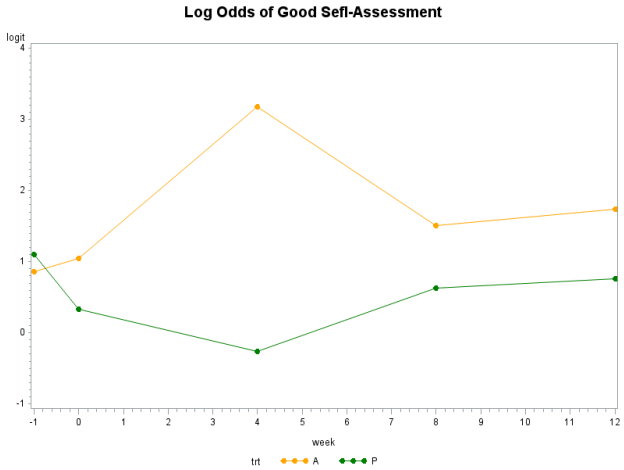
Question 1

a. Briefly summarize the outcome data from the study in a single table. Specifically, show for each treatment group at each scheduled assessment: how many provided an assessment and what number and percentage rated their arthritis status as "good". Based on the descriptive statistics in the table, briefly characterize the pattern of change in the proportions with a "good status" over time including a comparison of the two treatments.

Treatment		Week -1	Week 0	Week 4	Week 8	Week 12
A	N	27	27	25	22	20
	Proportion of "Good"	70%	74%	96%	82%	85%
P	N	24	24	23	23	22
	Proportion of "Good"	75%	58%	43%	65%	68%



For both groups, the proportions of "good" self-assessment over time have a non-linear trend, but with quite opposite patterns. For auranofin group, it goes up first and drops down later: during pre-randomization (week= -1, 0), it increases by a small proportion (4%); from randomization to week 4, it increases dramatically by 22%; after week 4, it drops about 14% and then stays relatively stable. For placebo group, it drops down first and goes up later: during pre-randomization and from randomization to week 4, it decreases dramatically (17% and 15%, respectively); after week 4, it goes up about 22% and stays quite stable. The difference between the two groups is most noticeable at week 4.

b. In the following questions, use PROC GENMOD in SAS to fit marginal logistic regression models using the approach of generalized estimating equations (GEE) to obtain parameter estimates. Define the time variable, T, to be a categorical variable with 4 levels: taking the value 99 for measurements at weeks -1 and 0 (i.e. for both "baseline" measurements prior to randomization) and values 4, 8 and 12 for the measurements at weeks 4, 8 and 12. The treatment variable, X, is an indicator variable taking the value 1 if a subject received auranofin (A) and the value 0 if placebo (P).

i. Consider a marginal logistic regression model for the repeated arthritis self-assessment measurements as the outcome variable and with just the T variable and a T*X interaction variable as categorical explanatory variables, using an independence working correlation matrix to describe the within-subject correlation structure (i.e. using the TYPE=IND on the REPEATED statement). Define appropriate notation and write down the algebraic form for the model that will be fitted using the GEE method, including any assumptions.

- The marginal expectation of the response $E(Y_{ij}|X_i, T_{ij})$ depends on covariates through a logit link function

$$logit\ E(Y_{ij}|X_i, T_{ij}) = \beta_1 + \beta_2 T_{ij} + \beta_3 X_i \times T_{ij}$$

(i.e., Logistic regression)

Where:

$Y_{ij} = 1$ if ith subject rated arthritis status as "good" on the jth week; $= 0$ otherwise

$T_{ij} = 99(reference\ level)$, if pre-randomization (week= -1 and 0); $= week_{ij}$ if after randomization; T_{ij} is categorical.

X_i=1 if ith subject received auranofin, X_i=0 if placebo.

- The marginal variance $var(Y_{ij}|X_i, T_{ij})$ depends on the marginal mean

$$var(Y_{ij}|X_i, T_{ij}) = E(Y_{ij}|X_i, T_{ij})\{1 - E(Y_{ij}|X_i, T_{ij})\}$$

(i.e., Bernoulli variance)

- Within-subject association is assumed to be independent

$$corr(Y_{ij}, Y_{ik}) = 0\ for\ j \neq k$$

(i.e., independent working correlation matrix)

ii. The model does not include the variable X as a main effect. What assumption is therefore being made? Briefly describe why this assumption might be reasonable in this study.

The main effect of X (treatment) is the effect of treatment at the reference level of the time variable, T. Since we code the two weeks prior to randomization as T=99, and treated T as categorical, now the reference level of T is T=99 (PROC GENMOD default uses the largest value as the reference level). Thus, the main effect of X is the effect of treatment at the "baseline" measurements prior to randomization (at both week= -1 and 0).

