Predicting Micronutrient Plasma Concentrations

Statistics 36-711: Applied Regression
Carnegie Mellon University

October 30, 2001

Abstract

Previous studies have found a relationship between cancer risk and micronutrient plasma concentration. In the present study, I looked for a relationship between micronutrient concentration and personal and dietary factors. I fit linear regression models to a data set containing 14 measurements—including micronutrient plasma concentrations, personal characteristics, and dietary habits—on each of 315 subjects. Beta-carotene concentration could be modeled reasonably well, and the model suggested that vitamin usage and never smoking can substantially increase the concentration. However, no model fit retinol concentration well, and I could draw no firm conclusions about its determinants. Making any inferences from this study is difficult, because the sampling process violated some important statistical assumptions.

1 Introduction

Previous studies have found a relationship between cancer risk and micronutrient plasma concentration. Specifically, lower plasma levels of beta-carotene and retinol seem to be associated with a higher risk of some kinds of cancer. In this study, I address the next logical question: What physical characteristics and dietary habits determine the plasma concentrations of these micronutrients? My goals are to improve our scientific understanding of the problem and to suggest ways for Americans to reduce their risk of cancer. My method is fitting linear regression models to data obtained from a cross-sectional study that measured micronutrient plasma concentrations, personal characteristics, and dietary habits for each of 315 subjects. As I discuss in Section 4, however, problems with the sampling process limit the inferences one can make from my results.

2 Data

2.1 Overview

Fourteen measurements were taken of each of 315 subjects, yielding a data set with 14 variables and 315 observations. The subjects were patients who had an elective surgical procedure to biopsy or remove a lesion of the lung, colon, breast, skin, or uterus that was found to be non-cancerous. The data set has no missing values, but it does have one anomalous observation (number 62):

```
age sex smok  quet vit    cal   fat  fib alco chol beta  ret betap retp
 65   1    3 23.38   3 6662.2 164.3 11.3  203  603 2893 1364    96  317
```

Both the 6662 calories per day and the 203 alcoholic drinks per week (an average of 29 per day) are extreme; the next-highest caloric intake is 4373, and the next-highest alcholic intake is 35. It is also odd that a person who consumes so many calories would have a relatively low body mass (23.38, well below the obesity cut-off of 28). Since I did not have any more information

Variable	Description
age	age in years
sex	1 = male; 2 = female
smok	smoking status: 1 = never; 2 = former; 3 = current
quet	quetelet (body mass) index: (weight in kg)/(height in m)^2
vit	vitamin usage: 1 = often; 2 = rarely; 3 = never
cal	number of calories consumed per day
fat	grams of fat consumed per day
fib	grams of fiber consumed per day
alco	number of alcoholic drinks consumed per week
chol	micrograms of cholesterol consumed per day
beta	micrograms of dietary beta-carotene per day
ret	micrograms of dietary retinol per day
betap	plasma concentration of beta-carotene in ng/ml
retp	plasma concentration of retinol in ng/ml

Table 1: Variables in the data set

about the observation, I assumed that it was miscoded, and thus I dropped it from the data set.

Table 1 describes the variables and gives their abbreviated names.

Because I wanted to explain the plasma concentration of beta-carotene and retinol, betap and retp were the natural dependent variables. Both variables, however, are truncated at 0. Since the classical regression model assumes that the dependent variable is normally distributed, and the support of a normal distribution is the whole real line, I looked for a transformation that would map the positive reals to the full real line. The logarithmic transformation was an obvious choice, and it brought the additional benefit of making the variables' distributions more symmetric. I called the transformed variables lbetap and lretp.

In the case of betap, the transformation involved a complication: one observation (number 257) had a value of 0, and $\log(0) = -\infty$. I therefore needed to replace the $-\infty$ with a real number that was still smaller than any other value of lbetap and that fit in with the other values in a plot of betap against lbetap. Zero is a simple number that meets both criteria, as

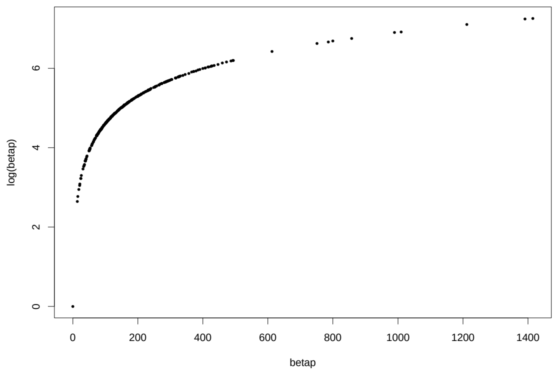

Figure 1: A plot of betap against lbetap, with (0, 0) substituted for $(0, -\infty)$.

Figure 1 shows, and so I chose it. However, this decision should be based on substantive knowledge, and so I invite my medical colleagues to suggest a better approach.

2.2 Exploratory Data Analysis

Most of the important information about the variables is best conveyed visually. Figure 2 shows histograms of betap, lbetap, retp, and lretp, as well as a scatterplot of lbetap and lretp. Figures 3, 4, and 5 show, for each of the nine continuous independent variables, a histogram of the variable, a scatterplot of the variable and lbetap, and a scatterplot of the variable and lretp. Figure 6 shows scatterplots among all of the continuous independent variables. Figure 7 uses parallel boxplots to compare the dependent variables to each of the three discrete independent variables. Finally, Table 2 summarizes the discrete independent variables numerically.

Rather than repeat in words what is displayed in the figures, I will summarize the main results.

- There is no visible relationship between the two dependent variables, lbetap and lretp. (Their correlation coefficient is 0.20.)

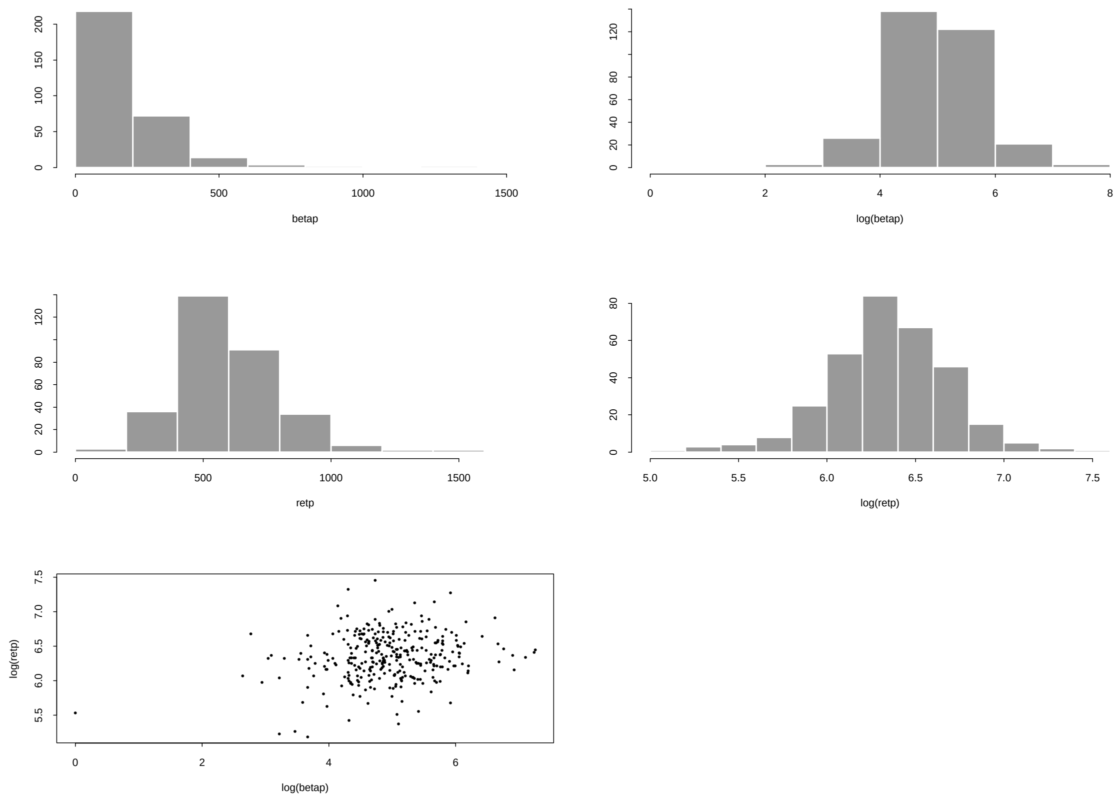

Figure 2: Graphical summaries of the two dependent variables, lbetap and lretp.

