02_question2

44

2022-06-01

Q2 Carry out the steps of the statistical analysis outlined below.

Turn in **commented statistical code** used to obtain the results reported. Use no more than **3 decimal places**. Lower-limb amputation (LLA) results primarily from complications of severe peripheral artery disease, diabetes mellitus, or trauma. Sustained exercise has been suggested by previous studies as a promising rehabilitation target to improve the long-term health after LLA. Using telehealth, a clinical trial has been conducted at the Veterans Affairs Eastern Colorado Regional Amputation Center to test the potential of sustaining walking exercise using **exercise self-management (EXP)** versus an **attention-control education (CTL)** program.

The **primary outcome** is *step count per day;* A few other **secondary outcomes** are also of interest including *presence of any adverse events*, as binary outcome. Outcomes were collected at **baseline (time=0)**, **6 months (time=6)** and **18 months (time=18)**. The primary hypothesis of the trial was that **the intervention would improve daily steps in at least 1000 steps**. Because randomization was **stratified by age (**< 60 **and** $geq 60$**)** and **level of amputation (below knee and above knee)**, these two variables should be adjusted for in the analyses. The longitudinal dataset of the study is in the table `data-lla.csv` with data dictionary below:

treat: `treat = EXP` for individuals in EXP group, and `treat = CTL` for individuals in CTL group
stepc: Step count
amp.above: amp.above=1 for amputation above the knee, amp.above=0 for amputation below the knee
time: Month from baseline
older: `older=1` for individuals 60 or older, and `older=0` for individuals younger than 60 years of age

Use R or SAS to answer the following questions regarding the hypothesis above.

Part 2a The primary hypothesis of the study was to test whether there was a difference of more than 1000 steps in the change of step count/day at 6 months between exposure groups. This particular type of trial is often referred to as a superiority trial by a margin. Carry out this hypothesis test and interpret your results for a clinical collaborator.

i. **Write the model formally in mathematical form. Allow for a difference in daily step count at baseline.**

There are four different types of models can be used to solve only for 2a(i) - (iv):

1. a linear change score model (R code fit_2_0)
2. a linear mixed model with random intercept (R code fit_2_1)
3. a generalized linear mixed model for Poisson regression with random intercept (R code fit_2_2)
4. a generalized estimate equation model for Poisson regression with compound symmetry correlation structure (not included due to only include marginal model)

However, if we still need to consider the question for 2b, a linear mixed model or a generalized linear mixed are preferred. There are pros and cons for each model: For linear mixed model, the outcomes are counts as integers (not continuous), so there will be violations on the assumption for normally distributed residuals. However, it is easier to calculate expected mean, standard error, and statistics for given groups, as well as, to test the primary hypothesis of the numeric difference between group. For generalized linear mixed model on Poisson regressor, the model is perfect for count-outcomes. However the model can only provide the risk rate (log risk rate), it will be hard to calculate a meaningful standard error for given groups, so it is hard to test the primary hypothesis. Besides there might be over- or under-dispersion situation to mitigate the performance of the model.

According to figures and tables below in 2a(v): There is nonlinear trend for the trajectories in treatment group, it is reasonable to treat time as a categorical variable with (0, 6, and 18 three levels) other than a continuous variable. Due to the parallel trajectory patterns for subjects from different groups, a random intercept will be applied for both models. The time effects are different between different intervention groups, so a time, intervention interaction term must be included in the model. I will also suggest building regression model on risk difference in future studies, however the related materials are not included in the lectures.

The $\mathbf{R}_i$ and $\mathbf{G}_i$ is identical for every individual subject i. Here we only use to distinguish the individual variance-covariance $\mathbf{R}_i$ from a population level varaince-covariance $\mathcal{R}$ and $\mathcal{G}$