The assumption made here is that, prior to randomization, patients in auranofin and placebo groups have the same proportion rated their arthritis status as "good". Since in randomized studies, baseline values are designed to be at the same level, this assumption might be reasonable in this study.

iii. **Fit the model described in (i). Provide the PROC GENMOD code used and the estimates of the parameters in the mean part of the model obtained as well as empirical standard errors for these estimates. [HINT: Add "/ param=ref " after your list of variables in your class statement to ensure that SAS doesn't try to outsmart you and add main effects of X back into the model].**

```
proc genmod data=lart descending;
      class id T(ref='99') X(ref='0') week/param=ref;
      model y=T T*X/dist=Bin link=logit  ;
      repeated subject=id/withinsubject=week type=ind modelSE;
run;
```

Analysis Of GEE Parameter Estimates					
Empirical Standard Error Estimates					
Parameter			**Estimate**	**Standard Error**	**Pr > |Z|**
Intercept			0.8287	0.2708	0.0022
T	4		-1.0911	0.4306	0.0113
T	8		-0.2001	0.4831	0.6787
T	12		-0.0666	0.4693	0.8872
T*X	4	1	3.4404	1.1039	0.0018
T*X	8	1	0.8755	0.7051	0.2144
T*X	12	1	0.9725	0.7757	0.2100

iv. How is the model fit in (iii) different from an ordinary logistic regression analysis that assumes that all observations are independent? Note here the question is not "how do the results from the two analyses differ for these data", but how the models themselves differ – that is, why would they give different results. Which approach do you prefer and why?

The marginal model fit in (iii) is inherently different from ordinary logistic regression. Although the independent working correlation matrix assumption in (iii) and the independence assumption in ordinary logistic regression, seem to be similar.

In ordinary logistic regression, we assume the responses are independent and have a Bernoulli or Binomial distribution. Hence we can explicitly write down the likelihood function for the observed data at hand, with likelihood function, we then use maximum likelihood approach to estimate the parameters and make inferences.

However, in the marginal model, the repeated measures are not mutually independent, with discrete response longitudinal data, there is no simple analogue of the multivariate normal distribution. No simple distributional assumption can be made, so we use an alternative approach -- Generalized Estimating Equations (GEE), to estimate.

In GEE, we have to assume a working within-subject association or correlation structure, in this case, we just assume "independence" for the working correlation structure, although the responses are not actually independent. The reason we can make this assumption is that, the GEE estimator of β is a consistent estimator whether the within-subject associations have been correctly modeled. That is, for $\hat{\beta}$ to provide a valid estimate we only require the model for the mean response has been correctly specified.

Although GEE estimator $\hat{\beta}$ is consistent under misspecification of the within-subject associations, the standard errors obtained under a mis-specified correlation structure are not valid. However, GEE by default replaces $cov(Y_{ij})$ by empirical variance estimator when estimating the standard errors of $\hat{\beta}$, hence by correcting for misspecification of within-subject association or overdispersion, the empirical variance estimator provides a valid estimate of $\widehat{SE}(\hat{\beta})$.

In summary, for repeated measures, we prefer GEE approach rather than ordinary logistic regression.

- For estimation:
  In large sample, the two models both provide similar and valid estimate of β ; in small sample, only GEE provides most efficient or precise estimate under a "best" working association structure.

- For inference:
  Only GEE provides valid inference, since $\widehat{SE}(\hat{\beta})$ in ordinary logistic regression is not valid, which leads to incorrect p-value and potential misleading inferences.

v. For the model in (i), conduct a Wald test of the joint hypothesis that there is no effect of auranofin compared with placebo at all three of weeks 4, 8 and 12. [HINT: Use the TYPE3 and WALD options together on the MODEL statement]. What do you conclude concerning the effect of auranofin versus placebo? [Note: do not just provide a p-value from the test of the joint hypothesis; instead provide a broader interpretation of the results concerning any treatment differences].

```
proc genmod data=lart descending;
      class id T(ref='99') X(ref='0') week/param=ref;
      model y=T T*X/dist=Bin link=logit type3 wald ;
      repeated subject=id/withinsubject=week type=ind modelSE;
run;
```

Analysis Of GEE Parameter Estimates
Empirical Standard Error Estimates

Parameter			Estimate	Standard Error	Pr > |Z|
Intercept			0.8287	0.2708	0.0022
T	4		-1.0911	0.4306	0.0113
T	8		-0.2001	0.4831	0.6787
T	12		-0.0666	0.4693	0.8872
T*X	4	1	3.4404	1.1039	0.0018
T*X	8	1	0.8755	0.7051	0.2144
T*X	12	1	0.9725	0.7757	0.2100

Wald Statistics For Type 3 GEE Analysis

Source	DF	Chi-Square	Pr > ChiSq
T	3	11.14	0.0110
T*X	3	11.25	0.0105

The (multivariate) Wald test of treatment*week interaction (T*X) yields $\chi^2 = 11.25$, with 3 degrees of freedom, p-value=0.0105, so we reject the null hypothesis and conclude that there is an effect of auranofin compared with placebo at some weeks during 4, 8, and 12 (at least one of the three weeks).