Variable	Breakdown
sex	13% male, 87% female
smok	14% currently, 50% formerly, 37% never
vit	35% never, 26% rarely, 39% often

Table 2: Numerical summaries of the three discrete independent variables.

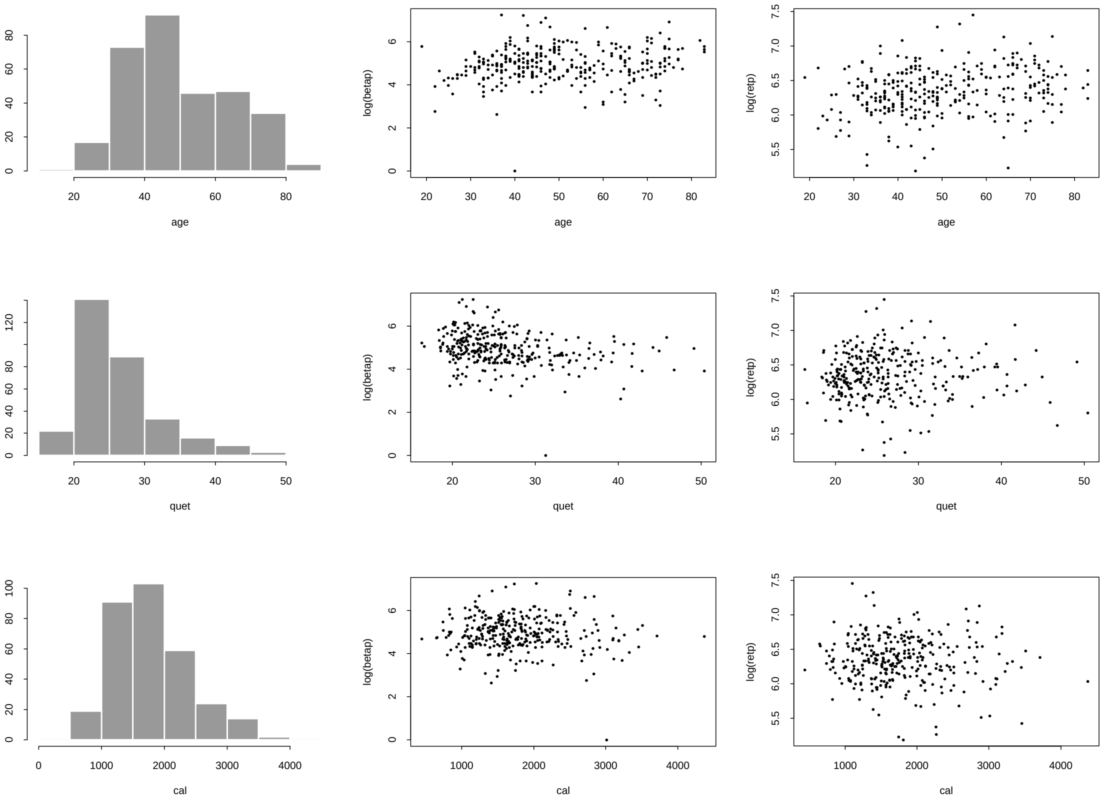

Figure 3: Graphical summaries of the independent variables age, quet, and cal.

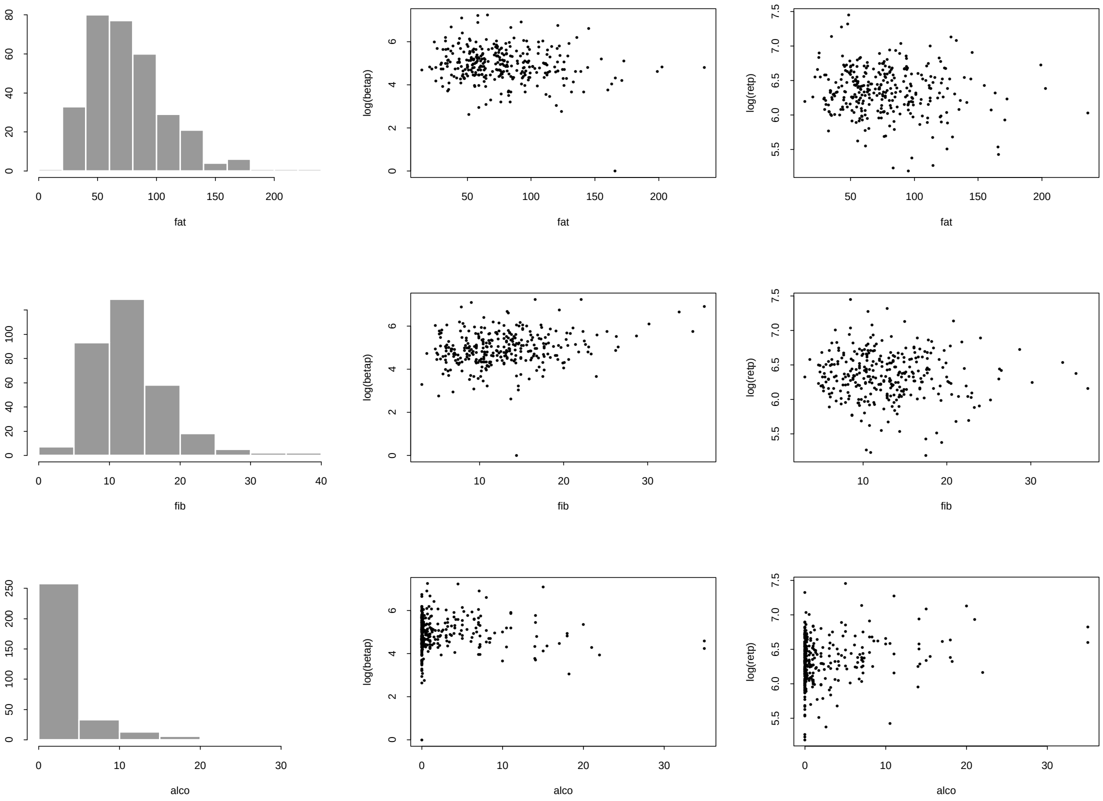

Figure 4: Graphical summaries of the independent variables fat, fib, and alco.

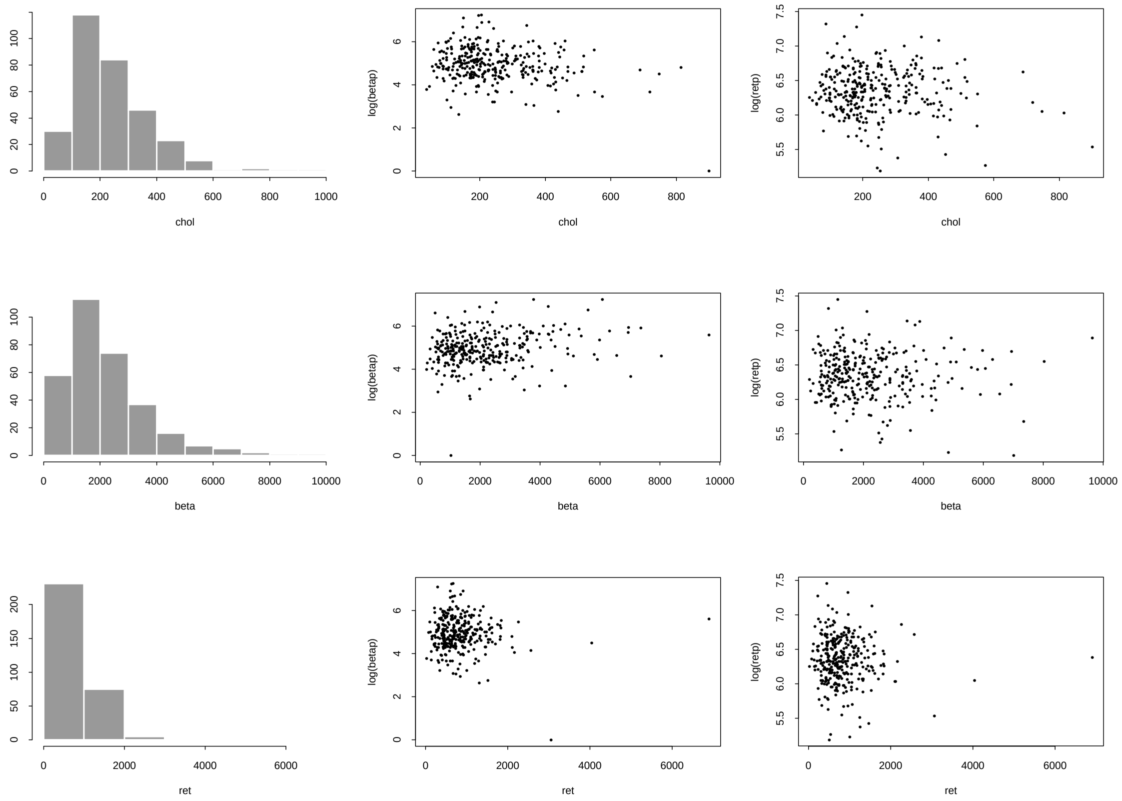

Figure 5: Graphical summaries of the independent variables chol, beta, and ret.

