**BIOS6643 HW4 2016**

1. Consider a study in which subjects’ blood pressures are observed over time (4 time points, equally spaced, no missing data). The model will have fixed effects for time, age (at start of experiment) and gender; a random intercept for subjects will also be included.
   1. Write out the mixed model, specifying all parts, if time is modeled as a continuous variable (linear term only) and the AR(1) structure is used to model the errors.

, where index age, subject, time and gender, respectively and b is associated with the random effect (intercept) for subjects; i=1,… ,n; j=1,…,4 (continuous); h =1, …,m; k = 1(male) and 2(female).

, (6611 way of doing it), where are fixed effects associated with time, age and gender, respectively and b is associated with the random effect (intercept) for subjects; i=1,… ,n; j=1,…,4 (continuous).

* 1. How many fixed effect (β) parameters are in the model (considering full-rank parameterization)? How many covariance (α) parameters?

Fixed effects:

1 for time, 1 for age, 1 for gender, and 1 for the intercept thus 4 fixed effect parameters considering full-rank parameterization.

Covariance parameters:

There will be 1 random intercept (for subject) and since the model is AR(1) structure, there will be 1 for AR(1) and 1 for thus in total there will be 3 covariance parameters.

* 1. Write the form of Var(**Y***i*) in terms of the covariance parameters.
  2. Repeat parts a-c if time is modeled as a class variable, the UN structure is used for the covariance matrix of the errors and there is no random intercept term.

where are fixed effects associated with time, age, and gender, respectively; i=1,…,n; j=1,…,4; k=range of age at start of experiment to end of experiment; l=1 or 2; let for time 2, 0 otherwise; for time 3, 0 otherwise for time 4, 0 otherwise (time 1 as control).

Fixed effects:

3 parameter for time (4-1=3 considering full rank), 1 for age, 1 for intercept, and 1 for gender, thus 6 fixed effect parameters considering full-rank parameterization.

Covariance parameters:

There will be 0 random intercept (for subject) and since the model is UN structure, there will be 10 for UN.

* 1. **Not to turn in.** In the model, age at start of experiment (i.e., baseline) was used. How would estimates change if you used continuous age in the model instead? In order to answer the question, write out the statistical models for both approaches.

If age was considered to be continuous then the model would incorporate the parameter for time and age into one fixed parameter.

where are fixed effects associated with time, age, and gender, respectively; i=1,…,n; j=1,…,4; k=range of age at start of experiment to end of experiment; l=1 or 2.

Where now the age of individuals change with time but since they are a fixed effect, the variance due to the change in age will be the same as the variance due to time thus the sums of squares of error will not change but the parameter will change to reflect the change in age due to time.

1. Consider the linear mixed model (in subject form): **Y***i*=**X***i***β**+**Z***i***b***i*+**ε***i*, where **ε**i~*N*(**0**,**R***i*) independent of **b***i*~N(**0**,**G***i*), and subjects *i*=1,…,*n* are independent. [Here, **G***i* and **R***i* are the same for all subjects; the index is just used to indicate dimensions.] A longitudinal experiment is conducted where subjects are observed at 3 equally spaced time points – for simplicity you can let the time points be 0,1, and 2. The data will be modeled using **b***i*=(*bi*0,*bi*1)*t* defined for subjects (*bi*0 = random intercept; *bi*1 = random slope for time) and **R**=*σε*2**I**. There will be fixed-effect terms for time and possibly other variables, but their specification isn’t relevant to the questions below. Justify all responses / show work.
2. Determine **V**i=*Var*(**Y**i) (the model covariance matrix for subject *i*) if the unstructured covariance structure is used for **G**.
3. Write the SAS, PROC MIXED code for the analysis of the data, if we allow for a simple linear time trend for fixed effects. (You can use generic names, e.g., *dat* for the data set, *time* for time.)

PROC MIXED DATA=dat;

MODEL y = time ;

RANDOM intercept time/subject=id type=un s;

RUN;

1. Often a realistic covariance structure for **V**i for longitudinal data is one where covariance between responses is positive but decays as time between responses increases. With the covariance structure you came up with in part a, is it possible for this structure to have covariance that decays as time between responses increases? Justify your response.

