# Longitudinal Effects of Beta Carotene Supplementation on Serum Levels

Dustin Pluta

## Contents

# 1 Introduction

Despite the many substantial improvements in cancer treatments, lung cancer remains a significant source of many deaths in the United States. The standard treatments of resection, chemotherapy, and radiotherapy have improved dramatically over time, but are still painful, expensive, and debiltating, and are by no means completely effective. Thus there is a continued need to better understand possible preventive measures and alternative treatment options, including environmental and dietary changes. Toward this end, a series of studies in the 1980s and 1990s sought to better understand the role and effects of beta carotene with respect to cancer risk.

An early review by Peto et al. (1981) considered evidence from a collection of prospective and retrospective studies that suggested an association between higher serum levels of beta-carotene and lower cancer incidence. Although the evidence was not conclusive, the need and hope of developing more effective cancer treatments prompted Peto et al. to propose and call for more targeted and controlled clinical trials to better determine the role and effects of beta carotene. [1]

A later review by Ziegler (1989) again found some weak evidence of a possible protective effect of beta carotene, largely on the evidence that lower serum beta carotene levels were consistently found to be associated with an increased incidence of lung cancer. Ziegler does however acknowledge that many of the relevant studies did not properly adjust for important possible confounders, such as smoking status and the other dietary benefits of eating fruits and vegetables that have high concentrations of beta carotene.

Continuing the effort to understand the effects of beta carotene, a study by Heinonen and Albanes (1994) examined both beta carotene and alpha-tocopherol (vitamin E) supplements on the incidence rates of lung cancer in male smokers. A total of 29,133 male smokers age 50 to 69 years were randomly assigned to supplements of vitamin E, beta carotene, a combination of both, or placebo, and followed for five to eight years. Surprisingly, the study observed an significant increase in the incidence of lung cancer among those subjects receiving beta carotene supplements. Additional cancers were also monitored, and no significant effect of beta carotene supplementation on the incidence rates of other cancers was observed.

A similar study by Omenn et a. (1996) also found possible adverse effects of beta carotene supplementation in smokers, former smokers, and workers exposed to asbestos, with a relative risk of lung cancer of 1.28 (95% CI 1.04, 1.57, $P = 0.02$) in the group receiving beta carotene supplments relative to the placebo group.

With this somewhat conflicting, but increasingly negative, evidence of beta carotene's effect on cancer incidence, additional trials considering the pharmokinetics of beta carotene supplementation in healthy non-smokers may help clinicians develop a more thorough understanding of the overall role and effects of beta carotene. The present study considers the data from a phase II trial designed to assess the impact of daily beta carotene supplementation on serum beta carotene (SBC) levels, as well as serum vitamin E levels. In particular, this study is motivated by four objectives: 1) Is supplementation of beta-carotene associated with a time-averaged increase in SBC? 2) Does beta-carotene supplementation affect the trend of SBC levels over time? 3) Does beta-carotene supplementation affect SVE levels over time? 4) Are SVE levels associated with SBC levels over time? In cases (1) - (3), in the presence of a significant effect, it is also of interest whether the effect is dose dependent.

# 2 Subjects and Methods

For this study, 46 volunteers were placed in one of five treatment groups by double-blind random assignment. All subjects were given a daily placebo for the first three months of the study, after which each subject received active daily beta carotene supplements in the amount of 15, 30, 45, or 60 mg/day, or placebo, depending on group assignment. Each month, SBC and serum vitamin E (SVE) levels were recorded. Other measured variables include age, gender, body mass index (BMI), and serum cholesterol level at baseline. Follow-ups were conducted for 15 months, although not all patients remained in the study for the entire

[1] Adding some context, it seems that, at the time of the article, developing cancer was such a salient fear that the editor of *Nature* felt the need to caution readers that the evidence of beta carotene benefits was still limited, and not yet sufficient to justify the consumption of large quantities of carrots.

duration. At the conclusion of the study, time averaged values of SBC and SVE were calculated for each subject as the area under the curve (AUC) of the longitudinal trend line for each serum measurement.
Two subjects had less than 3 months on study following the baseline period from months 0 - 3, and so were excluded from the analysis (although there are included in the summary statistics table, Table **??**tab:datasummary).

## 2.1 Objective 1 Inference Methods

Regarding the first objective, in order to test the effect of supplementation on time-averaged SBC (AUC), we start with a simple linear regression model with only a binary treatment indicator variable to test for an unadjusted relationship between treatment and mean-centered SBC AUC. We then extend this model to include the adjustment covariates for age, gender, BMI. Due to the randomized design, there should not be any substantial confounding effects, but including these covariates may improve precision of the resulting estimates by reducing the residual variance. It is reasonable to expect age, gender, and BMI to have a significant relationship with SBC, as these quantities are significantly associated with many physiologic characteristics. Since beta-carotene is fat-soluble, models which included cholesterol were also examined for the present study. However, cholesterol was not found to be significantly associated with SBC in those models, so we exclude it from consideration here and focus on the simpler adjustment models. After testing the binary treatment indicator, we test for significance of the supplement dosage level on SBC.
We code the dosage as a categorical rather than continuous variable in order to test and assess each dosage level separately.

Explicitly, the simple initial model (Model 1a) is

$$\begin{aligned} SBCAUC &\sim N(\mu, \Sigma) \qquad \text{(Model 1a)} \\ \mu &= \beta_0 + \beta_1 TRT, \\ \Sigma &= \sigma_\varepsilon^2 I_N. \end{aligned}$$

adjusted model for testing the effect of supplementation on the time-averaged SBC levels is

$$\begin{aligned} SBCAUC &\sim N(\mu, \Sigma) \qquad \text{(Model 1b)} \\ \mu &= \beta_0 + \beta_1 DOSE + \beta_2 AGE + \beta_3 MALE + \beta_4 BMI, \\ \Sigma &= \sigma_\varepsilon^2 I_N. \end{aligned}$$

## 2.2 Objective 2 Inference Methods

For the second objective, we wish to test whether beta-carotene supplementation affects the trend of beta-carotene levels over time. Examining the trends of the SBC levels over time, it is apparent that there is a significant jump in SBC for all treatment patients at 4 months, which is the first month of active treatment. Since we are interested in the trajectory of beta-carotene over time, we first calculate the baseline average for every patient using the SBC measures from months 0 - 3. Inspection the of the SBC trends shows that the baseline period measurements are nearly flat for the majority of subjects, and that there is a large, discontinuous jump at month 4 for those subjects on treatment. To address for this, we then subtract the baseline average from each subject's trend, and consider only measurements from months 4 - 15 for the model. This allows us to directly compare subjects' changes in SBC with respect to baseline, and avoids the issue of constructing a model that correctly accounts for the discontinuity at the 4th month.
Moreover, it is not clear how a subject's measurements in the baseline period will be related to the measurements during the treatment period, ledading to a risk of biasing the parameter estimates if we group the measurements from the entire study period.

In order to select appropriate random effects terms and covariance structure, we examine the trend plots, empirical covariance, and residuals from a fixed effects regression of SBC levels over time. This exploratory analysis suggests including subject specific random intercepts, and using an $AR(1)$ correlation structure. However, there is a marked change in trend at around Month 12, which is consistent across subjects in all

non-placebo treatment groups, indicating that a quadratic time term or a spline model with a knot at Month 12 should be used. These considerations lead to the following (unadjusted) linear mixed effects spline model

$$SBC_{i,t} = b_{0i} + \beta_0 + \beta_1 DOSE_i + \beta_2 t + \beta_3 DOSE_i * t + \beta_4 (t-12)_+ + \beta_5 DOSE_i * (t-12)_+ + \varepsilon_{it} \qquad \text{(Model 2a)}$$

where $t = 4, \cdots, T_i$ for $T_i$ the last month on study for subject $i$.

Examining the trend plots, it also appears that the variance increases after Month 12.
To reduce the impact of this heteroskedasticity, we also consider an adjusted log transformation to stabilize the variance. Specifically, letting $c$ be the magnitude of the minimum SBC across all patients after baseline normalization, we use the transformation

$$\log(SBC_{it} + c + 1) = \mu_i(t), \qquad \text{(Model 2b)}$$

where $\mu_i(t)$ follows one of the previous models. This transformation is effective at mitigating the change in variance, but unfortunatley changes the interpretation and implications of the model effects. We include consideration of this model, with the caveat that there is some dissent in the literature over the use of non-linear transformations in mixed effects models. Future work may instead consider using a generalized linear mixed model to address the heteroskedasticity.

The results of adjusting Model 2b with the covariates for age, gender, and BMI, are also provided below (labeled Model 2c).

To compare the effectiveness of the mixed effects model, we also fit a fixed effects spline model using ordinary least squares, and then adjust for the correlation in the data using the robust variance estimator given by

$$\hat{\Sigma} = (X^T X)^{-1} X^T V X (X^T X)^{-1},$$

where $V$ is the diagonal matrix of squared residuals from the OLS fit. This has the advantage that it does not require correct specification of the covariance structure, and should produce accurate inference given a sufficiently large sample size.

## 2.3 Objective 3 Inference Methods

Due to the few models are somewhat complicated relative to the small number of subjects in the data set, we also consider a fixed effects linear regression

For determining whether supplementation affects SVE levels over time, we follow the same exploratory procedures as for Objective 2. The SVE levels over time exhibit a shallow, roughly quadratic, trend for the treatment groups, and are mostly flat for the control group. The sharp bend at Month 12 seen for SBC levels is absent here, thus we adopt a quadratic time model rather the spline model. The primary model for baseline normalized SVE levels is

$$\begin{aligned}
&SVE_{i,t} = b_{0i} + \beta_0 + \beta_1 DOSE_i + \beta_2 t + \beta_3 DOSE * t + \beta_4 t^2 + \beta_5 DOSE_i * t^2 + \varepsilon_{it} \\
&b_{0i} \sim N(0, \sigma_b^2) \qquad \text{(Model 3a)} \\
&\varepsilon_i \sim \mathcal{N}_{T_i}(0, \sigma^2 I_{T_i}) \\
&Cov(SVE_{it}, SVE_{it'}) = \sigma^2 \rho^{|t-t'|},
\end{aligned}$$

where $t = 1, \cdots, T_i$ as above. The adjusted version of the model is given below as Model 3b.