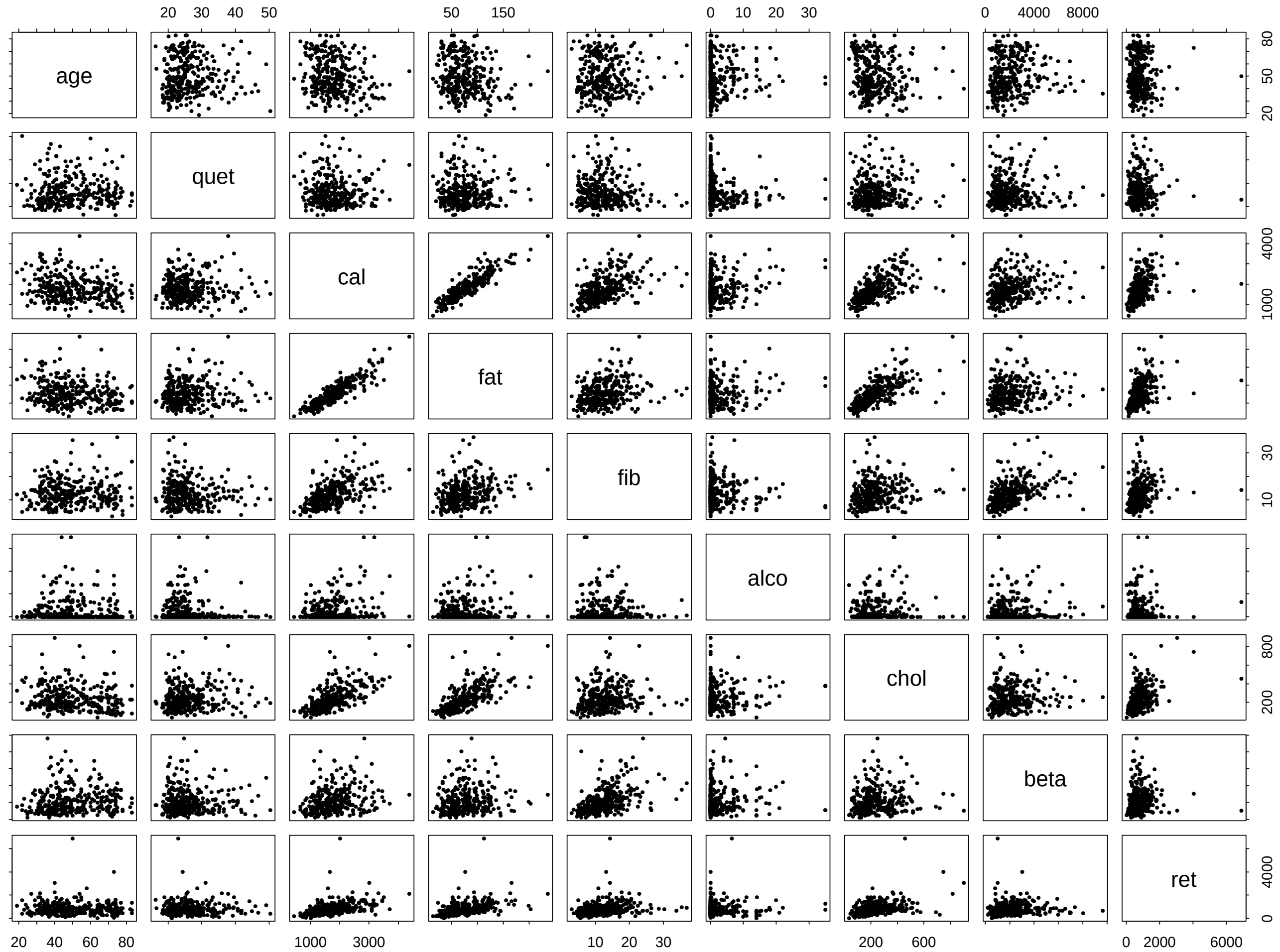

Figure 6: Relationships among the nine continuous independent variables.

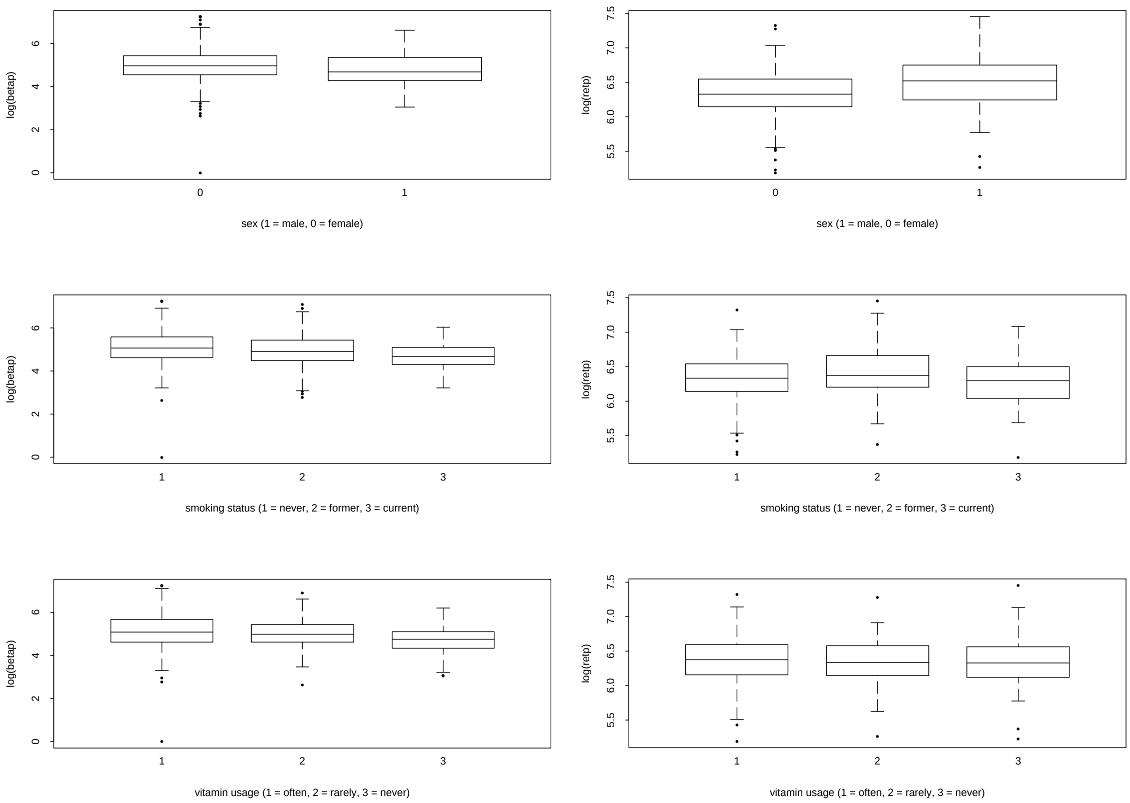

Figure 7: Graphical comparisons between the dependent variables and each of the three discrete independent variables, sex, smok, and vit.

- There is no visible relationship in almost every pairing of dependent and continuous independent variable. There seems to be a weak negative relationship between quet and lbetap, and a weak positive relationship between fib and lbetap and between beta and lbetap.

- There is no visible relationship in almost every pairing of dependent and discrete independent variable. Males, current smokers, and those who never take vitamins seem to have slightly lower levels of lbetap, while males have slightly higher levels of lretp.

