develop a 20 million sample size, which allows for high representation. This model accurately incorporates the distribution of all of the eight variables into one factor, effectively reducing a function of seven dimensions into a simple inequality.

5.5 Overview of Results

From Table 2, we observe that the entire United States exhibits a total **high school student satisfaction percentage of approximately 60%**. From this statistic, we can see that many students are not receiving the number of calories they need to sustain growth. We recognize that larger lunches would incur higher costs for school districts, but the trade-off is a healthier lifestyle for students.

Percentage of Individuals with Various Traits						
Region	Highly Active	Active	Low Active	Sedentary	Obese	Overweight
West	23.89	46.81	13.53	39.66	10.4	13.02
Midwest	25.725	46.3375	15.6125	38.05	12.9375	15.2
Southeast	22.63	40.59	18.57	40.84	14.27	15.49
Northeast	27.46666667	51.7	14.575	29.725	11.83333333	14.91666667

Figure 5: Frequency of various traits

When we looked at the geographic data from Figure 5, we found several interesting results. The west and the midwest regions of the United States had far higher values for the percentage of students that were satisfied (calorie-wise) with school meals than the southeast and the northeast. We wanted to see why this was the case. In order to do this, we broke down each region into the distribution of activity and obesity/overweight. Higher values of obese and overweight students would raise the caloric requirement at lunch, as would higher values of activity. Therefore, in regions where students were more overweight or in regions where students were more active, the percentage of students satisfied by school lunches would be lower. We can see in Table 5 that in the northeast, the percentage of students that are active are indeed higher relative to the other regions. In the southeast, the percentage of overweight and obese students are noticeably higher than those in other regions of the nation. From this simple analysis, we can see that **because students in the northeast and southeast are either more obese or more active than their counterparts in other regions, they are less likely to be satisfied with school lunches.**

After we did this, we wanted to know why the southeast had such a high percentage of students that were either overweight or obese. We hypothesized that obesity was influenced by the socio-economic environment of the region. In order to see if this was the case, we used US Census data to find the average household income for each region, and divided by the relative cost of living in each region to find a standardized average household income level [2, 4].

Region	State	Household Income	Relative Cost of Living	GDP/Cost of Living
West	Montana	61298.14286	98.4	622.9486063
	North Dakota	71242.85714	99.7	714.5722883
	Idaho	56801.57143	89.40	635.3643337
	Wyoming	71463.14286	99.50	718.2225413
	South Dakota	63622.14286	99.70	638.1358361
	Nevada	64151.85714	94.90	675.9942797
	Utah	68178.42857	93.00	733.1013825
	Colorado	70858	99.70	710.7121364
	Arizona	57530	100.80	570.734127
	New Mexico	53607.42857	92.50	579.5397683
	Average			**659.9325299**
Midwest	Texas	58501.42857	91.40	640.0593936
	Louisiana	58225.28571	94.20	618.1028207
	Oklahoma	56009	90.00	622.3222222
	Arkansas	51060	91.00	561.0989011
	Kansas	65363.14286	91.80	712.0168067
	Missouri	62174.14286	92.90	669.2588036
	Illinois	70409.57143	94.90	741.934367
	Wisconsin	67491.57143	95.10	709.6905513
	Average			**659.3104833**

Region	State	Household Income	Relative Cost of Living	GDP/Cost of Living
Southeast	Michigan	61509.57143	94.40	651.5844431
	Indiana	62075.28571	90.00	689.7253968
	Kentucky	56104.71429	90.10	622.6938322
	North Carolina	56329.57143	95.60	589.2214585
	Tennessee	55442.85714	89.70	618.0920529
	South Carolina	54767.42857	95.00	576.4992481
	Mississippi	47782.14286	88.70	538.6938315
	Alabama	54319.14286	92.40	587.8695114
	Georgia	58182.42857	92.00	632.4177019
	Florida	58060.57143	97.80	593.6663745
	Average			**610.0463851**
Northeast	West Virginia	55800.14286	96.60	577.6412304
	Maryland	91657	122.30	749.4439902
	Delaware	75792.57143	106.90	709.00441
	Pennsylvania	69628.28571	101.10	688.7070793
	New Jersey	91343.85714	129.50	705.3579702
	New York	73606.85714	134.50	547.2628784
	Connecticut	92645.57143	133.80	692.4183216
	Rhode Island	74178.42857	125.80	589.6536452
	Vermont	69786.42857	118.30	589.9106388
	Maine	64579.28571	109.00	592.4705111
	New Hampshire	86343.28571	120.20	718.330164
	Massachusetts	90030.14286	121.20	742.8229609
	Average			**658.5853167**