### 2.4 Objective 4 Inference Methods

Examining the sphaghetti plots of the SBC and SVE trends, a common downward trend over time is evident for both values in the treatment groups, particularly for the Doses 15-45 (Dose 60 seems to turn up near the end of the study). For this association, instead of relying on the choice of a sensible correlation structure and model, which may be complicated by the presence of SBC as a covariate, which is
itself time-varying and measured with error, we can fit a simpler fixed effects linear model and estimate the confidence interval of interest using the robust variance estimator (Huber 2005). Specifically, we fit the linear regression

$$SVE \sim SBC + MONTH + AGE + MALE + BMI, \qquad \text{(Model 4)}$$

and use the robust variance estimator as above.

For an alternative perspective, we can attempt to test the linear association over time using the SBC and SVE trend data with a block permutation procedure, wherein we permute the subject labels on the SBC vector and again compute the correlation with the SVE vector for each permutation. Permuting subject labels instead of individual time points retains the correlation structure between measures from a single subject while still generating a reference distribution under the null hypothesis of no association between SBC and SVE.
In order to simplify this calculation, we restrict the test to those subjects who did not drop out of the study early, which consists of 39 subjects.
This procedure should also be robust to non-trivial covariance structure in the data, but it will be underpowered relative to other methods for small sample sizes.

### 2.5 Computational Tools

All linear mixed effects models were fit using maximum likelihood procedures to facilitate the comparison of nested models.

Computations and modeling procedures were conducted in `R`. The `nlme` package was used for mixed effects modeling. Full code, documentation, and reproducible analyses are available on the author's GitHub.[2]

## 3 Results

### 3.1 Description of Study Sample

A summary description of the fixed variables in the data set is given in Table **??**, stratified by treatment group. Overall, we see the covariates are roughly balanced across the treatment groups. In particular, each group has, on average, at least 14 months on study per subject (the study duration was 16 months total).
The large difference in the SBC AUC levels between treatment and control groups strongly suggests that there is some relationship between beta carotene supplementation and SBC levels. This table also suggests that the treatment does not have a significant effect on time-averaged SVE levels.

### 3.2 Objective 1: Effects of Supplementation on Time-averaged Serum Beta Carotene Levels

The results for Model 1a are given in Table 2 below. This simple model gives a significant treatment effect of 932.45 ($95\%CI675.3, 1189.6$), indicating that, as expected, beta-carotene supplementation does produce a significantly higher time-averaged SBC.

Results for Model 1b are given in Table 3. This model includes treatment group as a factor, and the adjustment covariates age, gender, and BMI as adjustment.
As expected from the previous model, each dose level is individually significant. Testing the linear contrasts

[2] url: https://dspluta.github.com/DataAnalysisQual. Access can be requested at dpluta@uci.edu.

Table 1: Recorded patient characteristics by beta carotene dosage (in mg/day).

| | **Placebo** (N = 9) | **Dose 15** (N = 10) | **Dose 30** (N = 9) | **Dose 45** (N = 8) | **Dose 60** (N = 10) |
|---|---|---|---|---|---|
| Covariate | N or Mean (SD) | N or Mean(SD) | N or Mean(SD) | N or Mean(SD) | N or Mean(SD) |
| Months on Study | 14.33 (2.87) | 14.30 (3.65) | 15.78 (0.44) | 14.88 (2.32) | 14.40 (3.78) |
| SBC AUC | 384.19 (457.61) | 1131.81 (319.87) | 1309.30 (273.37) | 1324.29 (297.26) | 1522.56 (249.53) |
| SVE AUC | 7.98 (1.20) | 7.97 (0.86) | 8.23 (1.11) | 8.04 (0.49) | 8.35 (0.74) |
| Age | 56.11 (4.04) | 56.30 (4.64) | 57.44 (4.25) | 55.88 (3.14) | 56.50 (5.21) |
| Gender | | | | | |
| Female | 4 | 5 | 6 | 4 | 5 |
| Male | 5 | 5 | 3 | 4 | 5 |
| BMI | 26.18 (3.59) | 25.69 (3.58) | 25.83 (2.66) | 25.35 (3.32) | 24.94 (2.43) |
| Cholesterol | 216.00 (27.17) | 223.00 (29.72) | 214.44 (35.28) | 213.31 (33.54) | 238.05 (38.88) |

Table 2: Estimates from Model 1a

| | Est. | CI95 Lo | CI95 Hi | t value | P(>\|t\|) |
|---|---|---|---|---|---|
| Intercept | -745.9585 | -975.9861 | -515.931 | -6.5399 | < 1e-04 |
| Trt | 932.4481 | 675.2695 | 1189.6268 | 7.3119 | < 1e-04 |

for differences in dose effects, only D0se 15 and Dose 60 are significantly different, with an estimated difference of $-340.8(95\%CI - 622.7, -58.9)$.

### 3.3 Objective 2: Effects of Supplementation on Serum Beta Carotene Levels Over Time

Table **??** shows a highly significant effect of supplementation on SBC levels over time, averaged across all dose groups ($P < 3.6 \times 10^{-6}$), but does not show a significant interaction between treatment over time. This is likely due to the sharp change in the SBC trends around month 12. To adequately capture this change, Model 2b uses a first order spline with a single knot at month 12 (Table **??**). The treatment effect and interaction of the treatment with time prior to the knot are both highly significant; after month 12, the time effect itself is strongly significant ($P < 10^{-4}$), but the interaction with the treatment is not.

```
##                 Estimate  Robust SE     ci95.lo     ci95.hi    t value
## (Intercept)    6.2986378 0.29076821  5.72732562  6.86994992 21.6620580
```

Table 3: Estimates from Model 1b

| | Est. | CI95 Lo | CI95 Hi | t value | P(>\|t\|) |
|---|---|---|---|---|---|
| Intercept | 244.7417 | -1629.6794 | 2119.1628 | 0.2646 | 0.7928 |
| Dose 15 | 721.2984 | 443.1789 | 999.4179 | 5.2549 | < 1e-04 |
| Dose 30 | 874.4165 | 585.422 | 1163.411 | 6.1307 | < 1e-04 |
| Dose 45 | 902.3293 | 607.1775 | 1197.481 | 6.1944 | < 1e-04 |
| Dose 60 | 1062.0844 | 770.6271 | 1353.5417 | 7.3835 | < 1e-04 |
| Age | -0.187 | -24.6362 | 24.2623 | -0.0155 | 0.9877 |
| Male | -176.546 | -368.9619 | 15.8698 | -1.8591 | 0.071 |
| BMI | -33.6978 | -67.943 | 0.5474 | -1.9938 | 0.0536 |

Table 4: Model 4 results, testing association of SBC with SVE over time.

| | Est. | CI95 Lo | CI95 Hi | t value | P(>\|t\|) |
|---|---|---|---|---|---|
| Intercept | 4.7919 | 2.5055 | 7.0784 | 4.1179 | 0.0000 |
| SBC | -0.0003 | -0.0005 | -0.0001 | -3.0204 | 0.0027 |
| Male | -0.0639 | -0.3112 | 0.1834 | -0.5079 | 0.6117 |
| BMI | 0.0027 | -0.0408 | 0.0462 | 0.1214 | 0.9034 |
| Age | -0.0292 | -0.0568 | -0.0017 | -2.0855 | 0.0375 |
| Month | -0.3707 | -0.4064 | -0.3350 | -20.3890 | 0.0000 |

```
## month         -0.1048783 0.03493040 -0.17351085 -0.03624576 -3.0024937
## trtTRUE        1.3965758 0.29806095  0.81093455  1.98221696  4.6855375
## month:trtTRUE  0.0325285 0.03576859 -0.03775095  0.10280796  0.9094152
##                    Pr(>|t|)
## (Intercept)   2.071230e-73
## month         2.815125e-03
## trtTRUE       3.626389e-06
## month:trtTRUE 3.635799e-01
```

Testing for the dose effect in the spline model (Table 5), we find that all of the dose main effects are very significant (all $P < 10^{-8}$). The temporal main effect is moderately significant before and after the knot, with $P \approx 0.03$, and the time-dose interaction is also moderately significant prior to the knot.
The interactions are not significant after month 12.

For comparison, we also fit a mixed effects model version of the spline model, with an $AR(1)$ correlation structure (fixed effect estimates given in Table **??**).
The corresponding treatment group estimates match those of the fixed model in terms of significance. This is somewhat expected, but the validity of these estimates still rely on a correctly specified covariance structure, and is an inherently more complex model compared the fixed effects version, so it is not clear if applying the mixed effects model is required or justified for this data set.

## 3.4 Objective 3: Effects of Supplementation on Serum Vitamin E Levels Over Time

Similar to the results and methods of Objective 2, we consider both a mixed effects and fixed effects model. Results are given in @ref(tab:fit_3b).

## 3.5 Objective 4: Association of Serum Vitamin E with Serum Beta Carotene Over Time

The robust variance estimates for Model 4 are provided in Table 4.
Also fit but not shown here was the reduced model $SVE \sim SBC$, which did not yield a significant association ($P = 0.13$). However, from the 95% CI for the effect of SBC on SVE in Model 4 of $(-0.0001, -3.02)$, it is questionable whether this effect is practically significant. It should also be noted that, as with the above models, the data used here was from only the treatment period after month 3; when including the entire time series, the $P$-value for SBC drops to 0.018, which is still significant, but much reduced. Considering the trends of the treatment groups in the plots for SBC and SVE, it seems possible that the supplementation is actually driving this relationship by somehow causing both levels to drop after approximately a year of daily supplements.