- There are strong, positive relationships among the independent variables related to food intake: cal, fat, fib, chol, beta, and ret. This makes sense because cal measures overall food intake, while the others measure intake of specific foods. The formal correlation matrix is as follows:

	cal	fat	fib	chol	beta	ret
cal	1.0000000	0.8983502	0.5159527	0.6603640	0.25418928	0.41806787
fat	0.8983502	1.0000000	0.2818318	0.7031928	0.14103088	0.40968950
fib	0.5159527	0.2818318	1.0000000	0.1583234	0.48331001	0.21572316
chol	0.6603640	0.7031928	0.1583234	1.0000000	0.11284039	0.44143118
beta	0.2541893	0.1410309	0.4833100	0.1128404	1.00000000	0.05157041
ret	0.4180679	0.4096895	0.2157232	0.4414312	0.05157041	1.00000000

  Cal and fat are almost perfectly collinear, which violates an assumption of the linear regression model, and so I needed to drop one. I chose to drop cal for interpretational reasons: The effect of overall food intake is less interesting than the effect of fat intake. Also, "eat less" would be a less effective public health recommendation than "eat less fat."

- Surprisingly, there is no visible relationship between vitamin usage and micronutrient dietary intake (Figure 8). Since most multivitamins include beta-carotene and retinol, it seems that the survey distinguished between micronutrients consumed through regular food and micronutrients consumed through special sources, like vitamins or other external supplements.

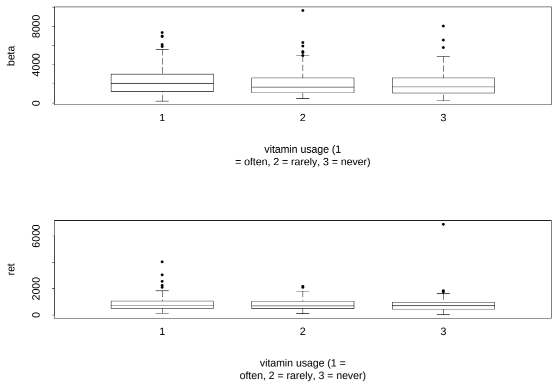

Figure 8: Graphical comparison of vitamin usage and micronutrient dietary intake.

3 Analysis

3.1 Explaining Beta-Carotene Plasma Concentration

First, I fit a linear regression model in which lbetap was regressed on all of the independent variables except cal (I discuss the reasons for omitting cal in Section 2). The model fit the data reasonably well. As Figure 9 shows, the fitted values tracked the true values fairly closely, the residuals seemed to be randomly distributed about 0, and the standardized residuals were approximately normally distributed. R^2 was 0.25, which means that the model accounted for about 25 percent of the variability in lbetap. For the full estimation results, please see Appendix A.

I hypothesized that both smoking and excessive alcohol drinking alter the body's physiological processes and thus influence the relationship between micronutrient plasma concentration and the other independent variables. To test this hypothesis, I added interaction terms to the original model; these terms represent the assumption that some coefficients are actually linear functions of other independent variables.[1]

[1] If X_1 has coefficient β_1 and I assume that $\beta_1 = \gamma_0 + \gamma_1 X_2$, then I will add the

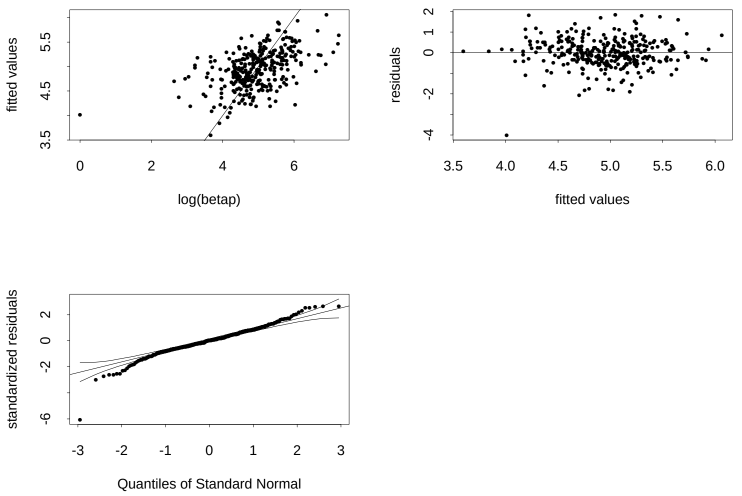

Figure 9: Diagnostic plots for regressing lbetap on all independent variables except cal.

Therefore, my second step was to add interaction terms between each independent variable and an indicator variable of whether the subject currently smokes. As Figure 10 shows, the model fit was about the same as in the first case. Furthermore, none of the interaction terms had a statistically significant coefficient at the 0.05 level (please see Appendix B for the details), and so I used an F test to formally test the hypothesis that all interaction terms had coefficients of 0. The test could not reject this hypothesis at the 0.05 level (please see Appendix C for the details), and so I concluded that smoking does not affect the relationships between lbetap and the other independent variables.

Third, I added (to the original model) interaction terms between each independent variable and alco. The model fit was about the same as in the first case (Figure 11), although two interaction terms had significant coefficients: vit.oft and ret (please see Appendix D for the details). I used an F test to test the hypothesis that all of the other interaction terms had coefficients of 0, and the test could not reject this hypothesis at the 0.05 level (please see

interaction term X_1X_2.

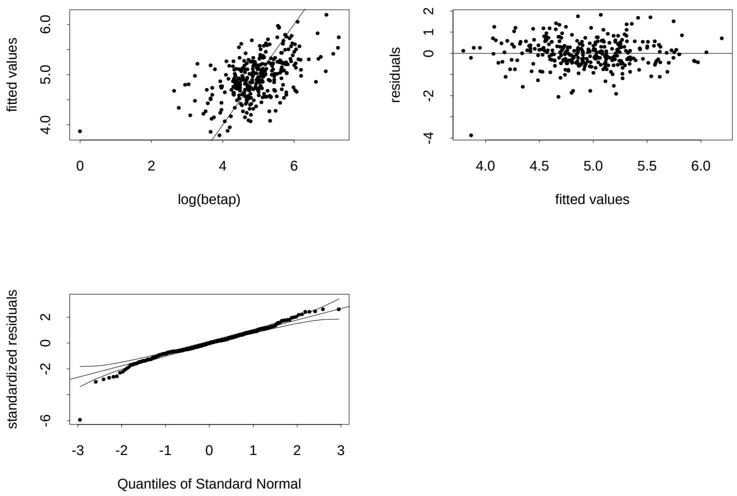

Figure 10: Diagnostic plots for regressing lbetap on all independent variables except cal, with a smoking interaction effect.

Appendix E for the details), so I concluded that alcohol drinking affects only the lbetap-vit.oft and lbetap-ret relationships.

All of this suggests that the best model is the original one plus interaction terms between alco and vit.oft and between alco and ret. Figure 12 shows the diagnostic plots for this model, which reveal a fairly good fit and suggest that the model assumptions hold. R^2 for this model was 0.26. Appendix F contains the full estimation results, and Table 3 lists the variables with significant coefficients and interprets their coefficients in substantive terms. Specifically, the table gives the change in betap implied by a 1-unit increase in the independent variable.[2] Most of these effects, while statistically significant, are substantively small; only vitamin usage and never smoking have substantively significant effects on the plasma concentration of beta-carotene.

Although the coefficient on vit.oft is not statistically significant, the coefficient on the alco-vit.oft interaction term tells us that each 1-unit increase in

[2] A 1-unit change in independent variable X_i leads to a β_i-unit change in $\log(Y)$, which means a change by a factor of $\exp(\beta_i)$ in Y.

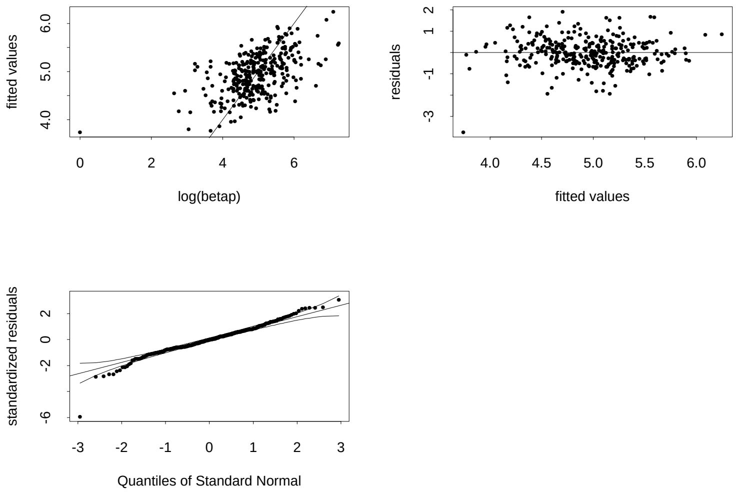

Figure 11: Diagnostic plots for regressing lbetap on all independent variables except cal, with an alcohol interaction.