Yes it is possible to have covariance structure to decay over time. For example, the **R** matrix with AR(1) covariance structure will have a decreasing covariance as time increases (assuming between responses increases. For the covariance structure depicted in part a, the covariance between random intercept and random slope (i.e.) would need to be a negative value thus as time increases between responses the covariance between responses will decrease.

.

1. Consider the peak flow data fit with a simple random effect model, with SAS output presented on pages 12-14 of the *LMM inference* notes (pages 122-124 of the complete PDF).
   1. Interpret the tests for random effects, e.g., interpret the effect for subject 101S.

Based on the partial output in pg 122-124 of appended LMM: inference notes, the random intercept for subject 101S is significantly different compared to the overall expected random intercept (0) with p-value 0.0412. Subject 102S has a larger estimated random intercept (108.63) away from the null, thus a higher p-value (<0.0001), while subject 110S has a smaller difference away from null (-28.8484) thus a non-significant p-value (0.2562). Subjects 103S, 104S, and 105S all have random intercepts that are significantly different from the null (.

* 1. Based on the estimated covariance parameters, are you surprised that the majority of kids have p-values below 0.05? Explain.

No, I’m not surprised because the is so much larger than the which is indicative that the between subject variability is much larger than the within subject variability and thus the resulting p-value for the random intercept should be relatively significant for majority of the kids. In terms of ICC (, this coefficient will be high, again indicating that the variability between subjects are higher compared to the variability within a subject.

* 1. **Not to turn in.** For subject 101S, determine the following, replacing unknown parameters with their estimates. The data is posted on the course web site; you only need the January values for subject 101S. The quantities can be found in the slides and notes as mentioned above.
     1. The shrinkage factor,.
     2. The random effect estimate, .
     3. The predicted PEF, . Is more weight given to the subject’s data or the population average (estimate)? Does the result surprise you? Explain.

1. (From Rosner, 2006.) A clinical trial was planned comparing the incidence of cancer in a group taking beta-carotene in capsule form compared with a group taking beta-carotene placebo capsules. One issue in planning such a study is which preparation to use for the beta-carotene capsules. Four preparations were considered: (1) Solatene (30mg capsules), (2) Roche (60mg capsules), (3) BASF (30mg capsules), (4) BASF (60mg capsules). To test efficacy of the four agents in raising plasma-carotene levels, a small bioavailability study was conducted. After two consecutive-day fasting blood samples, 23 volunteers were randomized to one of the four preparations mentioned above, taking 1 pill every other day for 12 weeks. The primary endpoint was level of plasma carotene attained after moderately prolonged steady ingestion. For this purpose, blood samples were drawn at 6, 8, 10 and 12 weeks. In order to model the data, consider group and time as class variables. Other model specifications may depend on the question. **Just use the second baseline measure as the ‘time 0’ measure for parts a through f.**
   1. Create a graph for the data. You have some flexibility on what type of graph to include.

Figure 1. Mean plasma beta-carotene level by time per group.