Linear mixed model with only random intercept

$$
\begin{aligned}
Y_{ij}|b_i &\sim Normal\left(\mu_{ij}, \sigma_\epsilon^2\right) \\
\mu_{ij} &= X_{ij}\beta + Z_i b_i \\
Y_{ij} &= \mu_{ij} + \epsilon_{ij} \\
&= X_{ij}\beta + Z_i b_i + \epsilon_{ij} \\
&= \beta_0 + \beta_1 \times treat_i + \beta_2 \times I(time6)_{ij} + \\
&\quad \beta_3 \times I(time18)_{ij} + \beta_4 \times old_i + \beta_5 \times amp_i + \\
&\quad \beta_6 \times [treat::I(time6)]_{ij} + \beta_7 \times [treat::I(time18)]_{ij} + Z_i b_i + \epsilon_{ij} \\
b_i &\sim Normal(0, \sigma_b^2) \\
\epsilon_{ij} &\sim Normal(0, \sigma_\epsilon^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Y}_i|\mathbf{b}_i &\sim Normal(\boldsymbol{\mu}_i, \mathbf{R}_i) \\
\boldsymbol{\mu}_i &= E[\mathbf{Y}_i|\mathbf{b}_i] = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i \\
\mathbf{R}_i &= Var[\mathbf{Y}_i|\mathbf{b}_i] = \sigma_\epsilon^2\mathbf{I} \\
\boldsymbol{\epsilon}_i &\sim Normal(\mathbf{0}, \mathbf{R}_i) \\
\mathbf{b}_i &\sim Normal(\mathbf{0}, \mathbf{G}_i) \\
\mathbf{V}_i &= Cov[\mathbf{Y}_i] = \mathbf{Z}_i\mathbf{G}_i\mathbf{Z}_i^T + \mathbf{R}_i \\
\mathbf{Y}_i &\sim Normal(\boldsymbol{\mu}_i, \mathbf{V}_i)
\end{aligned}
$$

Poisson generalized linear mixed model

$$
\begin{aligned}
Y_{ij}|b_i &\sim Poisson(\lambda) \\
\eta_{ij} &= log\left(E\left[Y_{ij}|b_i\right]\right) = log(\lambda|b_i) = X_{ij}\beta + Z_i b_i \\
&= \beta_0 + \beta_1 \times treat_i + \beta_2 \times I(time6)_{ij} + \\
&\quad \beta_3 \times I(time18)_{ij} + \beta_4 \times old_i + \beta_5 \times amp_i + \\
&\quad \beta_6 \times [treat::I(time6)]_{ij} + \beta_7 \times [treat::I(time18)]_{ij} + Z_i b_i \\
b_i &\sim Normal(0, \sigma_b^2)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{Y}_i|\mathbf{b}_i &\sim Poisson(\lambda_i) \\
\lambda_i &= E[\mathbf{Y}_i|\mathbf{b}_i] \\
\boldsymbol{\eta}_i &= \log(\lambda_i) = \log(E[\mathbf{Y}_i|\mathbf{b}_i]) = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i \\
\mathbf{b}_i &\sim Normal(\mathbf{0}, \mathbf{G}_i)
\end{aligned}
$$

In the Poisson regression case, the interpretation for βs are risk or risk ratios, so the variance is not meaningful for the variance for numeric difference (in primary hypothesis). Hence, we must use a first order Taylor series approximation to get the standard error for difference on step counts.

$$
\begin{aligned}
log(Z) &\approx log(E[Z]) + (Z - E[Z])E[Z] \\
E[log(Z)] &\approx logE[Z] \\
Var[log(Z)] &\approx E[Z]^{-2}Var[Z] \\
Var[Z] &= Var[log(Z)] \times e^{2E[logZ]} \\
sd[Z] &= sd[log(Z)] \times e^{E[logZ]}
\end{aligned}
$$

ii. Write down the null and alternative hypothesis of interest in terms of the model parameters specified in (a); keep the terms consistent with (a).

Linear mixed model with only random intercept

Null hypothesis: After adjusted for the other variables, there is no difference between LLA patients from the treatment group and control group, on *whether the change of step counts over 1000 at the 6th month*.

Alternative hypothesis: After adjustment for baseline step counts and stratification on age and amputation location, treatment as a statistically significant effect on *whether the step counts over 1000 at the 6th month* over control group.