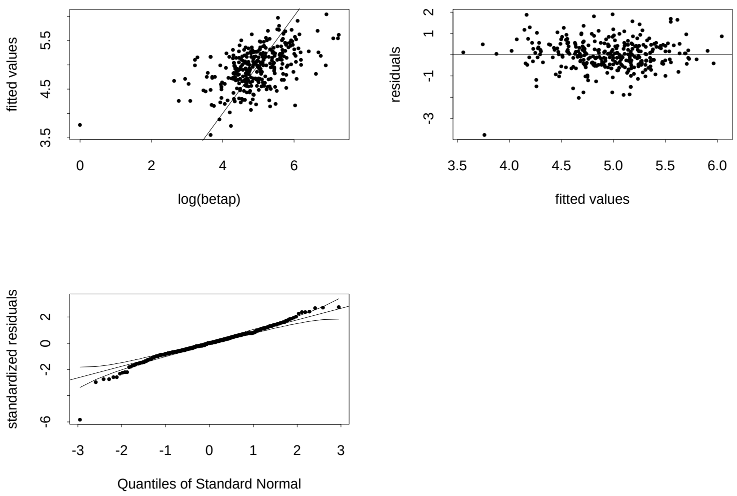

Figure 12: Diagnostic plots for the final lbetap model.

Variable	Factor by which betap is multiplied
age	1.0067
smo.nev	1.2906
quet	0.9681
vit.oft*	1.1876
vit.rar	1.3226
fib	1.0232
chol	0.9990
beta	1.0001

Table 3: The change in beta-carotene plasma level implied by a 1-unit increase in the significant independent variables.

weekly alcohol consumption multiplies the effect of vit.oft by 1.0432. The alco-ret interaction term had a coefficient of 0 in the final model, so alcohol does not seem to have an effect on the lbetap-ret relationship.

3.2 Explaining Retinol Plasma Concentration

First, I fit a linear regression model in which lretp was regressed on all of the independent variables except cal. The model fit was substantially worse than it was for lbetap; although the standardized residuals were almost normal, the fitted values did not track the true values very closely (Figure 13). R^2 was only 0.13. The full estimation results are in Appendix G.

Since only age and alco had coefficients that were significantly different from 0 at the 0.05 level, I used an F test to formally test the hypothesis that all of the other coefficients were 0. Although the test could not reject the hypothesis at the 0.05 level (see Appendix I for the details), the reduced model fit the data even more poorly. As Figure 14 shows, the fitted values hardly tracked the true values at all, and the standardized residuals were farther from being normal. I therefore chose to retain the full model.

Second, I looked for a smoking interaction effect. As with lbetap, the coefficients of the interaction terms were all nonsignficant (see Appendix J for the details), and an F test could not reject the hypothesis of no interaction

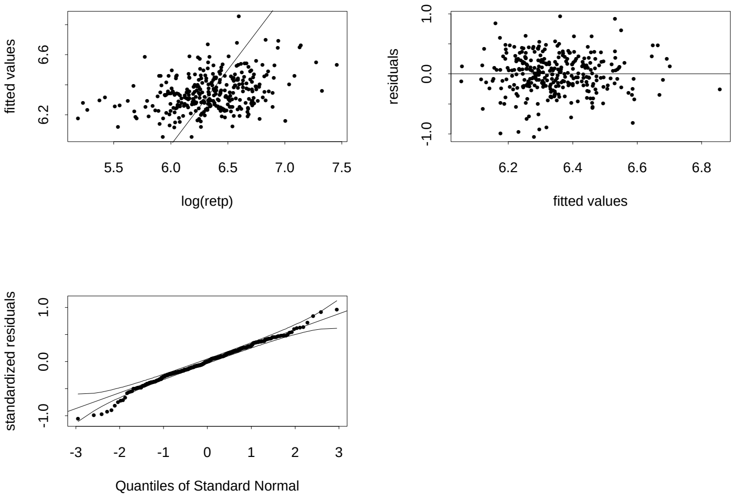

Figure 13: Diagnostic plots for regressing lretp on all independent variables except cal.

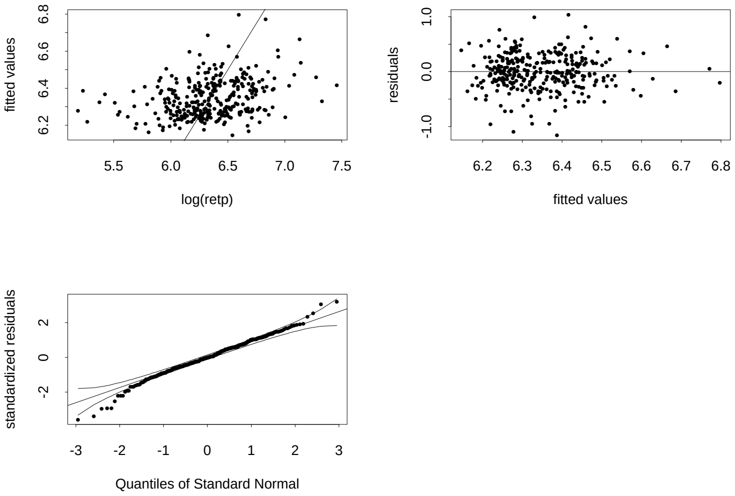

Figure 14: Diagnostic plots for regressing lretp on only age and alco.

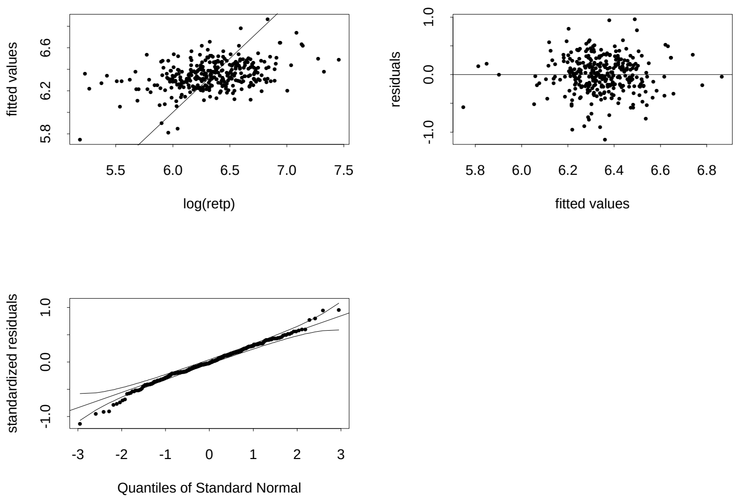

Figure 15: Diagnostic plots for regressing lretp on all independent variables except cal, with a smoking interaction.

effect (please see Appendix K for the details). Furthermore, the interaction terms did not improve the model fit, as Figure 15 shows, and so I concluded that smoking does not affect the relationship between lretp and the other independent variables.

Finally, I looked for an alcohol interaction effect. The interaction terms improved the model fit slightly, as Figure 16 shows; the fitted and true values were better synchronized and the residuals seemed more randomly distributed about 0. The full results are in Appendix L. Only one interaction term, vit.rar, had a significant coefficient, so I used an F test to formally test the hypothesis that the other interaction terms had coefficients of 0. The test could not reject the hypothesis (see Appendix M for the details), so I kept only the alco-vit.rar term.

This suggests that the best model is the original plus an alco-vit.rar interaction term. This model fits the data about as well as the original; one can from Figure 17 that the residuals are reasonably well distributed, although the fitted values still do not follow the true values very closely. R^2 was 0.13.