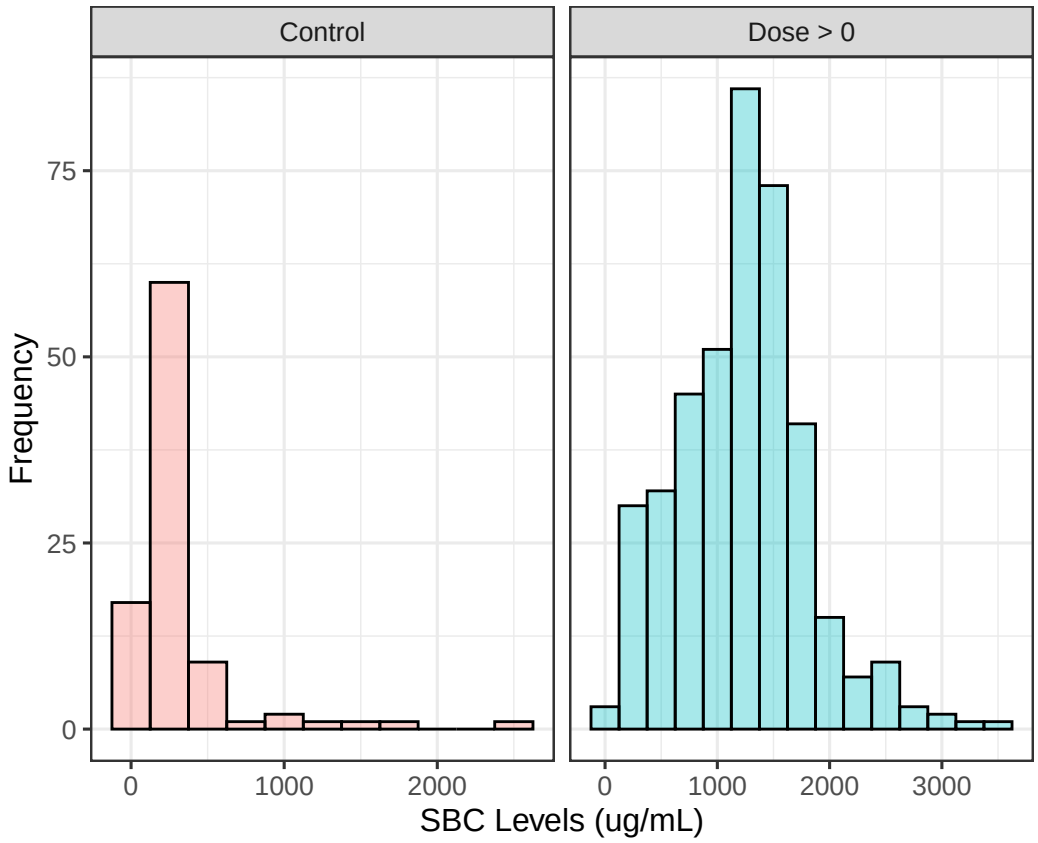
# 4 Discussion

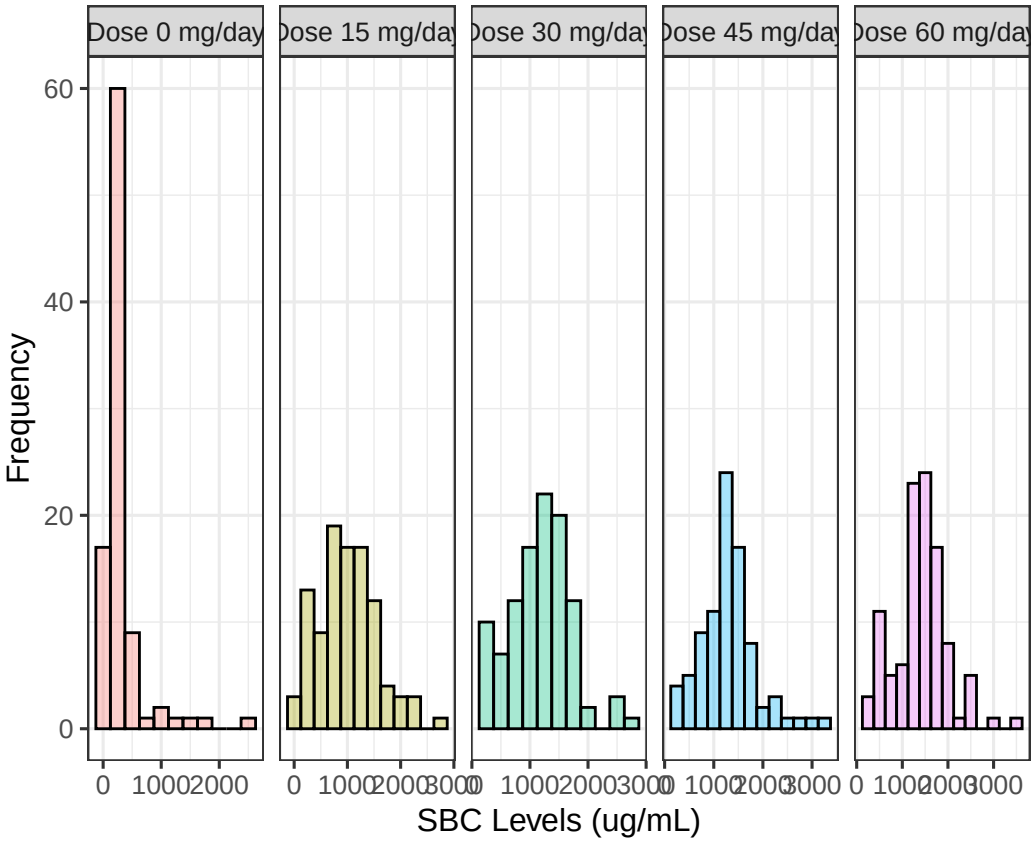
Given the results of the present analysis, there seems to be mounting evidence that supplementation with beta carotene can have serious effects on body chemistry, including an eventual decrease in overall serum beta carotene levels after a year of daily supplements. This effect was observed across all dose levels. This drop in beta carotene was also found to be significantly associated with to drops in serum vitamin E levels.

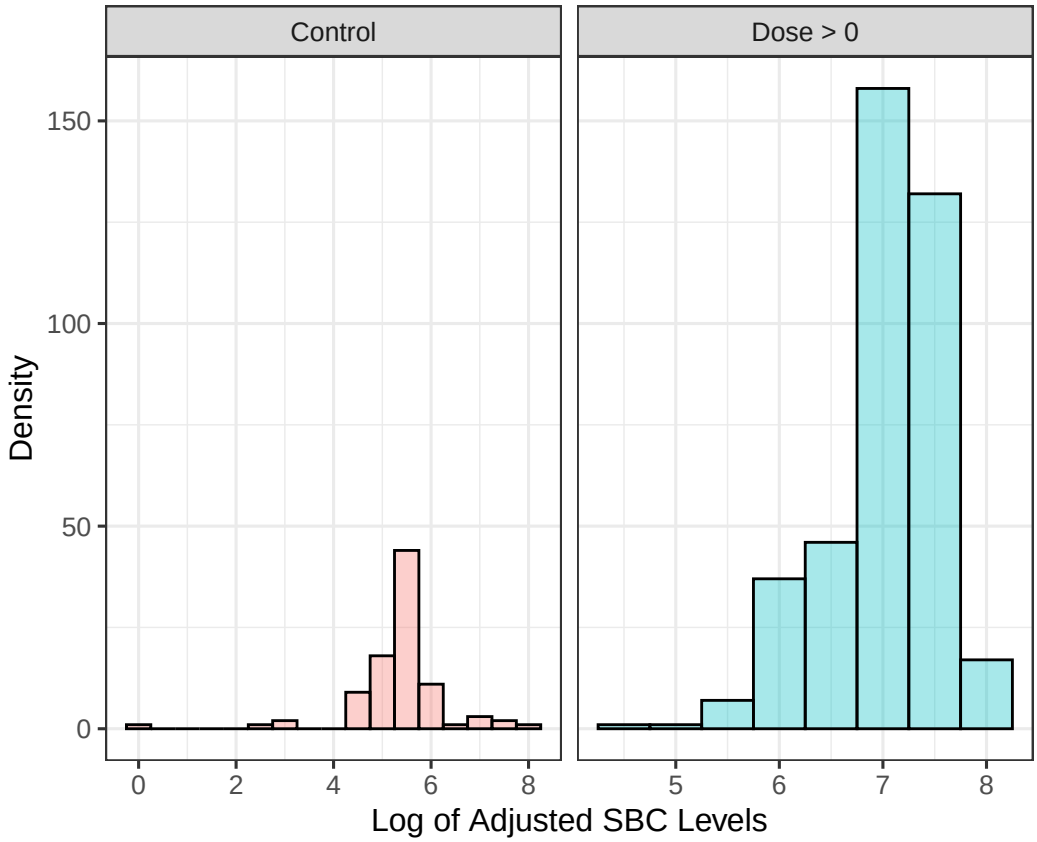
While further studies may be merited to better understand the mechanisms by which the beta carotene supplements are adversely affecting the serum levels, it seems clear that regular beta carotene supplements should not be advised for the average healthy adult.