H_0: $\beta_6 \leq 1000$

H_1: $\beta_6 > 1000$

Poisson generalized linear mixed model with only random intercept

For the older age group and amputation location above:

$H_0 : e^{\beta_0+\beta_1+\beta 4+\beta_5}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta 4+\beta_5}\left(e^{\beta_2}-1\right) \leq 1000$

$H_1 : e^{\beta_0+\beta_1+\beta 4+\beta_5}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta 4+\beta_5}\left(e^{\beta_2}-1\right) > 1000$

For the younger age group and amputation location above:

$H_0 : e^{\beta_0+\beta_1+\beta_5}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta_5}\left(e^{\beta_2}-1\right) \leq 1000$

$H_1 : e^{\beta_0+\beta_1+\beta_5}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta_5}\left(e^{\beta_2}-1\right) > 1000$

For the older age group and amputation location below:

$H_0 : e^{\beta_0+\beta_1+\beta_4}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta_4}\left(e^{\beta_2}-1\right) \leq 1000$

$H_1 : e^{\beta_0+\beta_1+\beta_4}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0+\beta_4}\left(e^{\beta_2}-1\right) > 1000$

For the younger age group and amputation location below:

$H_0 : e^{\beta_0+\beta_1}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0}\left(e^{\beta_2}-1\right) \leq 1000$

$H_1 : e^{\beta_0+\beta_1}\left(e^{\beta_2+\beta_6}-1\right) - e^{\beta_0}\left(e^{\beta_2}-1\right) > 1000$

iii. Calculate the change (from baseline) in expected means and standard error of the step counts at 6 months for each the intervention and the control group.

All three models are tested and presented below; details see the R code references.

Linear change score model

Intervention	Age	Amputation	Mean(t6-t0)	SD(t6-t0)
Treatment	Old	Above	1140.654	1.003
		Below	1139.431	0.794
	Young	Above	1141.197	0.889
		Below	1139.974	0.715
Control	Old	Above	59.498	0.97
		Below	58.275	0.774
	Young	Above	60.041	0.893
		Below	58.818	0.742

Linear mixed model with only random intercept

	Mean(t6-t0)	SD(t6-t0)
Treatment	1140.121	1.062
Control	-58.962	1.062

Poisson generalized linear mixed model with only random intercept

Intervention	Amputation	Age	log(Rate)	Rate	SD[log(Rate)]	Mean(t6-t0)	SD(t6-t0)
Treatment	Above	Old	0.349	1.418	0.002	1109.746	7.528
		Young	0.349	1.418	0.002	1115.309	7.566
	Below	Old	0.349	1.418	0.002	1078.026	7.313
		Young	0.349	1.418	0.002	1083.429	7.35
Control	Above	Old	-0.023	0.977	0.002	-50.768	4.921
		Young	-0.023	0.977	0.002	-51.022	4.945
	Below	Old	-0.023	0.977	0.002	-49.317	4.78
		Young	-0.023	0.977	0.002	-49.564	4.804

The standard deviation was calculated use the pooled standard deviation from both baseline and the 6th month step counts, with Taylor expansion:

$$\begin{aligned} log(Z) &\approx log(E[Z]) + (Z - E[Z])E[Z] \\ E[log(Z)] &\approx logE[Z] \\ Var[log(Z)] &\approx E[Z]^{-2}Var[Z] \\ Var[Z] &= Var[log(Z)] \times e^{2E[logZ]} \\ sd[Z] &= sd[log(Z)] \times e^{E[logZ]} \end{aligned}$$

The expected mean values from all three models are close to each other. According to the model, the effect size for older age and amputation location is very small, hence the expected mean values and the standard errors of *the change of step counts over 1000 at the 6th month from baseline* are very close to each other among different age group and amputation location groups.

The standard deviation from change score model and linear mixed model are close each other, however, due to the Poisson mixed model can only provide rate ratio for estimators, the standard deviations are much larger than the other models.

iv. Calculate the difference (EXP - CLT) in change of expected means in the two intervention groups. Report a 1-sided 95% confidence interval for this difference.

Linear change score model

The 95% confidence interval for *the change of step counts over 1000 at 6 month* is $(1079.822, \infty)$ for one-side approximated normal distribution.

Linear mixed model with only random intercept

The 95% confidence interval for *the change of step counts over 1000 at 6 month* is $(1196.614, \infty)$ for one-side approximated normal distribution.

Poisson generalized linear mixed model with only random intercept

Amputation	Age	Mean(Trt - Ctl)	SD(Trt-Ctl)	95% CI	
Above	Old	1160.514	8.994	1145.721	Inf
	Young	1166.331	9.039	1151.464	Inf
Below	Old	1127.343	8.737	1112.973	Inf
	Young	1132.993	8.781	1118.55	Inf

v. Make a graph, and include any appropriate descriptive statistics, that you would present as evidence related to the primary hypothesis of interest.