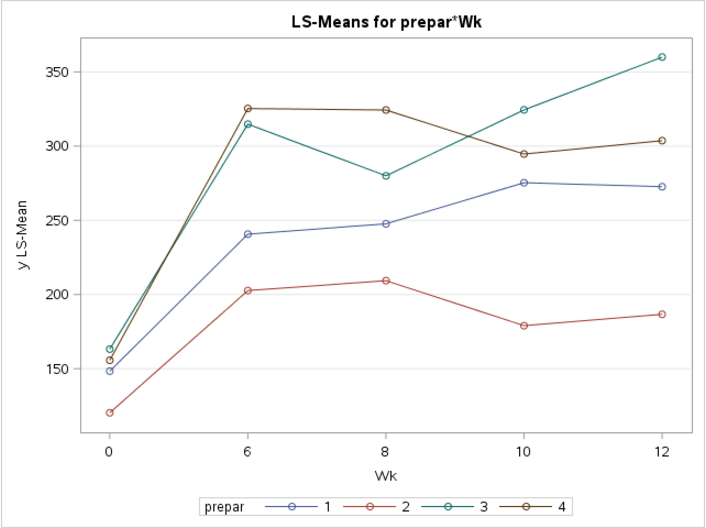
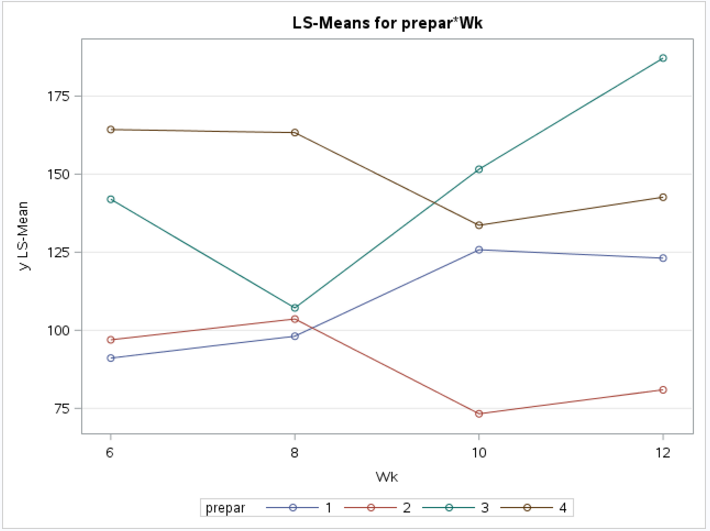


Figure 2. Mean difference in plasma beta-carotene level (from 6,8,10,12 weeks to baseline) per time point per group using the Hybrid Model.



* 1. Consider the model with a random intercept for subjects (in addition to other terms mentioned above); let **R** have the independent structure (**R**=*σε*2**I**). Mention 2 approaches to fitting the model and conducting tests.

Approach 1: Use RM ANOVA method.

We are able to use the RM ANOVA method because assumption of sphericity (at least induced as independent **R** structure) will be assumed.

proc glm order=data data=bclong;

class wk prepar;

model y=wk0 prepar wk prepar\*wk/ solution;

random boy / test; run;

Approach 2: Use Linear Mixed Model method.

proc mixed data=bclong;

class id Prepar wk;

model y= wk0 Prepar wk prepar\*wk / solution;

random intercept / subject=id(Prepar);

contrast 'linear' wk -2 -1 0 1 2;

contrast 'quadratic' wk 2 -1 -2 -1 2;

contrast 'cubic' wk -1 2 0 -2 1;

contrast 'quartic' wk 1 -4 6 -4 1 ;

contrast 'lxl' prepar\*wk -2 -1 0 1 2 2 1 0 -1 -2 0 0 0 0 0,

prepar\*wk -2 -1 0 1 2 0 0 0 0 0 2 1 0 -1 -2;

contrast 'qxq' prepar\*wk 2 -1 -2 -1 2 -2 1 2 1 -2 0 0 0 0 0,

prepar\*wk 2 -1 -2 -1 2 0 0 0 0 0 -2 1 2 1 -2;

contrast 'cxc' prepar\*wk -1 2 0 -2 1 1 -2 0 2 -1 0 0 0 0 0,

prepar\*wk -1 2 0 -2 1 0 0 0 0 0 1 -2 0 2 -1;

contrast '4x4' prepar\*wk 1 -4 6 -4 1 -1 4 -6 4 -1 0 0 0 0 0,

prepar\*wk 1 -4 6 -4 1 0 0 0 0 0 -1 4 -6 4 -1;

lsmeans prepar\*wk; run;

Approach 3 and 4: We can model the time and continuous or class.

Approach 5 and 6: We could add polynomial trends, seeing that there is potentially a cubic relationship.

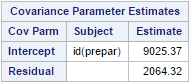
Approach 7 and 8: Estimation of the fixed effects using ML or REML.

Approach 9: We could also use the robust/sandwich variance to estimate the for estimation and hypothesis testing.