The joint test doesn't tell us specifically how and at which week(s) they are significantly different, look at the single degree of freedom test from the parameter estimate table, we can see that at week 4, the odds of "good" self-assessment for auranofin patients is $e^{3.4404} = 31.2$ times the odds for placebo patients, which is a highly significant effect (p=0.0018); while at week 8 and 12, the odds of "good" self-assessment for auranofin patients are $e^{0.8755} = 2.4$ *and* $e^{0.9725} = 2.6$ times the odds for placebo patients, respectively, which are not significant (p=0.2144, p=0.2100). So the auranofin is most effective at week 4, while at week 8 and 12, there are some effects but not statistically significant (if clinically effective).

vi. Now using PROC GENMOD, fit the same model but using an exchangeable log odds ratio structure for the working association structure. Write out how your model formulation given in (i) changes under this alternative assumption. Describe whether the conclusions of your GEE analyses in (v) are sensitive to the change in association structure.

The only model formulation in (i) change is the third part:

- Within-subject association is accounted for by assuming a common pair-wise log odds ratio

$$\log odd\ ratio\left(Y_{ij}, Y_{ik}\right) = \alpha\ for\ j \neq k$$

(i.e., exchangeable log odds ratio structure)

```
proc genmod data=lart descending;
      class id T(ref='99') X(ref='0') week/param=ref;
      model y=T T*X/dist=Bin link=logit type3 wald ;
      repeated subject=id/withinsubject=week logor=exch  modelSE;
run;
```

Analysis Of GEE Parameter Estimates					
Empirical Standard Error Estimates					
Parameter			**Estimate**	**Standard Error**	**Pr > |Z|**
Intercept			0.8287	0.2708	0.0022
T	4		-0.9827	0.4129	0.0173
T	8		-0.1086	0.5091	0.8312
T	12		0.0137	0.4696	0.9768
T*X	4	1	3.1819	1.0172	0.0018
T*X	8	1	0.5536	0.6562	0.3989
T*X	12	1	0.6774	0.6155	0.2711
Alpha1			1.7478	0.4253	<.0001

Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
T	3	9.99	0.0187
T*X	3	10.07	0.0180

The conclusions of the GEE analyses in (v) are not sensitive to the change in association structure.

The results from the exchangeable log odds ratio structure is qualitatively the same as the results from the independent structure, but quantitatively changed a little bit. The conclusions in (v) are still preserved here. The GEE empirical standard error estimates provides consistent estimates and valid standard errors under mis-specified within-subject association structures, so both models have valid and also similar estimates and inferences (p-values).

c. The investigator is interested in knowing whether the results comparing the effect of auranofin versus placebo on arthritis status over time is affected by any imbalance in sex between the treatment groups. Conduct analysis to address this specific issue. Briefly describe how you approached this (include your PROC GENMOD code) and what your key findings are.

Let's assume that female and male patients have different propensity for rating their arthritis status as "good"(this propensity is correspond to the main effect of gender), and this propensity doesn't change over time (no interaction between gender*time). The auranofin treatment has the same effect for both female and male (no interaction between gender*treatment).

So add the main effect of Gender into the model for the marginal mean

$$logit\ E(Y_{ij}|X_i, T_{ij}) = \beta_1 + \beta_2 T_{ij} + \beta_3 X_i \times T_{ij} + \beta_4 Gender_i$$

Where $Gender_i = 1\ if\ ith\ subject\ is\ female, 0\ if\ male$

In this way, we're controlling for gender difference in propensity of rating "good", that is, we always contrast patients in the two groups but with the same gender, at each week. So any treatment effect obtained by comparing patients with the same gender but different treatments shouldn't be confounded by any imbalance in gender.

```
proc genmod data=lart descending;
       class id T(ref='99') X(ref='0') gender(ref='0') week/param=ref;
       model y= gender T T*X/dist=Bin link=logit type3 wald ;
       repeated subject=id/withinsubject=week logor=exch  modelSE;
run;
```