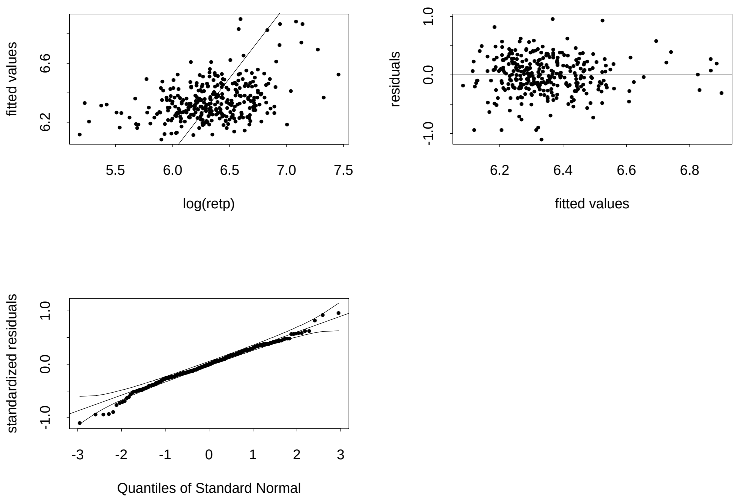

Figure 16: Diagnostic plots for regressing lretp on all independent variables except cal, with an alcohol interaction.

Variable	Factor by which retp is multiplied
age	1.0043
alco	1.0105

Table 4: The change in retinol plasma level implied by a 1-unit change in the significant independent variables.

Full results are in Appendix N. Table 4 interprets the significant coefficients in substantive terms. In this case, however, neither of the statistically significant coefficients translates into a substantively significant effect.

4 Discussion

This study analyzed a data set containing 14 measurements on each of 315 subjects, with the goal of finding a statistical link between personal and dietary factors and the plasma concentrations of two micronutrients that

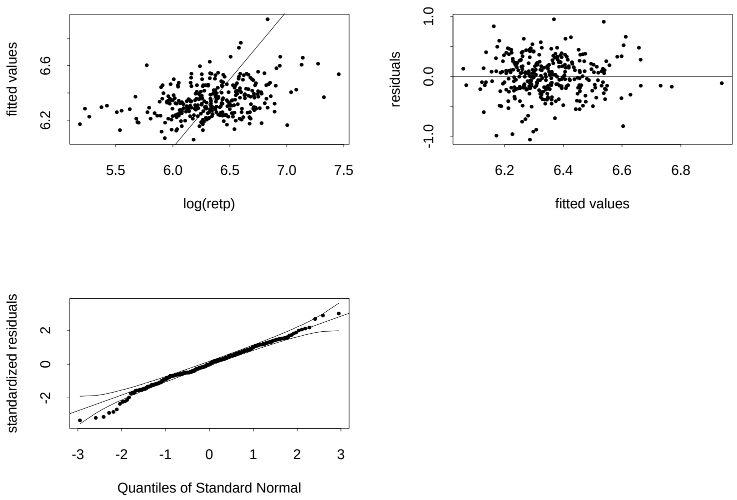

Figure 17: Diagnostic plots for final lretp model.

may reduce the risk of cancer.

4.1 Findings: Beta-Carotene

A linear regression model that predicted the log of beta-carotene plasma concentration from all of the personal and dietary factors in the data set, plus an interaction term between alcoholic intake and occasional vitamin usage, fit the data reasonably well. Estimation of the model led to the conclusion that only vitamin usage and never smoking have substantively large effects on beta-carotene concentration. Both taking a vitamin occasionally and never smoking cause one's plasma concentration to increase by a factor of about 0.3. Furthermore, drinking one extra alcoholic drink per week multiplies the effect of frequent vitamin usage by about 1.04.

4.2 Findings: Retinol

Linear regression with the variables in this data set appears to be the wrong approach to modeling retinol plasma concentration. None of the models I tried fit especially well, and even in the best of these, no variable had a substantively significant impact on retinol plasma concentration.

4.3 Limitations

The linear regression model assumes that the error term from one observation is unrelated to the error term from another, and that every error has approximately the same variance.[3] The model also assumes that the observations constitute a simple random sample from the conceptual population of interest.

In this case, the population of interest is the adult American population, from which the data clearly do not constitute a simple random sample; this sample includes only those adult Americans who sought treatment for a certain kind of health problem. We therefore cannot be sure that the results of any analysis of these data can be generalized to the general American public.

The error assumptions are tougher to check. The model fit diagnostics, at least for predicting beta-carotene plasma concentration, were relatively encouraging (Section 3). On the other hand, the subjects were selected based on a characteristic—seeking treatment for a non-cancerous lesion—that may be related to the dependent variable, and this could lead to a violation of the error assumptions.

4.4 Future Research

This kind of study would have greater validity if it were based on a true simple random sample from the adult American population. I believe, there-

[3]Technically, $\{\epsilon_i\}$ are independent and identically distributed members of the $N(0, \sigma^2)$ family of distributions.

fore, that the next step in studying this problem should be collecting data based on such a sample. The results of this study suggest that age, smoking status, vitamin usage, fiber consumption, and alcoholic intake are important variables to include. Exercise was not included in the current data set but might be worth including in the future, since it has been linked to numerous health benefits.

5 Acknowledgements

In this study, I used technical material from lectures by Professor Brian Junker and from the textbook *Applied Regression Analysis*[4]. I also benefited from private conversations with Professor Junker.

[4] J.O. Rawlings, S.G. Pantula, and D.A. Dickey. *Applied Regression Analysis.* Springer-Verlag, New York, Second Edition, 1998.

A Full Regression Results: Lbetap on All

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	5.0538	0.2841	17.7908	0.0000
age	0.0060	0.0030	1.9679	0.0500
male.2	-0.1674	0.1349	-1.2407	0.2157
smo.nev.2	0.2517	0.1301	1.9342	0.0540
smo.for.2	0.1917	0.1321	1.4515	0.1477
quet	-0.0322	0.0069	-4.6625	0.0000
vit.oft.2	0.2495	0.0965	2.5845	0.0102
vit.rar.2	0.2734	0.1057	2.5853	0.0102
fat	-0.0009	0.0018	-0.4953	0.6208
fib	0.0225	0.0092	2.4536	0.0147
alco	0.0024	0.0087	0.2719	0.7859
chol	-0.0011	0.0005	-2.3980	0.0171
beta	0.0001	0.0000	2.0237	0.0439
ret	0.0000	0.0001	-0.1391	0.8895

B Full Regression Results: Lbetap on All With Smoking Interaction

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	5.1646	0.3076	16.7894	0.0000
age	0.0063	0.0032	1.9560	0.0514
male.2	-0.2469	0.1456	-1.6959	0.0910
smo.cur.2	0.6767	1.1796	0.5736	0.5667
quet	-0.0310	0.0073	-4.2641	0.0000
vit.oft.2	0.3016	0.1036	2.9111	0.0039
vit.rar.2	0.2820	0.1151	2.4502	0.0149
fat	-0.0007	0.0018	-0.4043	0.6863
fib	0.0260	0.0095	2.7339	0.0066
alco	0.0061	0.0099	0.6158	0.5385
chol	-0.0013	0.0005	-2.5926	0.0100
beta	0.0001	0.0000	2.4632	0.0143
ret	0.0000	0.0001	-0.2639	0.7921
smo.cur.2:age	-0.0078	0.0122	-0.6408	0.5222

smo.cur.2:male.2	0.7193	0.5139	1.3996	0.1627
smo.cur.2:quet	0.0074	0.0282	0.2618	0.7936
smo.cur.2:vit.oft.2	-0.5161	0.3541	-1.4575	0.1461
smo.cur.2:vit.rar.2	-0.1501	0.3084	-0.4868	0.6267
smo.cur.2:fat	-0.0076	0.0080	-0.9499	0.3430
smo.cur.2:fib	-0.0297	0.0329	-0.9034	0.3670
smo.cur.2:alco	-0.0041	0.0224	-0.1849	0.8535
smo.cur.2:chol	0.0013	0.0015	0.9035	0.3670
smo.cur.2:beta	-0.0001	0.0001	-1.0294	0.3041
smo.cur.2:ret	0.0002	0.0003	0.7272	0.4677

C ANOVA Table: Lbetap, Smoking Interaction

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
age	1	4.0657	4.06574	8.23302	0.0044159
male.2	1	5.2154	5.21544	10.56114	0.0012908
smo.cur.2	1	3.4496	3.44962	6.98539	0.0086645
quet	1	16.6687	16.66869	33.75368	0.0000000
vit.oft.2	1	1.8227	1.82269	3.69090	0.0556903
vit.rar.2	1	3.6430	3.64298	7.37695	0.0070031
fat	1	1.5935	1.59353	3.22686	0.0734803
fib	1	7.6076	7.60760	15.40521	0.0001084
alco	1	0.0651	0.06514	0.13192	0.7167174
chol	1	2.8539	2.85389	5.77906	0.0168456
beta	1	2.0123	2.01226	4.07478	0.0444481
ret	1	0.0072	0.00722	0.01461	0.9038717
smo.cur.2:age	1	0.0109	0.01089	0.02206	0.8820309
smo.cur.2:male.2	1	1.1820	1.18196	2.39344	0.1229356
smo.cur.2:quet	1	0.0027	0.00274	0.00554	0.9407121
smo.cur.2:vit.oft.2	1	2.3263	2.32630	4.71070	0.0307868
smo.cur.2:vit.rar.2	1	0.0054	0.00536	0.01085	0.9171281
smo.cur.2:fat	1	0.7665	0.76649	1.55211	0.2138291
smo.cur.2:fib	1	1.4992	1.49925	3.03594	0.0824994
smo.cur.2:alco	1	0.0351	0.03507	0.07101	0.7900577
smo.cur.2:chol	1	0.4361	0.43612	0.88313	0.3481266
smo.cur.2:beta	1	0.6128	0.61283	1.24096	0.2662093

```
smo.cur.2:ret   1    0.2612  0.26118  0.52888 0.4676631
    Residuals 290  143.2116  0.49383
```

The table yielded an F statistic of 1.3 on 11 and 290 degrees of freedom, which has a p-value of 0.22.