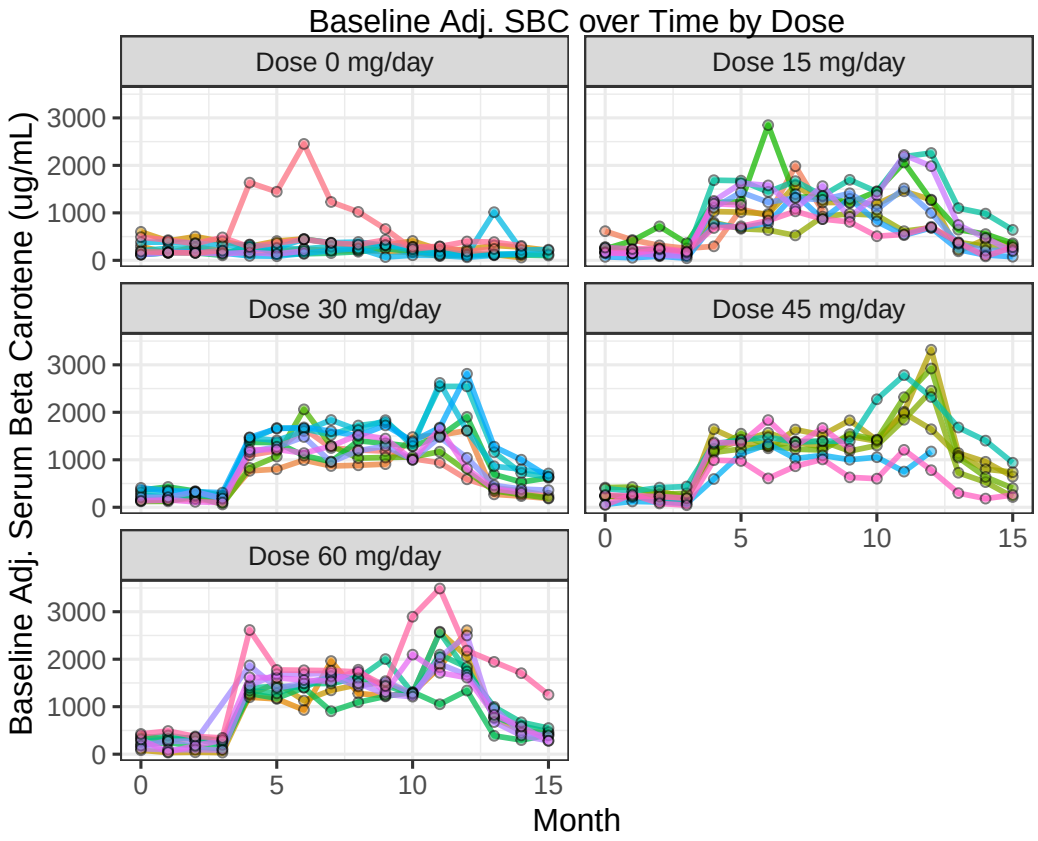
# 5 Figures

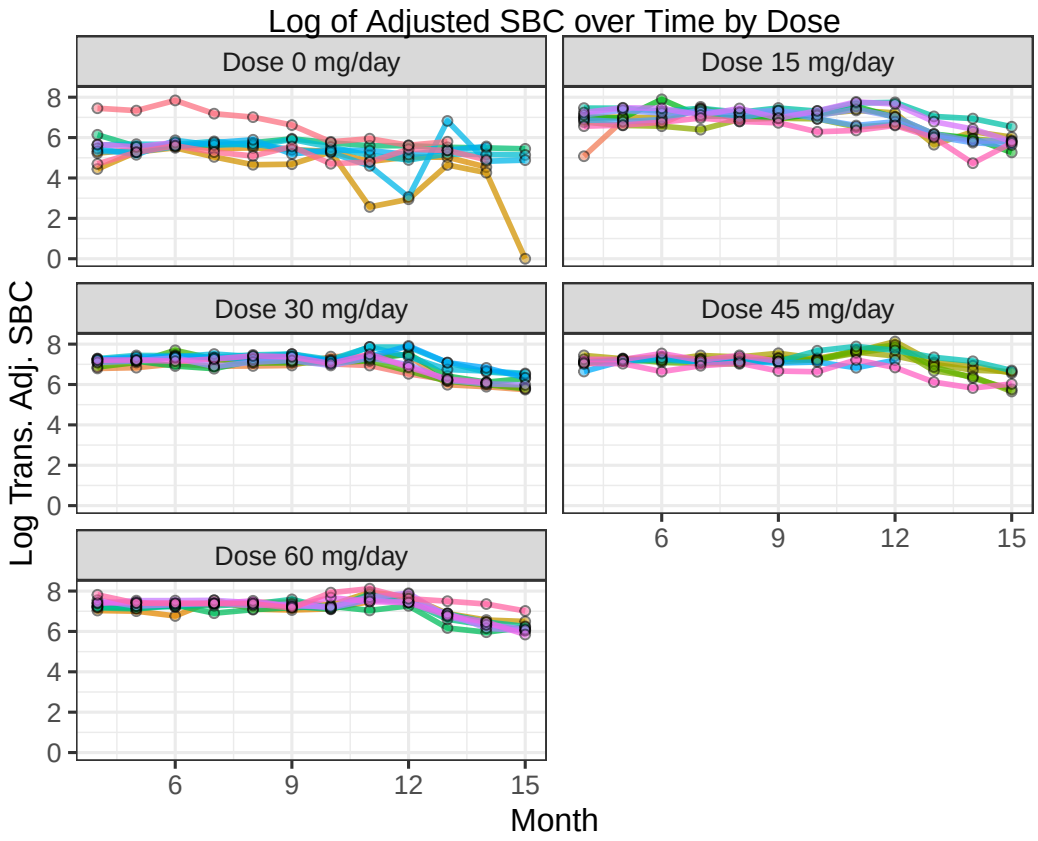
## 5.1 Descriptive Plots

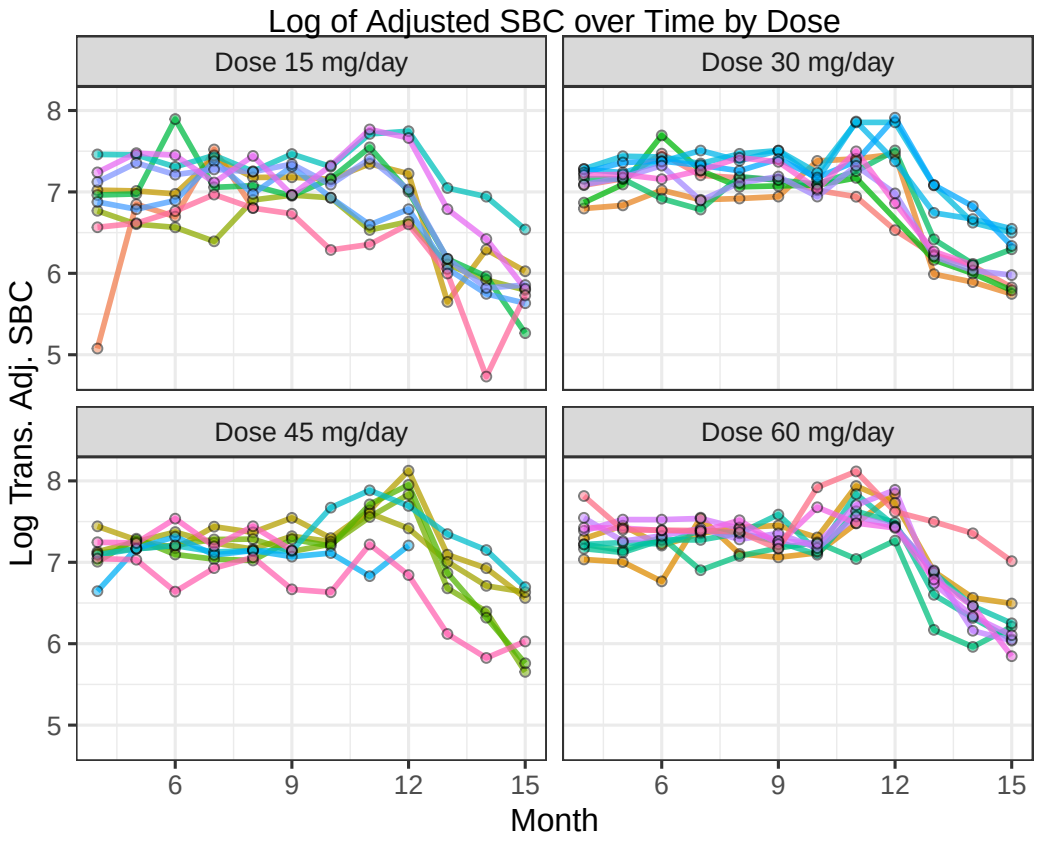


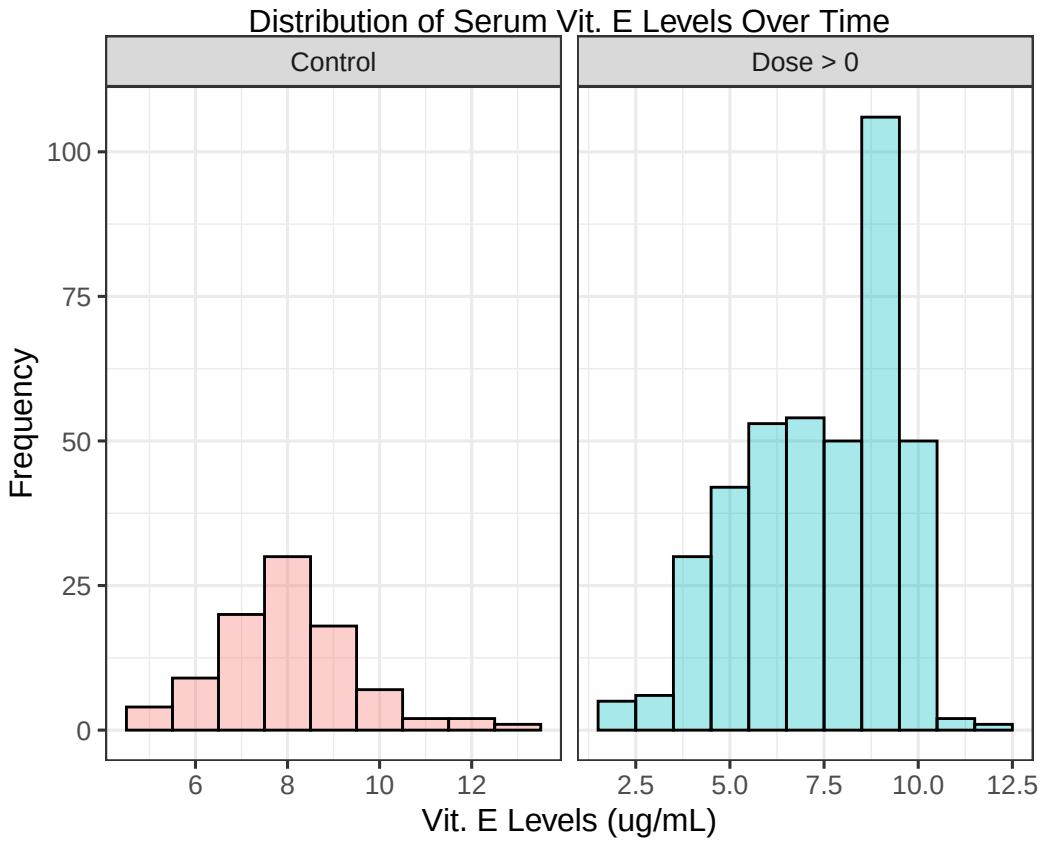

Dose 0 mg/day
Dose 15 mg/day
Dose 30 mg/day
Dose 45 mg/day
Dose 60 mg/day
Frequency
0
20
40
60
0
1000
2000
3000
SBC Levels (ug/mL)