\*4 different approach to set up baseline values: Baseline as outcome, baseline as outcome but assume same intercept, baseline as covariate, and change score model. Here we will use the baseline as outcome with no assumptions.

* 1. For the model above, determine the ICC and interpret the quantity.

Table 1. Covariance Parameter Estimate



The ICC (, (=81.4%. This indicates that the variability in the data that is due to between subject differences is 81.4%.

* 1. Now consider the same model but include a random slope for continuous (linear) time for subjects in addition to the random intercept. (NOTE: since we’re modeling time as a class variable in the fixed-effect part of the model, you’ll need to define a new time variable so that you can treat it as a continuous variable for the random effect.) Try the UN structure for **G** and compare results using the VC structure. Which model is better? Interpret the covariance term between the intercept and slope.

Both models converged…. But the UN **G** model did not have a positive definite **G** matrix. The UN structure for **G** had an AIC (1068.7) smaller compared to the VC structure AIC (1077.8) indicating that the covariance structure using UN for **G** allows for a better fit although the note in the SAS output indicates that the G matrix is not positive definite, the final V matrix indicates an increasing variance between responses thus the UN G model can be used for inferences.

For the UN structure model, the covariance term between the random intercept and random slope is UN(2,1) which is 344.51. This indicates that for subjects with greater random intercept values, they have a greater slope. Or this can be roughly interpreted as subjects with greater plasma betacarotene levels on average have a greater increase in plasma betacarotene level over time.

Table 2. Random intercept random slope with UN structure **G** matrix.

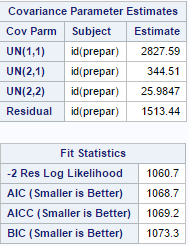
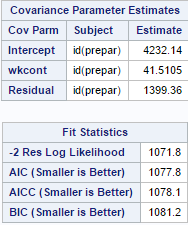


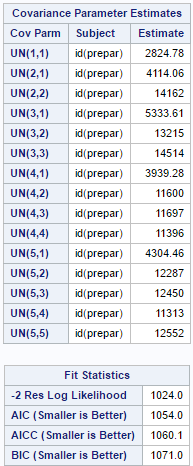
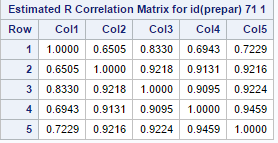
Table 3. Random intercept random slope with VC structure **G** matrix.



* 1. Now consider a model with no random effects, but specifying **R** to have the UN structure. Fit this model and compare with the previous ones.

The AIC for this model is smaller (1054.0) compared to the previous models with random intercept and random slope with UN G matrix and the model with VC structure G matrix. There were problems with the G matrix not being positive definite for the UN G matrix models but the V matrix indicated a variance structure that was reasonable. The fit is better with the no random effects UN R model but there are 10 parameters that need to be estimated which may be a problem with this small of a dataset. A close inspection on the covariance parameter indicates that the response by time increases then decreases compared to the baseline then decreases and it is hard to pinpoint the pattern in the change in variance using the no random effect UN R specification.

Table 4. No random effect UN structure for **R** model covariance parameter estimates, correlation matrix for subject 71 and fit statistics.

* 1. Write a paragraph about your findings. Include a couple of specific tests to support your story (with custom needs, if you’d like). You can also include a table with goodness-of-fit statistics for models fit in the previous parts. Out of all the models, which one would you use for your ‘final’ model? Why?

Based on the different fitting of the models, the addition of random intercept for subjects and random slope for time (in weeks) had a slightly higher AIC value compared to a model without random effects but inducing a **R** structure that has UN structure (Table 9). This model with no random effect was not used for the final model, in fact any model with UN **R** was not appropriate because too many covariance parameters are estimated for such a small sample size thus the next best fitting model was used which was a model with random intercept for subjects and random slope for time with UN **G** and VC **R.** In a sense, this model is a lot more parsimonious and has the lowest AIC value compared to the other models described in Table 9. Another issue with our final model was that the estimated **G** matrix was not positive definite but after examination of the **G** and the **V** matrix, which seemed reasonable, in that the variance between responses increased (as expected in our figure above), the final model with random intercept for subjects and random slope for time (in weeks, class) was still used in order to have a realistic fitted **V** matrix. Technically, it is really hard to decide on a model with such a short time period and a small sample set like this but with what is given and based on the trends visible in Figure 1, the V matrix makes sense and seems to be the next best fit to fitting a potentially problematic number (15) of covariance parameters.