Analysis Of GEE Parameter Estimates					
Empirical Standard Error Estimates					
Parameter			**Estimate**	**Standard Error**	**Pr > |Z|**
Intercept			0.9667	0.3145	0.0021
gender	1		-0.4872	0.4853	0.3154
T	4		-0.9982	0.4183	0.0170
T	8		-0.1208	0.5108	0.8131
T	12		-0.0171	0.4739	0.9713
T*X	4	1	3.2383	1.0526	0.0021
T*X	8	1	0.5675	0.6680	0.3955
T*X	12	1	0.7219	0.6062	0.2337
Alpha1			1.7190	0.4253	<.0001

Wald Statistics For Type 3 GEE Analysis			
Source	**DF**	**Chi-Square**	**Pr > ChiSq**
gender	1	1.01	0.3154
T	3	9.61	0.0222
T*X	3	9.85	0.0199

The estimate of gender is -0.4872, indicating that the odds of rating arthritis status as "good" in female patients is only $e^{-0.4872} = 0.6143$ the odds in male patients, at the same measurement time and with same treatment (either auranofin or placebo), but is not statistically significant (p=0.3154).

After controlling for sex, the sex-adjusted treatment effects still hold as the previous finding. Specifically, The Wald test of T*X yields $\chi^2 = 9.85$, with 3 degrees of freedom, p-value=0.0199, so we conclude that there is a significant effect of auranofin compared with placebo at some weeks during 4, 8, and 12 (at least one of the three weeks), after adjusting for sex. At week 4, the odds of "good" self-assessment for auranofin patients is $e^{3.2383} = 25.5$ times the odds for placebo patients with the same sex, which is highly significant effect (p=0.0021); while at week 8 and 12, the odds of "good" self-assessment for auranofin patients are $e^{0.5675} = 1.8\ and\ e^{0.7219} = 2.1$ times the odds for placebo patients with the same sex, respectively, which are not significant (p=0.3955, p=0.2337). So the auranofin is still most effective at week 4, while at week 8 and 12, there are some effects but not statistically significant, after adjusting for sex.

The results comparing the effect of auranofin versus placebo on arthritis status over time is affected by an imbalance in sex between the two treatment groups.

Question 2:

a) Provide an algebraic definition for a generalized linear marginal model in which the only effects are for the intercept and Year (as a continuous variable). Fit this model and provide a table from your SAS output which includes the estimates of the parameters in your model.

- The marginal expectation of the response $E(Y_{ij}|Year_{ij})$ depends on covariates through a log link function

$$logE(Y_{ij}|Year_{ij}) = \beta_1 + \beta_2 Year_{ij}$$

(i.e., Poisson regression)

Where:

Y_{ij} is the counts of new skin cancers for the ith subject in jth follow-up year

$Year_{ij}$=1,2,3,4, 5 is the year of follow-up

- The marginal variance $var(Y_{ij}|Year_{ij})$ depends on the marginal mean

$$var(Y_{ij}|Year_{ij}) = \emptyset E(Y_{ij}|Year_{ij})$$

(i.e., a scale parameter for overdispersion)

- Within-subject association is accounted for by assuming a common pair-wise correlation

$$corr(Y_{ij}, Y_{ik}) = \alpha$$

(i.e., exchangeable or compound symmetry correlation pattern)

```
proc genmod data=skin;
      class id cyear;
      model y=year /dist=poisson link=log scale=pearson;
      repeated subject=id/withinsubject=cyear type=exch;
run;
```

Analysis Of GEE Parameter Estimates						
Empirical Standard Error Estimates						
Parameter	**Estimate**	**Standard Error**	**95% Confidence Limits**		**Z**	**Pr > |Z|**
Intercept	-1.5410	0.1674	-1.8692	-1.2129	-9.20	<.0001
Year	-0.1065	0.0486	-0.2018	-0.0113	-2.19	0.0284

b) Provide an algebraic definition for a generalized linear mixed model (GLMM) in which the only fixed effects are for the intercept and Year (as a continuous variable), and the only random effect is the intercept. Assume there is no overdispersion. What is being assumed about how the distribution of risk among subjects changes with time?