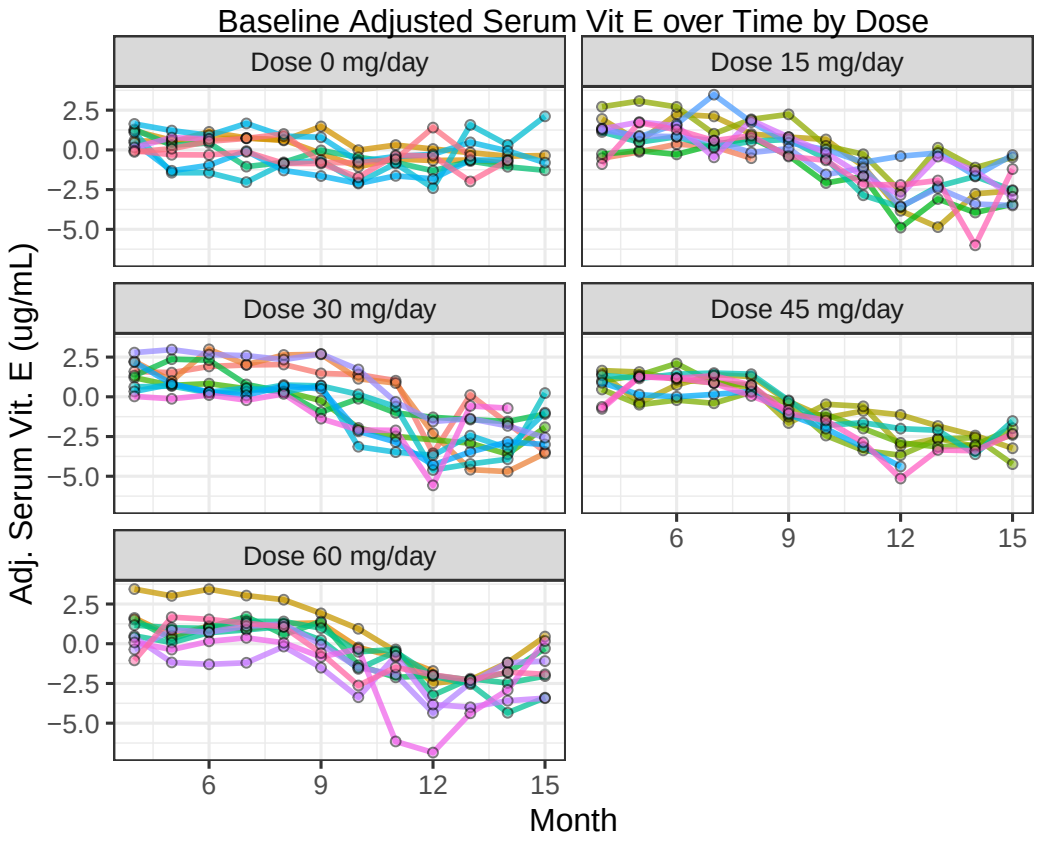

Control
Dose > 0
Density
150
100
50
0
0
2
4
6
8
5
6
7
8
Log of Adjusted SBC Levels


Baseline Adj. SBC over Time by Dose
Dose 0 mg/day
Dose 15 mg/day
Dose 30 mg/day
Dose 45 mg/day
Dose 60 mg/day
Baseline Adj. Serum Beta Carotene (ug/mL)
3000
2000
1000
0
0
5
10
15
Month

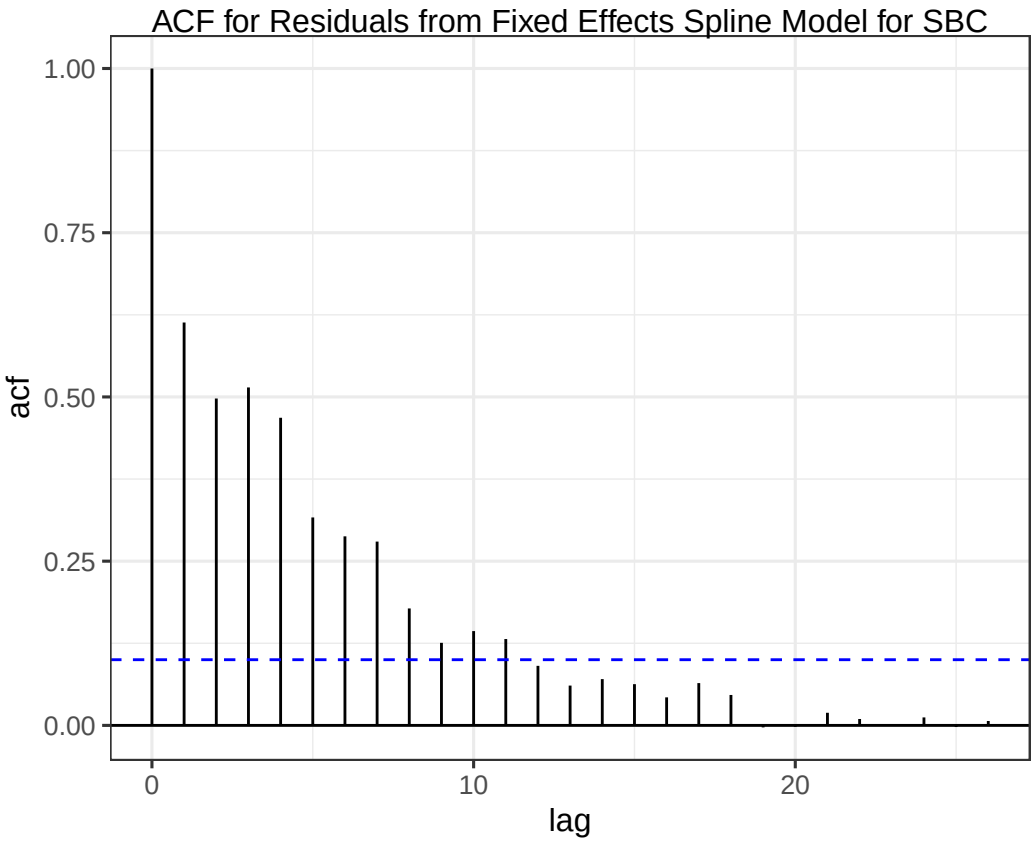

Log of Adjusted SBC over Time by Dose
Dose 0 mg/day
Dose 15 mg/day
Dose 30 mg/day
Dose 45 mg/day
Dose 60 mg/day
Log Trans. Adj. SBC
Month
0
2
4
6
8
6
9
12
15

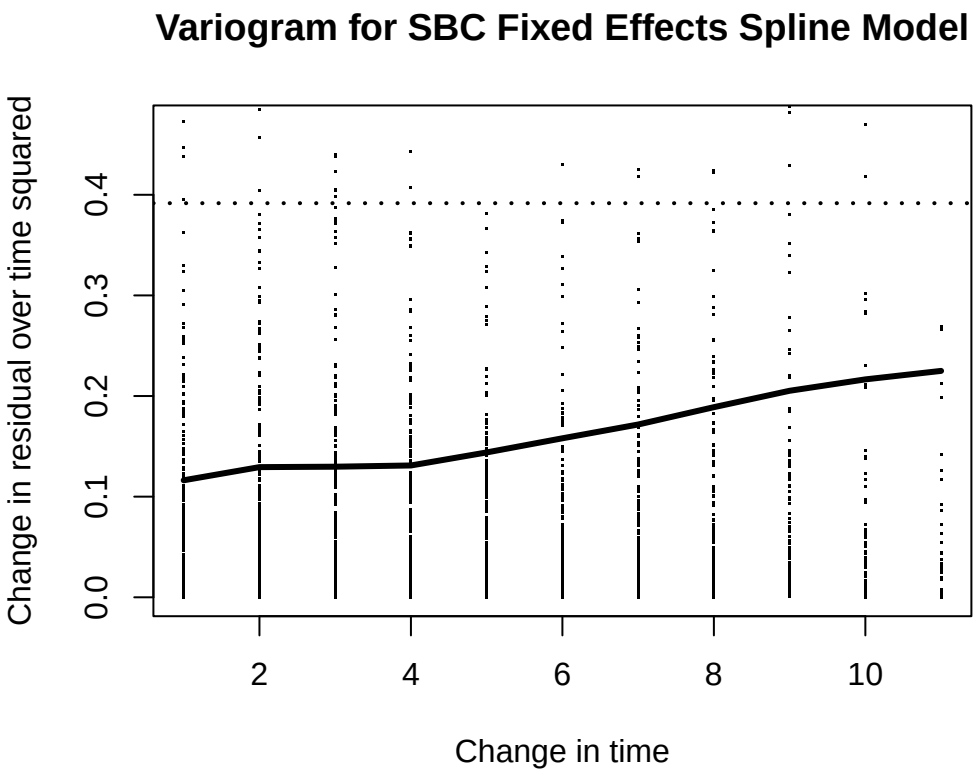

Log of Adjusted SBC over Time by Dose
Dose 15 mg/day
Dose 30 mg/day
Dose 45 mg/day
Dose 60 mg/day
Log Trans. Adj. SBC
8
7
6
5
6
9
12
15
Month


Distribution of Serum Vit. E Levels Over Time
Control
Dose > 0
Frequency
100
75
50
25
0
6
8
10
12
2.5
5.0
7.5
10.0
12.5
Vit. E Levels (ug/mL)