Figure 6: Household income relative to cost of living

From Figure 6, we can see that the average household income relative to cost of living values for the southeast are much lower than the values for the rest of the nation. We can claim that **there is some kind of correlation between socioeconomic environment and obesity and overweight**. This may be because processed foods (which are generally not very healthy) are cheaper than fresh vegetables and fruits.

6 There's No Such Thing as Free Lunch

There clearly exists a trade-off between the interests of the three stakeholders: the students, school districts, and federal government. While healthy food is no doubt better for the students, they often consist of leafy greens and whole grain products that students simply do not enjoy eating. The federal government wants students to be nutritionally well-rounded, but healthy foods also tend to be more expensive. Coupled with the effect of fewer students buying lunches, school districts are faced with a large financial burden.

We aim to **create a meal plan that will be (1) tasty and (2) nutritionally wholesome and (3) cost under \$6 a week per student.** While our desired result is the fulfillment of these three goals, we began by tackling the latter two—creating a meal plan that is within the budget of the school and provides students all essential servings of grains, proteins, dairy, vegetables, and fruit.

6.1 Assumptions

- There are an equal number of girls and boys enrolled in school.
- We define a nutritionally balanced diet using the recommendations of MyPlate created by the USDA.
- A cup of dairy is 8 fluid oz.
- A cup of fruit or vegetables is 8 oz. It is not considered to be 16 oz. to account for air space.

- All necessary oils are provided through the cooking oils used when preparing other foods. We can safely assume this because the recommended amount of oil in a healthy diet is relatively little (5 tsp.).

- Suggested portions are appropriate for moderately active students. By the definitions provided by MyPlate, sizes are intended for "individuals who get less than 30 minutes per day of moderate physical activity, beyond normal daily activities" [7].

6.2 Model Design

We begin by examining the nutritional guidelines set by MyPlate for middle- and high-schoolers of both genders (Table 3).

	Grains	Fruits	Dairy	Vegetables	Protein	Oils
Males 14–18	14 oz.	16 oz.	32 oz.	20 oz.	12 oz.	1 oz.
Females 14–18	10 oz.	10 oz.	32 oz.	17 oz.	9 oz.	.8 oz.
Males 9–13	10 oz.	10 oz.	32 oz.	17 oz.	9 oz.	.8 oz.
Females 9–13	8.5 oz.	10 oz.	32 oz.	13 oz.	9 oz.	.8 oz.

Table 3: MyPlate suggested portions for 1 working week of lunches

Values provided by MyPlate were reported in terms of portions per day. To arrive at the values in the table above, we divided the daily suggested portions by three, then multiplied by five to find the expected portions for five school lunches.

We can see that 14–18 year old boys consume the most food, so by creating a meal plan that is within budget, nutritionally whole, and tasty for this demographic, we can safely assume that all other students are also covered.