Poisson Mixed Effects Model:

Step 1:

Conditional on random effects (i.e., intercepts) b_i, Y_{ij} are independent and have a Poisson distribution, with $Var(Y_{ij}|b_i) = E(Y_{ij}|b_i),\ (i.e.\ \phi = 1\ no\ overdispersion)$

The conditional mean $E(Y_{ij}|b_i)$ is related to the linear predictor by a log link function:

$$\log E(Y_{ij}|b_i) = \beta_1 + \beta_2 Year_{ij} + b_i$$

Step 2:

The random effects b_i is assumed to have a normal distribution, with zero mean and variance σ_b^2

$$b_i \sim N(0, \sigma_b^2)$$

The distribution of risk among subjects in center 1 is assumed to change linearly over time, and with the same fixed rate of change for all subjects. But subjects have different starting risk at baseline (i.e. Year=0), because of subject-specific random intercepts.

c) Fit your chosen GLMM from (b) in SAS and provide the SAS code for the fitting the model (just provide the code from the PROC that you use). Provide a table from your SAS output which includes the estimates for the parameters in your GLMM, and provide careful interpretation of the Year term.

```
proc glimmix data=skin method=QUAD(Qpoints=50);
       class id;
       model y=year/dist=poisson link=log s;
       random intercept/subject=id;
run;
```

Solutions for Fixed Effects

Effect	Estimate	Standard Error	DF	t Value	Pr > |t|
Intercept	-2.2503	0.1716	421	-13.12	<.0001
Year	-0.1082	0.04340	1460	-2.49	0.0128

Covariance Parameter Estimates

Cov Parm	Subject	Estimate	Standard Error
Intercept	**id**	1.4678	0.2665

$\hat{\beta}_2 = -0.1082$:
For any patient enrolled in center 1, the patient-specific count of new skin cancer decreases multiplicatively by $e^{-0.1082} = 0.8974$ each year, or, the patient-specific count ratio of new skin cancer comparing 1-year after to the previous year is $e^{-0.1082} = 0.8974$.

d) Are the estimates for the fixed intercept terms the same or different in the GLMM compared with the marginal model fitted in question 2a? Why are they the same or different?

Solutions for Fixed Effects				
Model	Effect	**Estimate**	**Standard Error**	**Pr > |t|**
GLMM	Intercept	-2.2503	0.1716	<.0001
Marginal	Intercept	-1.5410	0.1674	<.0001

The fixed intercepts terms in the Marginal model and GLMM are different.

In the marginal model, the fixed intercept represents the log-transformed population-average expected counts of new skin cancers at baseline for all patients in center 1. In GLMM, the fixed intercept represents the log-transformed subject-specific expected counts of new skin cancers at baseline for a "typical" patient in center 1 with a random intercept $b_i = 0$. Quantitatively, in this example, the estimate from marginal model is about 30% smaller than the estimate from GLMM, but the standard errors have the same pattern. Thus, qualitatively, they yield similar p-values.

e) Use the parameter estimates from your GLMM and your model definition to characterize the distribution of expected counts of new skin cancers among subjects at center 1 during their first year of follow-up.

Solutions for Fixed Effects					
Effect	**Estimate**	**Standard Error**	**DF**	**t Value**	**Pr > |t|**
Intercept	-2.2503	0.1716	421	-13.12	<.0001
Year	-0.1082	0.04340	1460	-2.49	0.0128

Covariance Parameter Estimates			
Cov Parm	**Subject**	**Estimate**	**Standard Error**
Intercept	**id**	1.4678	0.2665

The expected counts of new skin cancers among subjects at center 1 during their first year of follow-up having the following distribution according to model definition and parameter estimates from GLMM:

$$\log E(Y_{i1}|b_i) = \widehat{\beta_1} + \widehat{\beta_2} * 1 + b_i$$
$$b_i \sim N(0, \widehat{\sigma_b^2})$$
$$\Rightarrow \log E(Y_{i1}|b_i) \sim N\left(\widehat{\beta_1} + \widehat{\beta_2} * 1, \widehat{\sigma_b^2}\right) = N(-2.3585, 1.4678)$$

So, at center 1, during the first year of follow-up, the distribution of log-transformed expected counts of new skin cancers among subjects at center 1 follows a (univariate) Normal distribution with mean -2.3585 and standard deviation $\sqrt{1.4678} = 1.21$.

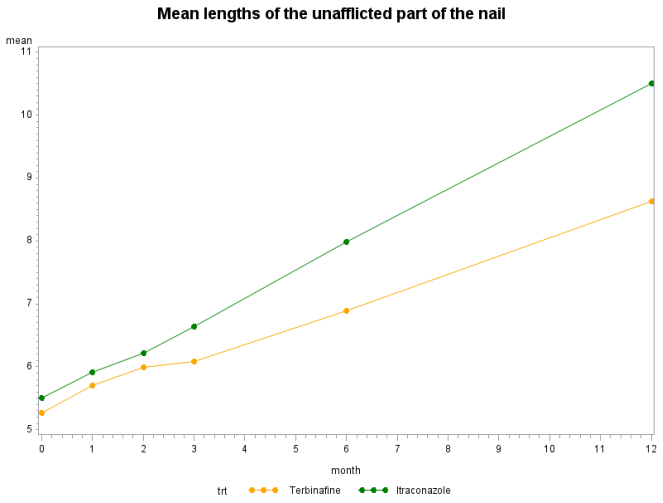
In un-transformed scale, a "typical" patient with random intercept $b_i = 0$ has the expected counts $e^{-2.3585} = 0.095$; and 95% patients would have the expected counts between $(0.0088,\ 1.0162) = (e^{-2.3585-1.96*\sqrt{1.4678}}, e^{-2.3585+1.96*\sqrt{1.4678}})$.