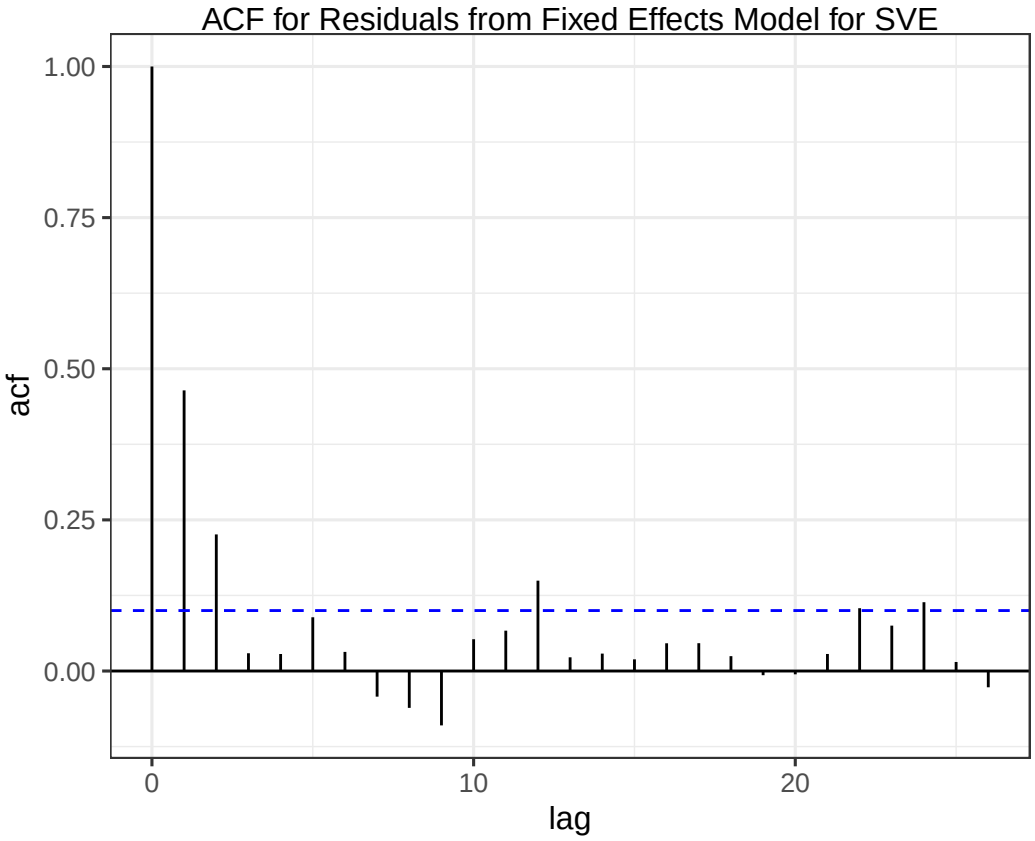

Baseline Adjusted Serum Vit E over Time by Dose
Dose 0 mg/day
Dose 15 mg/day
Dose 30 mg/day
Dose 45 mg/day
Dose 60 mg/day
Adj. Serum Vit. E (ug/mL)
2.5
0.0
−2.5
−5.0
6
9
12
15
Month

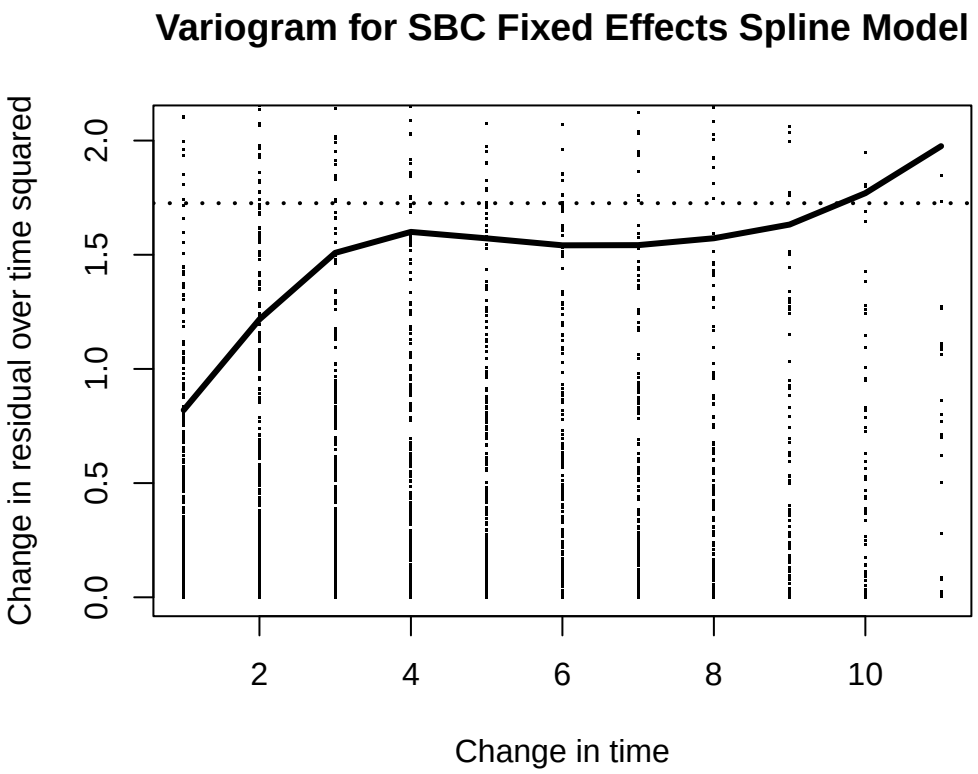
### 5.2 Figures: Objective 2



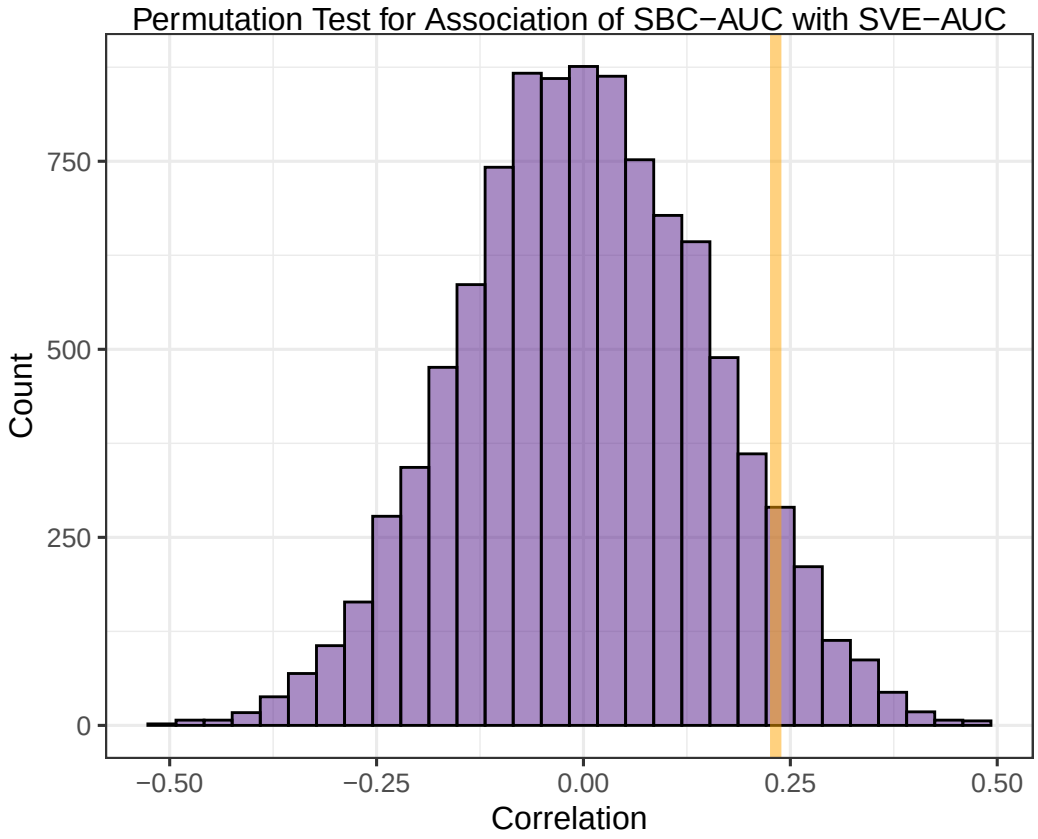

Variogram for SBC Fixed Effects Spline Model
Change in residual over time squared
0.0
0.1
0.2
0.3
0.4
2
4
6
8
10
Change in time

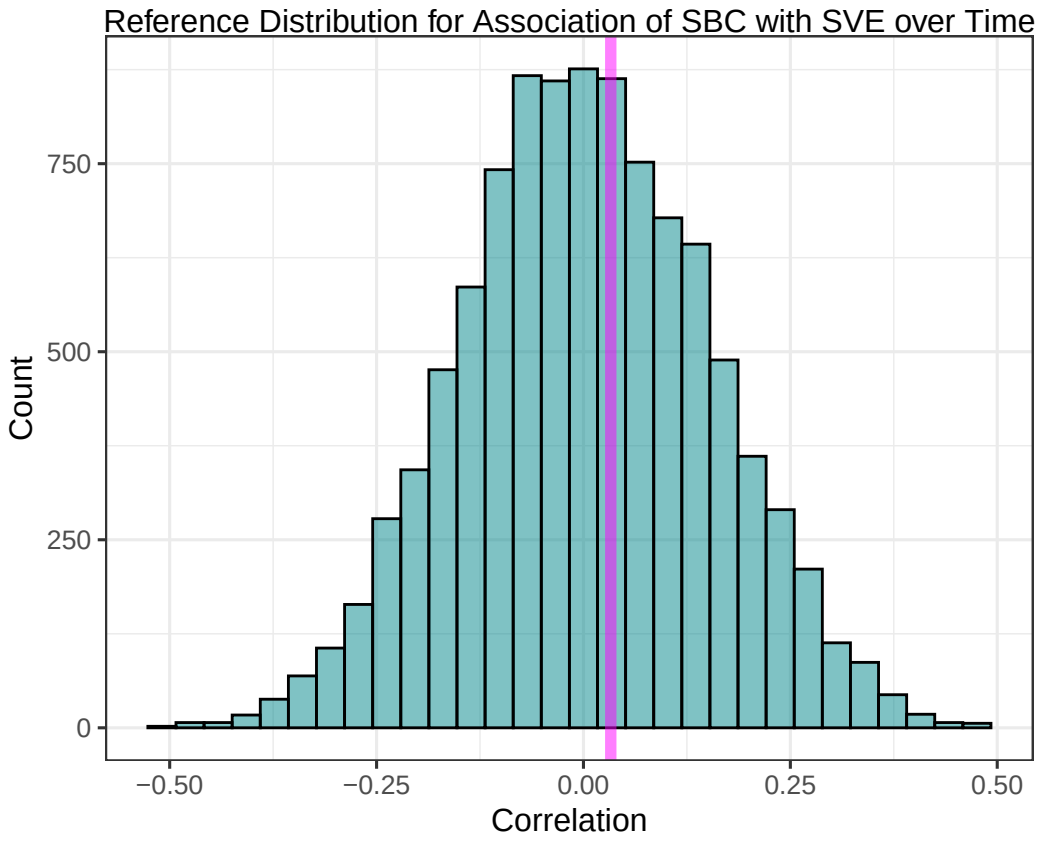
### 5.3 Figures: Objective 3




Variogram for SBC Fixed Effects Spline Model
Change in residual over time squared
0.0
0.5
1.0
1.5
2.0
2
4
6
8
10
Change in time