To create our meal plan, we chose a variety of different foods intended to represent their respective food groups. We chose the foods arbitrarily, but we believe they can all be commonly found in the menu of a standard high school. In addition, they are all relatively healthy, and are essentially staple foods. Our chosen representative items are shown below:

Representative Foods:

- Grains: rice, pasta, bread
- Vegetables: spinach, carrots, lettuce
- Protein: chicken, beef, pork, beans
- Dairy: milk, cheese, yogurt (not frozen)
- Fruit: bananas, honeydew, oranges, cantaloupe

6.2.1 Adherence to Budget

We first decided to check whether we could fulfill the $6 per student per week criterion. This was easiest done by finding the prices for each of our representative foods and then converting to a per-ounce standard. The price per ounce for each individual item, (Table 4) as well as the average price per ounce of food group is shown (Table 5) below:

Grains	Spaghetti	Rice	Bread	Average
Price per oz	$0.110	$0.057	$0.220	$0.129
Fruits	Bananas	Oranges	Honeydew	Average
Price per oz	$0.026	$0.056	$0.043	$0.050
Dairy	Milk	Yogurt	Cheese	Average
Price per oz	$0.218	$0.07	$0.047	$0.046
Vegetables	Lettuce	Baby Carrots	Spinach	Average
Price per oz	$0.0325	$0.050	$0.065	$0.049
Protein	All Meats*	Black Beans		Average
Price per oz	$0.180	$0.100		$0.140

Table 4: Average costs

Food Group	Grains	Fruits	Dairy	Vegetables	Protein
Average Price per oz.	$0.129	$0.05	$0.046	$0.049	$0.13
Suggested oz.	14 oz.	16 oz.	32 oz.	20 oz.	12 oz.
Total Price	$6.73				

Table 5: Total costs

Using the prices above and food group requirements from Tables 4 and 5, we find that to create a fully balanced and nutritional lunch meal plan, it would cost approximately $6.73 per male student per week. Since we are only allotted $6 per student, we need to deviate slightly from the best nutritional proportions to create viable yet still healthy lunches. Grains and protein are the most expensive food groups, so by replacing 4.5 oz. of each of those groups with 4.5 oz. of fruits or vegetables, our modified cost is now $5.97. We believe this approach is both fiscally efficient and logical because of the high demand for fresh fruits and vegetables by current high school students [14]. Our final portions and costs are detailed in Table 6.

Food Group	Grains	Fruits	Dairy	Vegetables	Protein
Average Price per oz.	$0.129	$0.05	$0.046	$0.049	$0.13
Suggested oz.	10 oz.	20 oz.	32 oz.	24 oz.	8 oz.
Total Price	$5.97				

Table 6: Adjusted food amounts for boys 14–18

However, it's important to keep in mind that we based our calculations on a 14–18 male standard, the high school demographic demanding the highest caloric intake per week. If we look at costs for the next highest demographic, 14–18 females, we it only costs $5.22 to purchase a perfectly nutritional meal for girls 14–18 (Table 7). If we average $6.73 (boys) and $5.22 (girls) using the assumption that the number of boys and girls is equal, the **average cost of providing an optimally nutritious meal to one male and one female aged 14–18 amounts to $5.98**, just under the budget.

Food Group	Grains	Fruits	Dairy	Vegetables	Protein
Average Price per oz.	$0.129	$0.05	$0.046	$0.049	$0.13
Suggested oz.	10 oz.	10 oz.	32 oz.	16 oz.	9 oz.
Total Price	$5.22				

Table 7: Total costs for girls

6.2.2 Adherence to Optimal Caloric Intake

Now that we had confirmed the fiscal viability of the $6 budget, we turned to ensuring that every student would be able to meet their necessary caloric intake in a single meal setting. Like our strategy for confirming the budget viability, we first found the calorie content of each representative food, converted to a per ounce basis, then found the average calorie per ounce of each food group. Our results are summarized in Table 8:

Grains	Spaghetti	Rice	Bread	Average
Calories per oz	42	45.3	70	52.4
Fruits	Bananas	Oranges	Honeydew	Average
Calories per oz	105	71	10	62
Dairy	Milk	Yogurt	Cheese	Average
Calories per oz	18	24	88	43
Vegetables	Lettuce	Baby Carrots	Spinach	Average
Calories per oz	4	13	20	12
Protein	All Meats*	Black Beans		Average
Calories per oz	63	47	55	

Table 8: Average calories

Again, following our template from our budget analysis, we then check to see whether the 14–18 male age group would be able to consume the ideal amount of calories during a single school lunch. If this is shown to be true, it indicates that there at least exists enough food for all other demographics to consume their ideal caloric intake as well. For the purposes of time, efficiency, and space, we chose not to include analysis of other demographics. Table 9 shows the total calories consumed in a week's worth of suggested lunch time portions.