Using the final model, comparison of different pill groups were conducted. There was no statistically significant difference in plasma beta carotene level over time between the group 4 vs. group 2 and group 1 vs. group 2 but there was a statistically significant difference in plasma beta carotene level over time for group 3 vs. group 2 (Table 5).

There was also no statistically significant difference in mean plasma beta carotene level between group 4 vs. group 2, group 3 vs. group 2, and group 1 vs. group 2 (Table 6).

The average of group 4, 3, and 1 vs. group 2 also indicated no difference in plasma beta carotene level over time (Table 7).

Lastly, comparing the 12 week vs. baseline measure of plasma beta carotene level of group 4,3,1 vs. group 2, the results indicate that only the difference between group 3 and group 2 were statistically significant (p=0.0113) indicated in Table 8.

Table 5. Contrast for difference between different pill groups changing over time.

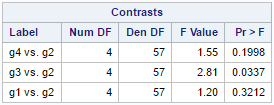


Table 6. Mean difference between groups in at least one time point.

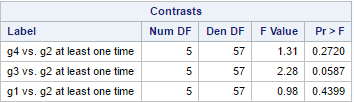


Table 7. Contrast for difference between average of g4,3,1 vs. g2 over time;

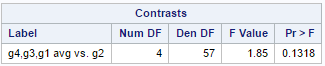


Table 8. Estimates for difference between 12 wk vs. baseline measures for g4,3,1, vs. g2.

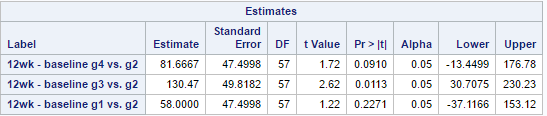


Table 9. Model specification and AIC for Final Model Selection.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model of week** | **Random effects** | **G** | **R** | **AIC** | **# of** | **Note** |
| Class | None | NA | UN | 1054 | 15 |  |
| Class | Rand. Int for id & Rand. slope for wk | UN | UN | 1060 | 18 | Hess and G not + def |
| Class | Rand. Int for id & Rand. slope for wk | UN | VC | 1068.7 | 4 | G not + def |
| Class | Rand. Int for id & Rand. slope for wk | UN | AR(1) | 1070.7 | 5 | G not + def |
| Class | Rand. Int for id & Rand. slope for wk | VC | VC | 1077.8 | 3 |  |
| Class | Rand. Int for id & Rand. slope for wk | VC | AR(1) | 1079.8 | 4 |  |
| Class | Random int for id | NA | VC | 1093.1 | 2 |  |
| Class | Random int for id | NA | AR(1) | 1094.9 | 3 |  |
| Class | None | NA | VC | 1191.3 | 1 |  |
| Cont. | Random int for id | NA | VC | 1245.1 | 2 |  |
| Class | Rand. Int for id & Rand. slope for wk | VC | UN | 1299 | 17 | G not + def |
| Class | Random int for id | NA | UN | 1299 | 16 | G not + def |
| Class | None | NA | AR(1) | 1097.8 | 2 |  |

* 1. **Not to turn in.** There are two baseline variables. There are different ways these variables could be used in an analysis. For example, you could simply use one or the other, or perhaps the average of the two. Suggest what you would use for a baseline measure. [Note: you receive the data not knowing for certain whether all the values are correct or not.]

We could check if the means are different from what is expected based on the overall mean for the subjects and pick the mean that is most consistent with the overall subject mean.