Optional Extra Credit Problem*

i) Is there a difference in the pattern of change of lengths of the unafflicted part of the nail between subjects receiving terbinafine and itraconazole over a 12 month period? Does one treatment show results more quickly?

Part I: Descriptive Statistics

treatment		Month 0	Month 1	Month 2	Month 3	Month 6	Month 12
Terbinafine	N	100	100	100	100	100	100
	mean	5.27	5.696	5.989	6.084	6.894	8.631
Itraconazole	N	100	100	100	100	100	100
	mean	5.501	5.907	6.215	6.637	7.984	10.508



For both treatment groups, the average length of the unafflicted part of the nail changes linearly over time. But subjects receiving terbinafine seem have a sharper increase trajectory over 12 month period.

Part II: Model

Choose a linear mixed effects model with random intercepts and slopes. This model serves 2 purposes:

- It accounts for correlation among repeated observation on the same subject via a random effect structure, in a way that that the correlation among repeated measures are not constant but changes as a function of time
- It allows for both population-average interpretation and subject-specific interpretation. Since the linear mixed effects model assume a linear or identity link function, the fixed effects have both population-average and subject-specific interpretations. But in this particular question, we're interested in a between-subject treatment effect, the population-average interpretation is more straightforward.

Due to randomization, the baseline measure should be the same for the treatment groups, we omit the main effect of treatment in the model. The length for ith subject on jth month:

$$Y_{ij} = \beta_1 + \beta_2 Month_{ij} + \beta_3 Month_{ij} \times Treatment_i + b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij}$$

- The Fixed Part:

$$\beta_1 + \beta_2 Month_{ij} + \beta_3 Month_{ij} \times Treatment_i$$

- The Random Part:

$$b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij}$$

- Where $Month_{ij}$ is a continuous variable indicating time since baseline

- b_{1i} is the subject-specific random intercept, b_{2i} is the subject-specific random slope. Assuming $b_{1i} \sim N(0, \sigma_{b_1}^2)$, $b_{2i} \sim N(0, \sigma_{b_2}^2)$, and $cov(b_{1i}, b_{2i}) = \sigma_{b_{1,b_2}}$. $\sigma_{b_1}^2$ is between-subject intercepts variation, $\sigma_{b_2}^2$ is between-subject slopes variation, $\sigma_{b_{1,b_2}}$ is between-subject co-variation of intercepts and slopes.

  Or we can re-write as, $(b_{1i}, b_{2i}) \sim MVN(0, G)$ where $G = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_{1,b_2}} \\ \sigma_{b_{1,b_2}} & \sigma_{b_2}^2 \end{bmatrix}$

- And assuming within-subject erroe $\varepsilon_{ij} \sim N(0, \sigma^2)$. σ^2 is the within-subject error variation.

Part III: Results and Interpretation

```
proc mixed data=toenail;
       class trt id;
       model length= month trt*month/s cl;
       random intercept month/type=un subject=id;
run;
```

Solution for Fixed Effects				
Effect	**trt**	**Estimate**	**Standard Error**	**Pr > |t|**
Intercept		5.3917	0.1121	<.0001
month		0.4254	0.03312	<.0001
month*trt	**1**	-0.1584	0.04460	0.0004
month*trt	**2**	0	.	.

Covariance Parameter Estimates		
Cov Parm	**Subject**	**Estimate**
UN(1,1)	**id**	2.2020
UN(2,1)	**id**	-0.1883
UN(2,2)	**id**	0.1103
Residual		0.9412

- $\hat{\beta}_3 = -0.1584$ $(p = .0004)$

The estimate for the difference in mean (population-average) slope between two treatments is -0.1584 (p=.0004). We can conclude that, on average, the growth rate in terbinafine group is 0.1584 mm/month slower than that in itraconazole group. And there is a 95% chance that the true difference of growth rate per month would be between -0.2460 and -0.07087.