Food Group	Grains	Fruits	Dairy	Vegetables	Protein
Average Calories per oz.	52.4	62	43	12	55
Suggested oz.	14 oz.	16 oz.	32 oz.	20 oz.	12 oz.
Total Calories	4001				

Table 9: Total calories

Table 9 shows that in a typical week of school lunches, we expect the typical 14–18-year-old male to consume about 4000 calories. This equates to about 800 calories per lunch, which is almost exactly on par with the suggested daily calorie intake by the Institute of Medicine estimated energy requirement of 2438.5 calories per day (calculated for a 16-year-old male at the 50th percentile for both weight and height: 1.72m height, 66kg weight). As a side note, our own mathematical model for the same parameters returned an ideal intake of about 2550 calories per day, and we believe the deficit of about 50 calories per meal can be compensated by some minor snack at some point throughout the day, or any kind of drink.

6.3 Additional Funding

We can also consider the effects of an increase in meal funding by pushing the budget per student per week up to $7. One notable strength of our current model is our ability to adhere to the more stringent $6/week limitation and still provide students with enough nutritional value to meet their individual needs. Therefore, we determined that **our extra dollar should be used to improve**

taste and encourage more students to eat school lunch. Some portion of this additional funding can also be used purchase additional food.

To improve taste, we looked at student preferences from a 2013 survey conducted in Iowa Public School Systems [14]. The survey indicated that 73.5% of students were dissatisfied with their school meals after the Healthy, Hunger-Free Kids Act of 2010 was implemented. In their suggestions for tastier food, a large number of children indicated that they wanted more fruit and fewer processed foods in their meals. Therefore, we decided to budget our extra $1 toward fruit (5 oz. per week) and higher quality meats.

6.4 Model Justification

Averaging the cost of food for boys and girls does not affect the integrity of the model because our costs account for the raw food materials the school will be purchasing. All food will be prepared and students will be able to choose for themselves how much they would like to consume.

Our costs are merely those of the food and do not include shipping or handling. Nevertheless, we noticed that when ordering over $100 of food, as a school of hundreds or thousands would do, shipping and handling are usually free. In addition, the website we used for cost data did not offer food in bulk, school-appropriate quantities, so we can assume that costs would be even lower than those in Table 4.

While the MyPlate suggested portions are designed for students with somewhat active lifestyles, serving active and highly active students should not affect how much food is needed for two reasons. First, the number of fully sedentary students should balance out the higher nutritional needs of more active students. Second, highly active students involved in a varsity sports are generally more aware of their necessary higher food consumption and usually bring snacks of their own to supplement the school lunch. Though this claim is somewhat unbased, we believe it to be extremely reasonable.

6.5 Overview of Results

Overall, we successfully developed a lunch plan that fell within the $6 per week per student budget, provided nutritionally balanced meals (as defined by the USDA) for all students, and appealed to the average student's palate.

Since all nutritional needs were met on a $6 budget, an extra dollar allocated to weekly student budgets would best be used by improving food quality and taste. This could be accomplished by purchasing more fruits, unprocessed meats, and making larger portion sizes.

Finally, we determined that even with the $6 budget constraint, the average caloric intake for a male would still hover around 800 calories per lunch, which lies almost perfectly in line with our daily calorie requirements suggested by the Institute of Medicine and our own mathematical models.

Therefore, we can conclude that we successfully created a student lunch meal plan that adequately meets the financial needs of school systems, nutritional requirements of the government, and finicky palate of the students.