Appendix Codes.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*HW4 Longitudinal Betacarotene\*

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

\*Step 1: Data Management and Checking;

\*Step 1a: Import Data;

proc import datafile="/folders/myshortcuts/FS2016/Longitudinal/HW/HW4/beta carotene.txt"

out=bc

dbms=dlm

replace;

getnames=yes;

delimiter=' ';

run;

\*Step 1b:Check if the data has right categories (continuous or categorical;

proc contents data=bc;

run;

\*Step 1c:id and prepare needs to change to categorical/factor variable;

data bc;

set bc;

char\_id = put(id, 7.) ;

drop id ;

rename char\_id=id ;

char\_prepar = put(prepar, 7.) ;

drop Prepar;

rename char\_prepar=prepar ;

run;

proc contents data=bc;

run;

\*Drop the 1st baseline measure, preparer and change name of wk variables to just number;

data bc;

set bc;

drop Base1lvl;

rename Base2lvl = Wk0;

rename Wk6lvl = Wk6;

rename Wk8lvl = Wk8;

rename Wk10lvl = Wk10;

rename Wk12lvl = Wk12;

run;

proc sort data=bc;

by id;

run;

proc transpose data=bc out=bclong2;

by id;

run;

data bclong2;

set bclong2(rename=(col1=y));

Wk=input(substr(\_name\_, 3), 5.);

drop \_name\_;

run;

data bclong2;

merge bclong2

bc(keep=id

Prepar); /\*\* these variables are not transposed \*\*/

by id;

run;

proc contents data=bclong;

run; \*Step 2: Data visualization

\*HW4a: Create a graph for the data.

You have some flexibility on what type of graph to include;

\*baseline as outcome FINAL MODEL;

Proc Glimmix data=bclong2;

class id Prepar wk;

model y = Prepar wk Prepar\*wk;

lsmeans Prepar\*wk / plots=(meanplot(join sliceby=Prepar)) bylevel; \* no CI or SE bar yet;

run;

\*HW4c: analysis methods for efficacy of preparations over time;

proc mixed order=data data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept / subject=id(Prepar);

contrast 'linear' wk -3 -1 1 3;

contrast 'quadratic' wk 1 -1 -1 1;

contrast 'cubic' wk -1 3 -3 1;

lsmeans prepar\*wk; run;

\*Cubic trend for wk and lxl trend significant for prepar\*wk significant thus keep just the

lxl term and maybe add the wk wk^2 wk^3 in the model but not specified thus keep simple;

\*HW4d: analysis using random slope for time;

\*Define new time term for continuous slope;

data bclong2;

set bclong2;

wkcont = wk;

run;

\*Using UN for G matrix VC for R;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=UN subject=id(Prepar) v g ;

repeated /subject=id(Prepar);

lsmeans prepar\*wk; run;

\*Using VC for G matrix VC R;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=VC subject=id(Prepar) v g;

repeated /subject=id(Prepar);

lsmeans prepar\*wk; run;

\*Contrast for difference in groups changing over time;

\*Using UN for G matrix VC for R Final Model;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=UN subject=id(Prepar);

repeated /subject=id(Prepar) type=VC;

lsmeans prepar\*wk;

contrast 'g4 vs. g2' Prepar\*wk 0 0 0 0 0 1 -1 0 0 0 0 0 0 0 0 -1 1 0 0 0,

Prepar\*wk 0 0 0 0 0 1 0 -1 0 0 0 0 0 0 0 -1 0 1 0 0,

Prepar\*wk 0 0 0 0 0 1 0 0 -1 0 0 0 0 0 0 -1 0 0 1 0,

Prepar\*wk 0 0 0 0 0 1 0 0 0 -1 0 0 0 0 0 -1 0 0 0 1;

contrast 'g3 vs. g2' Prepar\*wk 0 0 0 0 0 1 -1 0 0 0 -1 1 0 0 0 0 0 0 0 0,

Prepar\*wk 0 0 0 0 0 1 0 -1 0 0 -1 0 1 0 0 0 0 0 0 0,

Prepar\*wk 0 0 0 0 0 1 0 0 -1 0 -1 0 0 1 0 0 0 0 0 0,

Prepar\*wk 0 0 0 0 0 1 0 0 0 -1 -1 0 0 0 1 0 0 0 0 0;

contrast 'g1 vs. g2' Prepar\*wk -1 1 0 0 0 1 -1 0 0 0 0 0 0 0 0 0 0 0 0 0,

Prepar\*wk -1 0 1 0 0 1 0 -1 0 0 0 0 0 0 0 0 0 0 0 0,

Prepar\*wk -1 0 0 1 0 1 0 0 -1 0 0 0 0 0 0 0 0 0 0 0,

Prepar\*wk -1 0 0 0 1 1 0 0 0 -1 0 0 0 0 0 0 0 0 0 0;

run;