- $\widehat{Var}(b_{2i}) = 0.1103$

There is substantial variability between subjects in the growth rate. For terbinafine group, we would expect 95% patients have growth rates between (-0.38, 0.92); for itraconazole group, we would expect 95% patients have growth rates between (-0.23, 1.08).

Part VI: Sensitive Analysis

Now add the main effect of treatment to the above model:

$$Y_{ij} = \beta_1 + \beta_2 Month_{ij} + \beta_3 Treatment_i + \beta_4 Month_{ij} \times Treatment_i + b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij}$$

```
proc mixed data=toenail;
       class trt id;
       model length= month trt trt*month/s cl;
       random intercept month/type=un subject=id;
run;
```

Solution for Fixed Effects				
Effect	**trt**	**Estimate**	**Standard Error**	**Pr > |t|**
Intercept		5.4421	0.1588	<.0001
month		0.4208	0.03464	<.0001
trt	**1**	-0.1007	0.2246	0.6541
trt	**2**	0	.	.
month*trt	**1**	-0.1493	0.04899	0.0024
month*trt	**2**	0	.	.

The main effect of treatment is not significant (p=0.6541), which confirms the validity of the model without the main effect of treatment. And this model yields similar estimate and p-value for $Month_{ij} \times Treatment_i$ (both quantitatively and qualitatively).

Part V: Conclusion:

To answer the main question of interest, there is a significant difference in the pattern of change of lengths of the unafflicted part of the nail between subjects receiving terbinafine and itraconazole over a 12 month period. On average (population-average), the growth rate in terbinafine group is 0.1584 mm/month slower than that in itraconazole group. So itraconazole shows results more quickly.

ii) Is there an association between the pattern of change of nail lengths and gender and/or health club frequency in subjects taking terbinafine? This might indicate that this drug brings about relief more swiftly in some kinds of subject versus others.

We use the following linear mixed effects model for subjects taking terbinafine only:

$$Y_{ij} = \beta_1 + \beta_2 Month_{ij} + \beta_3 Gender_i + \beta_4 Gender_i \times Month_{ij} + \beta_5 Club_i + \beta_6 Club_i \times Month_{ij} + b_{1i} + b_{2i} t_{ij} + \varepsilon_{ij}$$

Where:
$Gender_i$ is the gender indicator for ith subject taking terbinafine (=0 if female, =1 if male)
$Club_i$ is the Health club frequency indicator for ith subject taking terbinafine (=0 if once a week or less, =1 if more than once a week)

Solution for Fixed Effects

Effect	gender	club	Estimate	Standard Error	Pr > |t|
Intercept			5.1682	0.2209	<.0001
month			0.2530	0.03587	<.0001
gender	**0**		-0.5118	0.2978	0.0865
gender	**1**		0	.	.
month*gender	**0**		0.04583	0.04834	0.3437
month*gender	**1**		0	.	.
club		**0**	1.2847	0.3166	<.0001
club		**1**	0	.	.
month*club		**0**	-0.01204	0.05139	0.8148
month*club		**1**	0	.	.

$\widehat{\beta_4}$=0.04583 (p=0.3437):

The estimate for the difference in mean slope between male and female subjects taking the same treatment terbinafine is 0.04583 (p=.3437). We can conclude that, for subjects taking terbinafine and with the same health club frequency status,the expected growth rate for female taking terbinafine is 0.04583 mm/month faster than that in male taking terbinafine, but is not statistically significant.

$\widehat{\beta_6}$=-0.01204 (p=0.8148):

The estimate for the difference in mean slope between subjects taking the same treatment terbinafine but differ in terms of health club frequency is -0.01204 (p=.8148). We can conclude that, for subjects taking terbinafine and with the same gender, the expected growth rate for subjects visiting health club more than once a week is 0.01204 mm/month faster than that in those visiting once a week or less, but is not statistically significant.