D Full Regression Results: Lbetap on All With Alcohol Interaction

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	5.0051	0.3081	16.2463	0.0000
age	0.0083	0.0034	2.4138	0.0164
male.2	-0.0434	0.1638	-0.2648	0.7914
smo.nev.2	0.2552	0.1484	1.7200	0.0865
smo.for.2	0.0901	0.1522	0.5923	0.5541
quet	-0.0326	0.0074	-4.3955	0.0000
vit.oft.2	0.1670	0.1106	1.5094	0.1323
vit.rar.2	0.2396	0.1234	1.9422	0.0531
fat	0.0009	0.0021	0.4044	0.6863
fib	0.0222	0.0106	2.0833	0.0381
alco	-0.0372	0.0935	-0.3980	0.6909
chol	-0.0011	0.0005	-2.0161	0.0447
beta	0.0001	0.0000	1.5849	0.1141
ret	-0.0002	0.0001	-1.5727	0.1169
alco:vit.oft.2	0.0663	0.0231	2.8735	0.0044
alco:ret	0.0000	0.0000	2.3516	0.0194
alco:age	-0.0009	0.0008	-1.0208	0.3082
alco:male.2	-0.0370	0.0263	-1.4094	0.1598
alco:smo.nev.2	-0.0107	0.0314	-0.3411	0.7333
alco:smo.for.2	0.0423	0.0274	1.5440	0.1237
alco:quet	0.0028	0.0023	1.2275	0.2206
alco:vit.rar.2	0.0193	0.0262	0.7365	0.4621
alco:fat	-0.0006	0.0004	-1.4448	0.1496
alco:fib	0.0004	0.0029	0.1324	0.8948
alco:chol	0.0000	0.0001	0.0114	0.9909
alco:beta	0.0000	0.0000	-0.3142	0.7536

E ANOVA Table: Lbetap, Alcohol Interaction

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
age	1	4.0657	4.06574	8.33594	0.0041813
male.2	1	5.2154	5.21544	10.69317	0.0012059
smo.nev.2	1	1.1916	1.19163	2.44318	0.1191346
smo.for.2	1	2.4284	2.42845	4.97902	0.0264260
quet	1	16.9675	16.96749	34.78826	0.0000000
vit.oft.2	1	1.6796	1.67956	3.44359	0.0645190
vit.rar.2	1	3.5934	3.59337	7.36745	0.0070421
fat	1	1.4872	1.48718	3.04915	0.0818455
fib	1	7.5237	7.52373	15.42582	0.0001075
alco	1	0.0905	0.09054	0.18564	0.6668922
chol	1	2.8821	2.88205	5.90904	0.0156738
beta	1	2.0918	2.09181	4.28881	0.0392546
ret	1	0.0097	0.00968	0.01985	0.8880496
alco:vit.oft.2	1	1.6821	1.68213	3.44885	0.0643165
alco:ret	1	1.4765	1.47650	3.02724	0.0829452
alco:age	1	0.7342	0.73418	1.50528	0.2208631
alco:male.2	1	0.6978	0.69780	1.43069	0.2326359
alco:smo.nev.2	1	0.8812	0.88121	1.80674	0.1799572
alco:smo.for.2	1	1.5285	1.52853	3.13393	0.0777364
alco:quet	1	0.6698	0.66982	1.37333	0.2422093
alco:vit.rar.2	1	0.2111	0.21111	0.43284	0.5111254
alco:fat	1	1.7252	1.72522	3.53720	0.0610155
alco:fib	1	0.0052	0.00524	0.01075	0.9174835
alco:chol	1	0.0001	0.00006	0.00012	0.9913908
alco:beta	1	0.0481	0.04814	0.09870	0.7536195
Residuals	288	140.4680	0.48774		

The table yielded an F statistic of 1.3 on 11 and 289 degrees of freedom, which has a p-value of 0.23.

F Full Regression Results: Final Lbetap Model

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	5.1169	0.2836	18.0437	0.0000
age	0.0067	0.0030	2.2144	0.0276
male.2	-0.1849	0.1349	-1.3704	0.1716
smo.nev.2	0.2551	0.1292	1.9741	0.0493
smo.for.2	0.2144	0.1314	1.6309	0.1040
quet	-0.0324	0.0069	-4.7315	0.0000
vit.oft.2	0.1719	0.1061	1.6194	0.1064
vit.rar.2	0.2796	0.1054	2.6524	0.0084
fat	-0.0011	0.0018	-0.6322	0.5278
fib	0.0229	0.0091	2.5170	0.0124
alco	-0.0325	0.0183	-1.7739	0.0771
chol	-0.0010	0.0005	-2.2239	0.0269
beta	0.0001	0.0000	2.0164	0.0447
ret	-0.0001	0.0001	-1.0499	0.2946
alco:vit.oft.2	0.0423	0.0200	2.1109	0.0356
alco:ret	0.0000	0.0000	1.7303	0.0846

G Full Regression Results: Lretp on All

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	6.0991	0.1300	46.9067	0.0000
age	0.0045	0.0014	3.2139	0.0015
male.2	0.0671	0.0618	1.0868	0.2780
smo.nev.2	0.0097	0.0596	0.1628	0.8708
smo.for.2	0.0854	0.0605	1.4118	0.1590
quet	0.0017	0.0032	0.5304	0.5962
vit.oft.2	0.0482	0.0442	1.0897	0.2767
vit.rar.2	0.0517	0.0484	1.0691	0.2859
fat	-0.0008	0.0008	-0.9902	0.3229
fib	-0.0015	0.0042	-0.3580	0.7206
alco	0.0136	0.0040	3.4309	0.0007
chol	-0.0001	0.0002	-0.4819	0.6303
beta	0.0000	0.0000	-0.7972	0.4260
ret	0.0000	0.0000	-0.0801	0.9362

H Full Regression Results: Small Lretp Model

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	6.0470	0.0663	91.1938	0.0000
age	0.0052	0.0013	4.1700	0.0000
alco	0.0141	0.0037	3.8082	0.0002

I ANOVA Table: Small Lretp Model

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
age	1	1.81568	1.815682	17.31722	0.0000414
alco	1	1.52247	1.522467	14.52066	0.0001682
male.2	1	0.04717	0.047170	0.44989	0.5029040
smo.nev.2	1	0.14022	0.140223	1.33739	0.2484153
smo.for.2	1	0.20439	0.204386	1.94935	0.1636890
quet	1	0.01164	0.011641	0.11102	0.7392156
vit.oft.2	1	0.02364	0.023643	0.22550	0.6352262
vit.rar.2	1	0.10165	0.101652	0.96951	0.3255947
fat	1	0.49873	0.498726	4.75664	0.0299618
fib	1	0.06683	0.066828	0.63738	0.4252922
chol	1	0.03397	0.033969	0.32398	0.5696499
beta	1	0.06596	0.065960	0.62910	0.4283114
ret	1	0.00067	0.000673	0.00642	0.9361810
Residuals	300	31.45450	0.104848		

This table yielded an F statistic of 0.95 on 11 and 290 degrees of freedom, which has a p-value of 0.49.