\*Contrast for mean difference between groups in at least one time point;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=UN subject=id(Prepar);

repeated /subject=id(Prepar) type=VC;

lsmeans prepar\*wk;

contrast 'g4 vs. g2 at least one time'

Prepar 0 1 0 -1 Prepar\*wk 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1 0 0 0 0,

Prepar 0 1 0 -1 Prepar\*wk 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1 0 0 0,

Prepar 0 1 0 -1 Prepar\*wk 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1 0 0,

Prepar 0 1 0 -1 Prepar\*wk 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1 0,

Prepar 0 1 0 -1 Prepar\*wk 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 -1;

contrast 'g3 vs. g2 at least one time'

Prepar 0 1 -1 0 Prepar\*wk 0 0 0 0 0 1 0 0 0 0 -1 0 0 0 0 0 0 0 0,

Prepar 0 1 -1 0 Prepar\*wk 0 0 0 0 0 0 1 0 0 0 0 -1 0 0 0 0 0 0 0,

Prepar 0 1 -1 0 Prepar\*wk 0 0 0 0 0 0 0 1 0 0 0 0 -1 0 0 0 0 0 0,

Prepar 0 1 -1 0 Prepar\*wk 0 0 0 0 0 0 0 0 1 0 0 0 0 -1 0 0 0 0 0,

Prepar 0 1 -1 0 Prepar\*wk 0 0 0 0 0 0 0 0 0 1 0 0 0 0 -1 0 0 0 0;

contrast 'g1 vs. g2 at least one time'

Prepar -1 1 0 0 Prepar\*wk -1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0,

Prepar -1 1 0 0 Prepar\*wk 0 -1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0,

Prepar -1 1 0 0 Prepar\*wk 0 0 -1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0,

Prepar -1 1 0 0 Prepar\*wk 0 0 0 -1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0,

Prepar -1 1 0 0 Prepar\*wk 0 0 0 0 -1 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0;

run;

\*Contrast in avg g4,3,1 vs. g2;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=UN subject=id(Prepar);

repeated /subject=id(Prepar) type=VC;

lsmeans prepar\*wk;

contrast 'g4,g3,g1 avg vs. g2' Prepar\*wk 1 -1 0 0 0 -3 3 0 0 0 1 -1 0 0 0 1 -1 0 0 0,

Prepar\*wk 1 0 -1 0 0 -3 0 3 0 0 1 0 -1 0 0 1 0 -1 0 0,

Prepar\*wk 1 0 0 -1 0 -3 0 0 3 0 1 0 0 -1 0 1 0 0 -1 0,

Prepar\*wk 1 0 0 0 -1 -3 0 0 0 3 1 0 0 0 -1 1 0 0 0 -1;

run;

\*Estimates for mean change difference;

proc mixed data=bclong2;

class id Prepar wk;

model y= Prepar wk prepar\*wk / solution;

random intercept wkcont/ type=UN subject=id(Prepar);

repeated /subject=id(Prepar) type=VC;

lsmeans prepar\*wk;

estimate '12wk - baseline g4 vs. g2' Prepar\*wk 0 0 0 0 0 1 0 0 0 -1 0 0 0 0 0 -1 0 0 0 1 / cl;

estimate '12wk - baseline g3 vs. g2' Prepar\*wk 0 0 0 0 0 1 0 0 0 -1 -1 0 0 0 1 0 0 0 0 0 / cl;

estimate '12wk - baseline g1 vs. g2' Prepar\*wk -1 0 0 0 1 1 0 0 0 -1 0 0 0 0 0 0 0 0 0 0 / cl;

run;