7 Strengths and Weaknesses

7.1 Strengths

- Our equations for calculating lunch caloric intake takes into account several additional factors that current equations do not, namely ethnicity, amount of sleep, and breakfast consumption. This allows for more inclusive and accurate results for lunchtime calorie needs.
- We indicate that our calorie intake model is in fact better than the IOM model by showing that our model has a lower cost function.
- Our distributions were created using an extremely large amount of data (20 million) and essentially factors were considered (gender, age, physical activity multiplier, height, weight, time slept, and breakfast consumption)
- Our meal plan is very inclusive and flexible because we use average costs for a food group rather than individual food costs. Depending on the preferences and availability of foods in different areas, our plan can be easily adjusted to accommodate all school systems.
- Our meal plan is able to satisfy the interests of all thre stakeholders. Students have tasty meals, school districts have relatively low food cost, and the federal government can ensure the nutritional balance of K-12 students across the nation.

7.2 Weaknesses

- When individuals are hungry and are consuming food, a lag period exists in which the brain needs to communicate to the rest of the body that it is "full." Therefore, if a student does not eat breakfast, s/he is more likely to overeat during lunch.
- In creating our distributions, we did not take into account race because it was not as available for state and regional data.
- In part 2, our model develops discrete values for all of the variables rather than utilizing probabilistic distributions. This might pose a problem with inaccuracy, but should not be very significant due to our large sample size.

7.3 Sensitivity Analysis

Changing coefficients of our mathematical model determined from Section 4 could prove detrimental to our data analysis conducted in Section 5. This is because the 800 calories provided by school lunches is close to the average caloric requirement for lunch for high school students. As a result, if either the coefficients of the model are manipulated in a fashion to increase the caloric intake by some amount or the amount of calories provided by a school lunch is decreased, we may observe a large perceived change in satisfaction ratings. However, our model is resilient to change in data, especially due to the nature of the model developed in part 2. A small change in variables will have almost no effect on the satisfaction percentage because we have a sample size of 20 million individuals.

To test the sensitivity of our distribution model, we chose to alter the average level of physical activity and calculated percent change in our satisfaction result. Table 10 illustrates these percentages. In the real world, if the bulk of students suddenly become more or less active, the school lunch will become less or more satisfactory, calorie-wise. Because the quality of the school lunch depends on this, it is important to calculate this sensitivity.

Physical Activity Multiplier offset	National Satisfaction Percentage
–0.20	82.92%
–0.10	73.30%
+0.00	60.01%
+0.10	42.91%
+0.20	26.38%

Table 10: Change in US satisfaction percentage vs. physical activity multiplier offset

As we can see in Table 10, changing the physical activity multiplier of the entire sample has a significant impact on the percentage of students that are satisfied with school meals. However, this will most likely not happen. Changing the PAM by 0.1 or 0.2 as done above means that the entire nation's physical activity level shifts. This will only happen on a long-term time scale. In the short run, it is more likely that a school's activity level shifts—which will not influence the national satisfaction percentage by much at all.

If the cost of food increases, our meal plan would need to change. Currently, our cost is just under the $6 budget cutoff, so any increase in resource price would necessitate sacrificing some nutritional value to meet our budget. If this were to happen, we suggest portions of grains and proteins, the more expensive food groups, to be cut and portions of fruits and vegetables to be increased. Similarly, if a particular nutritional or caloric minimum were to be imposed, we suggest an increase in portions of grains, fruits and proteins, food groups with a higher calorie per ounce ratio

8 Future Work

- In our mathematical model devised in Part 1, we neglected to include consideration of body fat percentage versus basal metabolic rate. Studies have shown that different amounts of caloric intake are required based on body-muscle-to-fat ratio [1], and we wish to incorporate this into our model in the future.

- Socioeconomic standing is also a large factor in determining the diet of individuals: those who are at a lower socioeconomic standing will tend toward buying less expensive, less nutritious, but more caloric foods. This causes a higher incidence of obesity in their population, greatly influencing their daily required caloric intake.

- We want to incorporate ethnicity as a factor in our distributions in Section 5 to more accurately represent the diverse populations in this country.

- We would like more raw data in order to further train the model that we created in Section 4. Doing so would result in a more accurate predictor of caloric needs.

References

[1] Basal metabolic rate. Retrieved from: `http://www.shapefit.com/basal-metabolic-rate.html`, 2010.