J Full Regression Results: Lretp on All With Smoking Interaction

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	6.1534	0.1416	43.4660	0.0000

age	0.0048	0.0015	3.1933	0.0016
male.2	0.0722	0.0670	1.0780	0.2819
smo.cur.2	-0.6354	0.5429	-1.1705	0.2428
quet	0.0004	0.0033	0.1132	0.9100
vit.oft.2	0.0175	0.0477	0.3677	0.7134
vit.rar.2	0.0389	0.0530	0.7351	0.4629
fat	-0.0004	0.0008	-0.4582	0.6472
fib	-0.0013	0.0044	-0.3081	0.7583
alco	0.0131	0.0046	2.8566	0.0046
chol	-0.0003	0.0002	-1.1704	0.2428
beta	0.0000	0.0000	-0.0319	0.9745
ret	0.0000	0.0000	0.0407	0.9676
smo.cur.2:age	0.0016	0.0056	0.2859	0.7752
smo.cur.2:male.2	-0.1255	0.2365	-0.5306	0.5961
smo.cur.2:quet	0.0227	0.0130	1.7490	0.0814
smo.cur.2:vit.oft.2	0.2311	0.1630	1.4181	0.1572
smo.cur.2:vit.rar.2	-0.0511	0.1419	-0.3601	0.7190
smo.cur.2:fat	-0.0034	0.0037	-0.9240	0.3562
smo.cur.2:fib	-0.0014	0.0152	-0.0895	0.9288
smo.cur.2:alco	0.0051	0.0103	0.4980	0.6189
smo.cur.2:chol	0.0013	0.0007	1.8396	0.0669
smo.cur.2:beta	-0.0001	0.0001	-1.9300	0.0546
smo.cur.2:ret	0.0001	0.0001	0.7234	0.4700

K ANOVA Table: Lretp, Smoking Interaction

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
age	1	1.81568	1.815682	17.35917	0.0000409
male.2	1	0.27207	0.272070	2.60117	0.1078722
smo.cur.2	1	0.06718	0.067179	0.64227	0.4235450
quet	1	0.00415	0.004147	0.03965	0.8423083
vit.oft.2	1	0.00062	0.000624	0.00597	0.9384883
vit.rar.2	1	0.06908	0.069084	0.66049	0.4170544
fat	1	0.31051	0.310506	2.96865	0.0859585
fib	1	0.12110	0.121100	1.15780	0.2828161
alco	1	1.43161	1.431608	13.68716	0.0002583

chol	1	0.03919	0.039186	0.37465	0.5409615
beta	1	0.04759	0.047587	0.45497	0.5005231
ret	1	0.00185	0.001846	0.01765	0.8943978
smo.cur.2:age	1	0.08174	0.081744	0.78153	0.3774073
smo.cur.2:male.2	1	0.07238	0.072379	0.69200	0.4061710
smo.cur.2:quet	1	0.06234	0.062339	0.59600	0.4407367
smo.cur.2:vit.oft.2	1	0.05236	0.052365	0.50064	0.4797856
smo.cur.2:vit.rar.2	1	0.01021	0.010211	0.09762	0.7549263
smo.cur.2:fat	1	0.00002	0.000021	0.00020	0.9887728
smo.cur.2:fib	1	0.27748	0.277484	2.65294	0.1044445
smo.cur.2:alco	1	0.01535	0.015347	0.14673	0.7019629
smo.cur.2:chol	1	0.42122	0.421216	4.02711	0.0457025
smo.cur.2:beta	1	0.42652	0.426522	4.07785	0.0443687
smo.cur.2:ret	1	0.05474	0.054740	0.52335	0.4699997
Residuals	290	30.33254	0.104595		

This table yielded an F statistic of 1.4 on 11 and 290 degrees of freedom, which has a p-value of 0.17.

L Full Regression Results: Lretp on All With Alcohol Interaction

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	6.1765	0.1432	43.1220	0.0000
age	0.0045	0.0016	2.8525	0.0047
alco	-0.0390	0.0434	-0.8984	0.3697
vit.rar.2	-0.0168	0.0574	-0.2931	0.7696
male.2	0.0148	0.0762	0.1945	0.8459
smo.nev.2	0.0576	0.0690	0.8355	0.4041
smo.for.2	0.1031	0.0707	1.4568	0.1463
quet	-0.0008	0.0034	-0.2179	0.8277
vit.oft.2	-0.0025	0.0514	-0.0481	0.9616
fat	-0.0007	0.0010	-0.7509	0.4533
fib	-0.0029	0.0049	-0.5920	0.5543
chol	0.0000	0.0003	-0.1598	0.8731
beta	0.0000	0.0000	-0.8199	0.4130

ret	0.0000	0.0001	-0.0373	0.9703
alco:vit.rar.2	0.0273	0.0122	2.2452	0.0255
alco:age	0.0000	0.0004	0.0543	0.9568
alco:male.2	0.0106	0.0122	0.8704	0.3848
alco:smo.nev.2	-0.0150	0.0146	-1.0271	0.3052
alco:smo.for.2	-0.0022	0.0127	-0.1761	0.8603
alco:quet	0.0017	0.0011	1.6352	0.1031
alco:vit.oft.2	0.0203	0.0107	1.8883	0.0600
alco:fat	-0.0001	0.0002	-0.5987	0.5498
alco:fib	0.0013	0.0014	0.9428	0.3466
alco:chol	0.0000	0.0001	-0.4150	0.6784
alco:beta	0.0000	0.0000	-0.3558	0.7222
alco:ret	0.0000	0.0000	0.3391	0.7348

M ANOVA Table: Lretp, Alcohol Interaction

	Df	Sum of Sq	Mean Sq	F Value	Pr(F)
age	1	1.81568	1.815682	17.22179	0.0000438
alco	1	1.52247	1.522467	14.44064	0.0001765
vit.rar.2	1	0.03983	0.039826	0.37775	0.5392961
male.2	1	0.05510	0.055095	0.52258	0.4703299
smo.nev.2	1	0.13809	0.138095	1.30983	0.2533751
smo.for.2	1	0.20330	0.203299	1.92830	0.1660174
quet	1	0.00955	0.009552	0.09060	0.7636285
vit.oft.2	1	0.08285	0.082847	0.78581	0.3761095
fat	1	0.49873	0.498726	4.73043	0.0304470
fib	1	0.06683	0.066828	0.63387	0.4265955
chol	1	0.03397	0.033969	0.32220	0.5707323
beta	1	0.06596	0.065960	0.62564	0.4296114
ret	1	0.00067	0.000673	0.00639	0.9363589
alco:vit.rar.2	1	0.16829	0.168289	1.59622	0.2074609
alco:age	1	0.00278	0.002779	0.02635	0.8711506
alco:male.2	1	0.20793	0.207933	1.97225	0.1612872
alco:smo.nev.2	1	0.02975	0.029750	0.28218	0.5956840
alco:smo.for.2	1	0.00102	0.001021	0.00968	0.9216922
alco:quet	1	0.14284	0.142844	1.35488	0.2453900

alco:vit.oft.2	1	0.32065	0.320648	3.04135	0.0822352
alco:fat	1	0.06836	0.068361	0.64840	0.4213486
alco:fib	1	0.11006	0.110057	1.04389	0.3077756
alco:chol	1	0.01441	0.014409	0.13667	0.7118838
alco:beta	1	0.01264	0.012644	0.11993	0.7293647
alco:ret	1	0.01213	0.012125	0.11501	0.7347592
Residuals	288	30.36364	0.105429		

This table yielded an F statistic of 0.69 on 11 and 289 degrees of freedom, which has a p-value of 0.75.

N Full Regression Results: Final Lretp Model

	Value	Std. Error	t value	Pr(>|t|)
(Intercept)	6.1033	0.1299	46.9713	0.0000
age	0.0043	0.0014	3.1013	0.0021
alco	0.0104	0.0047	2.2325	0.0263
male.2	0.0812	0.0627	1.2952	0.1963
smo.nev.2	0.0176	0.0598	0.2945	0.7686
smo.for.2	0.0917	0.0606	1.5128	0.1314
quet	0.0019	0.0032	0.6016	0.5479
vit.oft.2	0.0437	0.0443	0.9863	0.3248
vit.rar.2	0.0215	0.0539	0.3978	0.6911
fat	-0.0008	0.0008	-0.9983	0.3189
fib	-0.0013	0.0042	-0.3138	0.7539
chol	-0.0001	0.0002	-0.4990	0.6181
beta	0.0000	0.0000	-0.8517	0.3951
ret	0.0000	0.0000	-0.0710	0.9434
alco:vit.rar.2	0.0108	0.0085	1.2682	0.2057