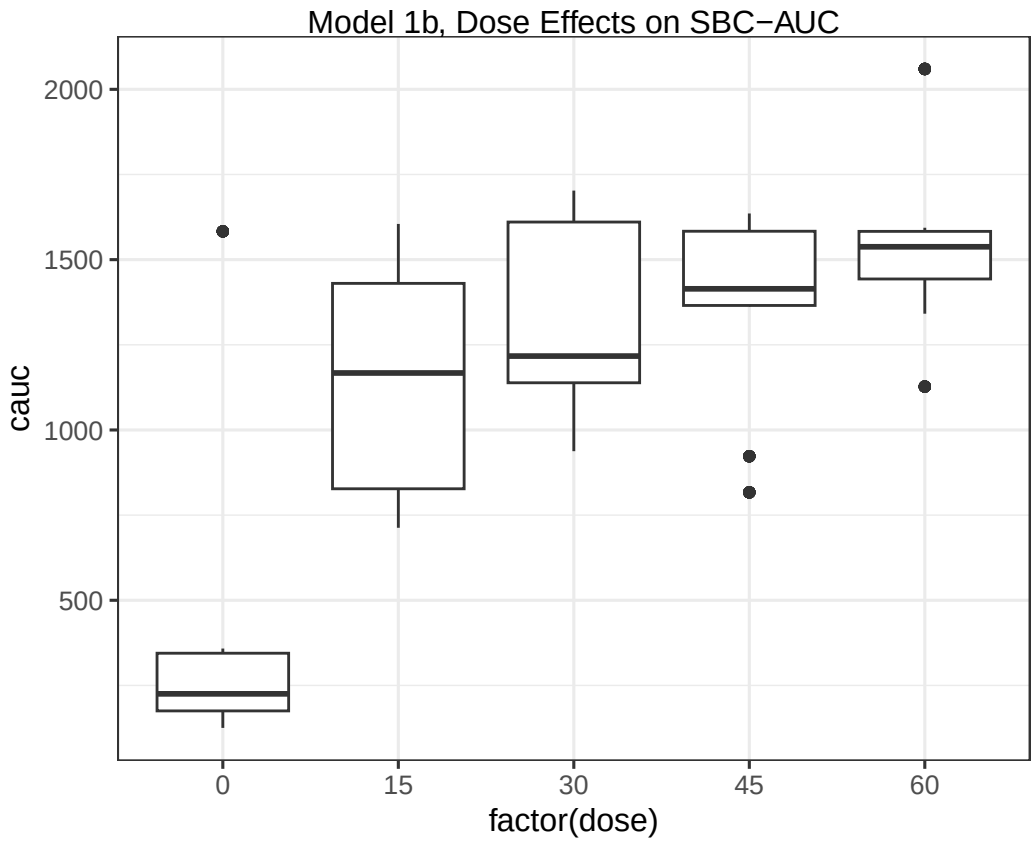
### 5.4 Figures: Objective 4




Reference Distribution for Association of SBC with SVE over Time
Count
750
500
250
0
−0.50
−0.25
0.00
0.25
0.50
Correlation

# 6 Appendix

## 6.1 Additional Figures




# 7 Additional Tables

```
##
## Call:
## lm(formula = bcarot_trans ~ trt + trt * bs(month, degree = 1,
##     knots = c(12)), data = dat_post)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -4.5049 -0.1853  0.0285  0.2309  2.1894
##
## Coefficients:
##                                                 Estimate Std. Error t value
## (Intercept)                                       5.8176     0.1044  55.715
## trtTRUE                                           1.3185     0.1168  11.285
## bs(month, degree = 1, knots = c(12))1            -0.6658     0.1727  -3.856
## bs(month, degree = 1, knots = c(12))2            -1.3127     0.2042  -6.430
## trtTRUE:bs(month, degree = 1, knots = c(12))1     0.8127     0.1930   4.212
## trtTRUE:bs(month, degree = 1, knots = c(12))2     0.1204     0.2225   0.541
##                                                 Pr(>|t|)
## (Intercept)                                      < 2e-16 ***
```

```
## trtTRUE                                          < 2e-16 ***
## bs(month, degree = 1, knots = c(12))1          0.000131 ***
## bs(month, degree = 1, knots = c(12))2          3.06e-10 ***
## trtTRUE:bs(month, degree = 1, knots = c(12))1 3.02e-05 ***
## trtTRUE:bs(month, degree = 1, knots = c(12))2 0.588711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5086 on 486 degrees of freedom
## Multiple R-squared:  0.6925, Adjusted R-squared:  0.6893
## F-statistic: 218.9 on 5 and 486 DF,  p-value: < 2.2e-16
##
##                                                Estimate Robust SE    ci95.lo
## (Intercept)                                   5.8176465 0.1629849  5.4974045
## trtTRUE                                       1.3184587 0.1668348  0.9906522
## bs(month, degree = 1, knots = c(12))1        -0.6658147 0.3175650 -1.2897845
## bs(month, degree = 1, knots = c(12))2        -1.3127339 0.6189449 -2.5288722
## trtTRUE:bs(month, degree = 1, knots = c(12))1  0.8126836 0.3246093  0.1748726
## trtTRUE:bs(month, degree = 1, knots = c(12))2  0.1204039 0.6232349 -1.1041637
##                                                   ci95.hi    t value
## (Intercept)                                   6.13788855 35.6943939
## trtTRUE                                       1.64626520  7.9027819
## bs(month, degree = 1, knots = c(12))1        -0.04184490 -2.0966252
## bs(month, degree = 1, knots = c(12))2        -0.09659571 -2.1209223
## trtTRUE:bs(month, degree = 1, knots = c(12))1  1.45049452  2.5035744
## trtTRUE:bs(month, degree = 1, knots = c(12))2  1.34497143  0.1931918
##                                                    Pr(>|t|)
## (Intercept)                                   6.544560e-138
## trtTRUE                                        1.841064e-14
## bs(month, degree = 1, knots = c(12))1          3.654310e-02
## bs(month, degree = 1, knots = c(12))2          3.443407e-02
## trtTRUE:bs(month, degree = 1, knots = c(12))1  1.262143e-02
## trtTRUE:bs(month, degree = 1, knots = c(12))2  8.468895e-01
```

```
##                                                          Estimate Robust SE
## (Intercept)                                           5.81764654 0.1629849
## factor(dose)15                                        1.18198432 0.1931985
## factor(dose)30                                        1.37551122 0.1694298
## factor(dose)45                                        1.26252690 0.1697964
## factor(dose)60                                        1.44867733 0.1685527
## bs(month, degree = 1, knots = c(12))1                -0.66581472 0.3175650
## bs(month, degree = 1, knots = c(12))2                -1.31273395 0.6189449
## factor(dose)15:bs(month, degree = 1, knots = c(12))1  0.72460787 0.3604161
## factor(dose)30:bs(month, degree = 1, knots = c(12))1  0.68667963 0.3345214
## factor(dose)45:bs(month, degree = 1, knots = c(12))1  0.99230398 0.3375252
## factor(dose)60:bs(month, degree = 1, knots = c(12))1  0.85351192 0.3288451
## factor(dose)15:bs(month, degree = 1, knots = c(12))2 -0.05263819 0.6403396
## factor(dose)30:bs(month, degree = 1, knots = c(12))2  0.05233370 0.6285679
## factor(dose)45:bs(month, degree = 1, knots = c(12))2  0.39387653 0.6358637
## factor(dose)60:bs(month, degree = 1, knots = c(12))2  0.15900045 0.6293105
##                                                          ci95.lo     ci95.hi
## (Intercept)                                           5.49738944  6.13790364
## factor(dose)15                                        0.80235890  1.56160974
## factor(dose)30                                        1.04259020  1.70843224
```

Table 5: Model 4 results, testing association of SBC with SVE over time.