[2] Family income per state, US census. Retrieved from: `http://www.census.gov/hhes/www/income/data/Fam_Inc_SizeofFam1.xls`, 2010.

[3] Kids push back on new school lunches. Retrieved from: `http://usatoday30.usatoday.com/news/nation/story/2012/09/28/kids-push-back-on-new-school-lunch/57842204/1`, 2012.

[4] Cost of living by state. Retrieved from: `http://www.top50states.com/cost-of-living-by-state.html`, 2013.

[5] The Case for Eating Breakfast. Retrieved from: `http://www.healthychildren.org/English/healthy-living/nutrition/pages/The-Case-for-Eating-Breakfast.aspx`, 2013.

[6] Defining Obesity. Retrieved from: `http://www.cdc.gov/obesity/adult/defining.html`, 2014.

[7] Food groups overview. Retrieved from: `http://www.choosemyplate.gov/food-groups/`, 2014.

[8] Metabolic Equivalents. Retrieved from: `http://www.globalrph.com/metabolic_equivalents.htm`, 2014.

[9] NM Astbury, MA Taylor, and IA Macdonald. Breakfast consumption affects appetite, energy intake, and the metabolic and endocrine responses to foods consumed later in the day in male habitual breakfast. *The Journal of Nutrition*, pages 1381–1389, 2011.

[10] Ogden CL, Carroll MD, Kit BK, and Flegal KM. Prevalence of childhood and adult obesity in the united states, 2011-2012. *JAMA*, 311(8):806–814, 2014.

[11] Stephen R Daniels, Donna K Arnett, Robert H Eckel, Samuel S Gidding, Laura L Hayman, Shiriki Kumanyika, Thomas N Robinson, Barbara J Scott, Sachiko St Jeor, and Christine L Williams. Overweight in children and adolescents: pathophysiology, consequences, prevention, and treatment. *Circulation*, 111(15):1999–2012, April 2005.

[12] WH Dietz. Overweight in childhood and adolescence. *New England Journal of Medicine*, 2004.

[13] David S Freedman, Zuguo Mei, Sathanur R Srinivasan, Gerald S Berenson, and William H Dietz. Cardiovascular risk factors and excess adiposity among overweight children and adolescents: the Bogalusa Heart Study. *The Journal of Pediatrics*, 150(1):12–17.e2, January 2007.

[14] Elizabeth Golembiewski. Lunch Survey: Formative Research for Communication about School Meal Pattern Changes. *University of Iowa Public Policy Center*, 2013.

[15] LH Kushi, T Byers, and C Doyle. Cancer Society Guidelines on Nutrition and Physical Activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. *CA: A Cancer Journal ...*, 2006.

[16] C Li, ES Ford, G Zhao, and AH Mokdad. Prevalence of pre-diabetes and its association with clustering of cardiometabolic risk factors and hyperinsulinemia among US adolescents National Health and Nutrition Examination Survey 2005–2006. *Diabetes Care*, 32(2), 2009.

[17] Rachel R Markwald, Edward L Melanson, Mark R Smith, Janine Higgins, Leigh Perreault, Robert H Eckel, and Kenneth P Wright. Impact of insufficient sleep on total daily energy expenditure, food intake, and weight gain. *Proceedings of the National Academy of Sciences of the United States of America*, 110(14):5695–700, April 2013.

[18] Institute of Medicine of the National Academies. Dietary reference intakes for energy, carbohydrate, fiber, fat, fatty acids,cholesterol, protein,and amino acids. *The National Academies Press*, 2005.

[19] Teresa F Rutledge and Martha F Boyd. Morbidity and Mortality Weekly Report Youth Risk Behavior Surveillance — United States , 2009. 59, 2010.

[20] National Center for Health Statistics (US). Health, United States. 2004.

[21] C Weyer, S Snitker, C Bogardus, and E Ravussin. Energy metabolism in African Americans: potential risk factors for obesity. *The American Journal of Clinical Nutrition*, 70(1):13–20, July 1999.