In conclusion, terbinafine brings about relief more swiftly in female than male with the same health club frequency status, and more swiftly in subjects visiting health club more than once a week than in those visiting once a week or less with the same gender. Although they are not statistically significant, they could be clinically significant.

```
*******************************
Appendix: SAS codes
*******************************;

*****************************************Q1;
*import dataset;
data art;
        infile  'C:\data\Projects\APCD High Cost\Longitudinal\arthritis.txt';
        input id sex $ age trt $ y_1 y0 y4 y8 y12;
proc print;run;

*Transpose to long format;
data lart;
        set art;
        y=y_1;week=-1;output;
        y=y0;week=0;output;
        y=y4;week=4;output;
        y=y8;week=8;output;
        y=y12;week=12;output;
        drop y_1 y0 y4 y8 y12;
proc print;run;

*Q1 a ;
proc sort data=lart;by trt week;run;
proc means data=lart noprint;
        by trt;class week;
        var y;
        output out=meandata(drop=_type_ _freq_) mean=mean N=N ;
run;
proc transpose data=meandata out=outmean;
        by trt;
        id week;
var N mean ;
proc print data=outmean; run;

* Plot the untransformed proportion by treatment;
proc gplot data=meandata ;
        symbol1 color=orange interpol=join value=dot;
        symbol2 color=green interpol=join value=dot;
        plot mean*week=trt;
        title 'Proportion of Good Sefl-Assessment';
run;

* After logit transform;
data meandata;
        set meandata;
        logit=log(mean/(1-mean));
run;
proc gplot data=meandata;
        symbol1 color=orange interpol=join value=dot;
        symbol2 color=green interpol=join value=dot;
        plot logit*week=trt;
        title 'Log Odds of Good Sefl-Assessment';
run;

*Q1 b Marginal Model;
data lart;
        set lart;
        if week in (-1,0) then T=99;else T=week;
        if trt='A' then X=1;else X=0;
run;

proc genmod data=lart descending;
        class id T(ref='99') X(ref='0') week/param=ref;
        model y=T T*X/dist=Bin link=logit  ;
        repeated subject=id/withinsubject=week type=ind modelSE;
run;

* Wald test;
proc genmod data=lart descending;
        class id T(ref='99') X(ref='0') week/param=ref;
```

```
        model y=T T*X/dist=Bin link=logit type3 wald ;
        repeated subject=id/withinsubject=week type=ind modelSE;
run;

* exchangeable log odds ratio;
proc genmod data=lart descending;
        class id T(ref='99') X(ref='0') week/param=ref;
        model y=T T*X/dist=Bin link=logit type3 wald ;
        repeated subject=id/withinsubject=week logor=exch  modelSE;
run;

*imbalance in sex between the treatment groups;
data lart;
        set lart;
        if sex='F' then gender=1;else gender=0;
run;
proc genmod data=lart descending;
        class id T(ref='99') X(ref='0') gender(ref='0') week/param=ref;
        model y= gender T T*X/dist=Bin link=logit type3 wald ;
        repeated subject=id/withinsubject=week logor=exch  modelSE;
run;


****************************************Q2 ;
*import dataset;
data data;
        infile  'C:\data\Projects\APCD High Cost\Longitudinal\skin.txt';
        input id center Age Skin Gender Exposure Y Treatment Year;
run;

proc format ;
        value treatment_
        0='Placebo'
        1='beta-carotene'
;
run;

data skin;
        set data;
        cyear=year;format treatment treatment_.;
        where center=1;
proc print;run;

proc genmod data=skin;
        class id cyear;
        model y=year /dist=poisson link=log scale=pearson;
        repeated subject=id/withinsubject=cyear type=exch modelSE;
run;

proc glimmix data=skin method=QUAD(Qpoints=50);
        class id;
        model y=year/dist=poisson link=log s;
        random intercept/subject=id;
run;


***********************************Q3;
data toenail;
        infile  'C:\data\Projects\APCD High Cost\Longitudinal\toenail.dat';
        input id club gender month length trt;
run;

proc format;
        value club
        0='Once a week or less'
        1='More than once a week'
;
run;
proc format;
        value gender_
        0='Female'
```

```
        1='Male'
;
run;
proc format;
        value trt
        1='Terbinafine'
        2='Itraconazole'
;
run;

*Despripctive Summary;
proc sort data=toenail;by trt month;run;
proc means data=toenail n mean nway;
        format trt trt_.;
        by trt;class month;
        var length;
        output out=meandata(drop=_type_ _freq_) mean=mean  N=N;
run;
proc transpose data=meandata out=outmean;
        by trt;
        id month;
        var N mean ;
proc print data=outmean;run;

* Plot means by treatment gorup;
proc gplot data=meandata ;
        symbol1 color=orange interpol=join value=dot;
        symbol2 color=green interpol=join value=dot;
        plot mean*month=trt;
        title 'Mean lengths of the unafflicted part of the nail';
run;

*Linear Mixed Effects Model;
proc mixed data=toenail;
class trt id;
model length= month trt*month/s cl;
random intercept month/type=un subject=id;
run;
proc mixed data=toenail;
class trt id;
model length= month trt trt*month/s cl;
random intercept month/type=un subject=id;
run;

* Risk factors;
data terb;
set toenail;
where trt=1;
run;
proc mixed data=terb;
class trt id gender club;
model length= month gender gender*month club club*month/s;
random intercept month/type=un subject=id;
run;
```