| | Est. | CI95 Lo | CI95 Hi | t value | P(>\|t\|) |
|---|---|---|---|---|---|
| Intercept | 5.8176 | 5.4974 | 6.1379 | 35.6944 | 0.0000 |
| Dose 15 | 1.1820 | 0.8024 | 1.5616 | 6.1180 | 0.0000 |
| Dose 30 | 1.3755 | 1.0426 | 1.7084 | 8.1185 | 0.0000 |
| Dose 45 | 1.2625 | 0.9289 | 1.5962 | 7.4355 | 0.0000 |
| Dose 60 | 1.4487 | 1.1175 | 1.7799 | 8.5948 | 0.0000 |
| Month < 12 | -0.6658 | -1.2898 | -0.0418 | -2.0966 | 0.0366 |
| Month >= 12 | -1.3127 | -2.5289 | -0.0965 | -2.1209 | 0.0344 |
| Dose 15 * Mo. < 12 | 0.7246 | 0.0164 | 1.4328 | 2.0105 | 0.0449 |
| Dose 30 * Mo. < 12 | 0.6867 | 0.0294 | 1.3440 | 2.0527 | 0.0406 |
| Dose 45 * Mo. < 12 | 0.9923 | 0.3291 | 1.6555 | 2.9399 | 0.0034 |
| Dose 60 * Mo. < 12 | 0.8535 | 0.2073 | 1.4997 | 2.5955 | 0.0097 |
| Dose 15 * Mo. >= 12 | -0.0526 | -1.3109 | 1.2056 | -0.0822 | 0.9345 |
| Dose 30 * Mo. >= 12 | 0.0523 | -1.1828 | 1.2874 | 0.0833 | 0.9337 |
| Dose 45 * Mo. >= 12 | 0.3939 | -0.8556 | 1.6433 | 0.6194 | 0.5359 |
| Dose 60 * Mo. >= 12 | 0.1590 | -1.0776 | 1.3956 | 0.2527 | 0.8006 |

```
## factor(dose)45                                          0.92888551  1.59616830
## factor(dose)60                                          1.11747978  1.77987489
## bs(month, degree = 1, knots = c(12))1                  -1.28981392 -0.04181551
## bs(month, degree = 1, knots = c(12))2                  -2.52892948 -0.09653842
## factor(dose)15:bs(month, degree = 1, knots = c(12))1    0.01640832  1.43280742
## factor(dose)30:bs(month, degree = 1, knots = c(12))1    0.02936181  1.34399745
## factor(dose)45:bs(month, degree = 1, knots = c(12))1    0.32908391  1.65552405
## factor(dose)60:bs(month, degree = 1, knots = c(12))1    0.20734786  1.49967599
## factor(dose)15:bs(month, degree = 1, knots = c(12))2   -1.31087327  1.20559690
## factor(dose)30:bs(month, degree = 1, knots = c(12))2   -1.18277070  1.28743810
## factor(dose)45:bs(month, degree = 1, knots = c(12))2   -0.85556359  1.64331664
## factor(dose)60:bs(month, degree = 1, knots = c(12))2   -1.07756312  1.39556402
##                                                            t value      Pr(>|t|)
## (Intercept)                                           35.69439390 8.486857e-137
## factor(dose)15                                         6.11797749  1.977056e-09
## factor(dose)30                                         8.11847334  4.056289e-15
## factor(dose)45                                         7.43553385  4.869997e-13
## factor(dose)60                                         8.59480423  1.200510e-16
## bs(month, degree = 1, knots = c(12))1                 -2.09662518  3.655286e-02
## bs(month, degree = 1, knots = c(12))2                 -2.12092230  3.444362e-02
## factor(dose)15:bs(month, degree = 1, knots = c(12))1   2.01047577  4.494321e-02
## factor(dose)30:bs(month, degree = 1, knots = c(12))1   2.05272232  4.064452e-02
## factor(dose)45:bs(month, degree = 1, knots = c(12))1   2.93994034  3.442124e-03
## factor(dose)60:bs(month, degree = 1, knots = c(12))1   2.59548327  9.736903e-03
## factor(dose)15:bs(month, degree = 1, knots = c(12))2  -0.08220355  9.345193e-01
## factor(dose)30:bs(month, degree = 1, knots = c(12))2   0.08325862  9.336808e-01
## factor(dose)45:bs(month, degree = 1, knots = c(12))2   0.61943551  5.359252e-01
## factor(dose)60:bs(month, degree = 1, knots = c(12))2   0.25265817  8.006412e-01
```

```
## Linear mixed-effects model fit by maximum likelihood
##   Data: dat_post
##        AIC      BIC    logLik
##   540.8082 578.5945 -261.4041
```

```
##
## Random effects:
##  Formula: ~1 | ptid
##         (Intercept)  Residual
## StdDev:   0.3358288 0.3897854
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~month | ptid
##  Parameter estimate(s):
##      Phi1
## 0.2428149
## Fixed effects:  bcarot_trans ~ trt + trt * bs(month, degree = 1, knots = c(12))
##                                                 Value Std.Error  DF  t-value
## (Intercept)                                  5.797473 0.1475491 444 39.29181
## trtTRUE                                      1.334230 0.1653384  42  8.06969
## bs(month, degree = 1, knots = c(12))1       -0.657207 0.1590729 444 -4.13148
## bs(month, degree = 1, knots = c(12))2       -1.160267 0.1816100 444 -6.38878
## trtTRUE:bs(month, degree = 1, knots = c(12))1  0.794588 0.1773820 444  4.47953
## trtTRUE:bs(month, degree = 1, knots = c(12))2 -0.016493 0.1983754 444 -0.08314
##                                                p-value
## (Intercept)                                     0.0000
## trtTRUE                                         0.0000
## bs(month, degree = 1, knots = c(12))1           0.0000
## bs(month, degree = 1, knots = c(12))2           0.0000
## trtTRUE:bs(month, degree = 1, knots = c(12))1   0.0000
## trtTRUE:bs(month, degree = 1, knots = c(12))2   0.9338
##  Correlation:
##                                                (Intr) trTRUE b(,d=1,k=c(12))1
## trtTRUE                                        -0.892
## bs(month, degree = 1, knots = c(12))1          -0.517  0.461
## bs(month, degree = 1, knots = c(12))2          -0.309  0.275  0.331
## trtTRUE:bs(month, degree = 1, knots = c(12))1   0.463 -0.519 -0.897
## trtTRUE:bs(month, degree = 1, knots = c(12))2   0.283 -0.319 -0.303
##                                                b(,d=1,k=c(12))2
## trtTRUE
## bs(month, degree = 1, knots = c(12))1
## bs(month, degree = 1, knots = c(12))2
## trtTRUE:bs(month, degree = 1, knots = c(12))1 -0.297
## trtTRUE:bs(month, degree = 1, knots = c(12))2 -0.915
##                                                tTRUE:(,d=1,k=c(12))1
## trtTRUE
## bs(month, degree = 1, knots = c(12))1
## bs(month, degree = 1, knots = c(12))2
## trtTRUE:bs(month, degree = 1, knots = c(12))1
## trtTRUE:bs(month, degree = 1, knots = c(12))2  0.348
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -9.09028803 -0.40772736  0.02702968  0.42851681  4.93112411
##
## Number of Observations: 492
## Number of Groups: 44

## Linear mixed-effects model fit by maximum likelihood
```

```
##   Data: dat_post
##        AIC      BIC    logLik
##   1522.252 1572.633 -749.1258
##
## Random effects:
##  Formula: ~1 | ptid
##         (Intercept) Residual
## StdDev:   0.2064235 1.295942
##
## Correlation Structure: ARMA(1,0)
##  Formula: ~month | ptid
##  Parameter estimate(s):
##      Phi1
## 0.5449465
## Fixed effects:  vite_norm ~ month + dose + month * dose + I(month^2) * dose +      male + age + bmi
##                     Value Std.Error  DF    t-value p-value
## (Intercept)      4.128967 2.3147496 444  1.7837640  0.0751
## month           -0.293885 0.2268838 444 -1.2953122  0.1959
## dose             0.015364 0.0262202  39  0.5859730  0.5613
## I(month^2)       0.003747 0.0119707 444  0.3129865  0.7544
## male             0.033702 0.2234218  39  0.1508454  0.8809
## age             -0.030076 0.0275979  39 -1.0897815  0.2825
## bmi             -0.013456 0.0393668  39 -0.3418054  0.7343
## month:dose      -0.002075 0.0061250 444 -0.3387531  0.7350
## dose:I(month^2) -0.000040 0.0003222 444 -0.1252007  0.9004
##  Correlation:
##                 (Intr) month  dose   I(m^2) male   age    bmi    mnth:d
## month           -0.412
## dose            -0.385  0.782
## I(month^2)       0.390 -0.982 -0.734
## male            -0.006 -0.014 -0.007  0.007
## age             -0.805  0.003  0.026 -0.007  0.131
## bmi             -0.621  0.026  0.060 -0.028 -0.290  0.269
## month:dose       0.341 -0.822 -0.955  0.808  0.017  0.001 -0.033
## dose:I(month^2) -0.326  0.811  0.895 -0.826 -0.015  0.001  0.038 -0.982
##
## Standardized Within-Group Residuals:
##         Min          Q1         Med          Q3         Max
## -3.82626472 -0.59671393  0.07178257  0.58569890  2.76317081
##
## Number of Observations: 492
## Number of Groups: 44
```

## 7.1 Notes on Data Processing

1. ID 26 had an initial dose listed as "30", with all subsequent doses listed as "0". Changed initial value to "0", as the "30" is most likely a typo (previous patient ID 25 has a dose of "30").

2. IDs 44 & ID 45 were excluded from the analysis due to having too few months on study.

3. After removing ID 45, there were 6 `NA` measures, simultaneous for `bcarot` and `vite`.
   All `NA`s were imputed with dosage group average at respective month

4. A number of patients had two measurements recorded for a given month. It is not clear why these repeated measures were taken, but the values are not obviously erroneous. In order to facilitate modeling the autocorrelation structure of the data set, and to make the data more consistent across subjects,

these repeated measures within a single month for a patient were replaced by the average of the two values.

There may be a better was of utilizing these extra measurements if more information regarding their occurrence can be obtained from the data collectors.

# References


Heinonen, Olli P, and Demetrius Albanes. 1994. "The Effect of Vitamin e and Beta Carotene on the Incidence of Lung Cancer and Other Cancers in Male Smokers." *The New England Journal of Medicine (USA).*

Huber, Peter J. 2005. *Robust Statistics*. Vol. 579. John Wiley & Sons.

Omenn, Gilbert S, Gary E Goodman, Mark D Thornquist, John Balmes, Mark R Cullen, Andrew Glass, James P Keogh, et al. 1996. "Effects of a Combination of Beta Carotene and Vitamin a on Lung Cancer and Cardiovascular Disease." *New England Journal of Medicine* 334 (18): 1150–55.

Peto, R, R Doll, J Del Buckley, and MB Sporn. 1981. "Can Dietary Beta-Carotene Materially Reduce Human Cancer Rates?" *Nature* 290 (5803): 201–8.

Ziegler, Regina G. 1989. "A Review of Epidemiologic Evidence That Carotenoids Reduce the Risk of Cancer." *The Journal of Nutrition* 119 (1): 116–22.