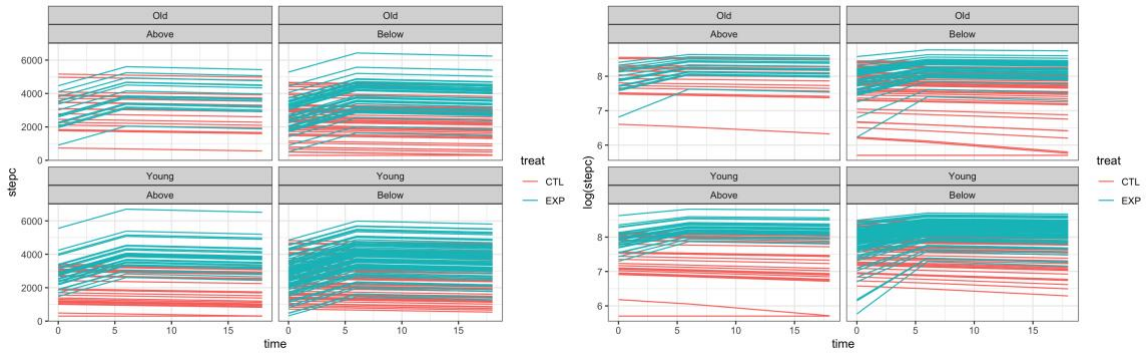

Figure1: Trajectories for each subject over time with Step and log(Step) stratified by Age group and Amputation location

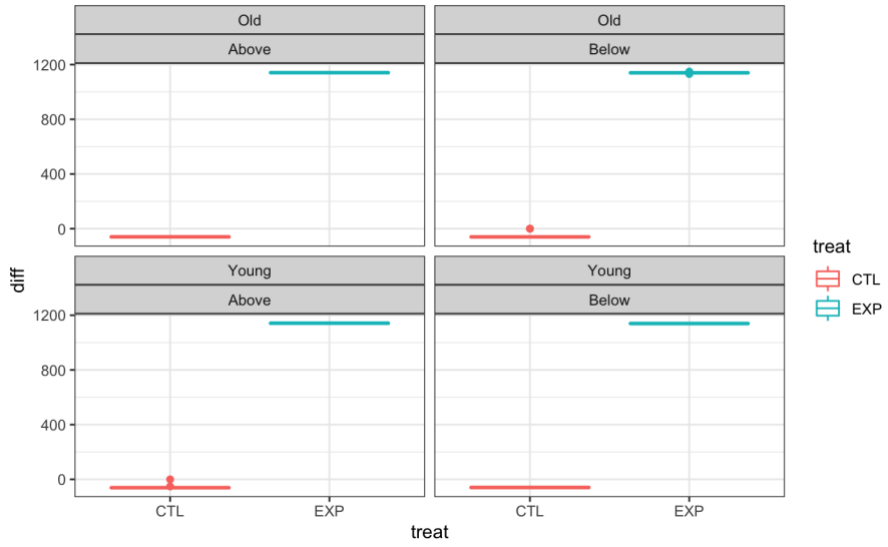

Figure2: Boxplot for changes of steps over baseline and the 6th month, stratified by Age group and Amputation location. If the mean value for different intervention (the gap) is larger than 1000 steps, then we can reject the null hypothesis.

vi. Write your conclusion of the hypothesis test in (b).

Based on the linear mixed model `fit_2_1` as well as the generalized linear mixed model `fit_2_2`, conditional on given individual adjusted for age group and operational amputation location, we can reject the null hypothesis and claim there is a very highly statistically significant difference on *the change of step counts over 1000 at the 6th month* between two intervention groups (p-value < 0.05).

vii. Comment on your assumptions of the correlation structure for the repeated measurements within subject.

For the linear mixed model and generalized linear mixed model, the conditional model with only random intercept is equivalent to a marginal GEE model with compound symmetry covariance structure. The correlation structure for repeated measure is assumed to be compound symmetric structure, with ICC as $\rho = \frac{\sigma_b^2}{\sigma_b^2+\sigma_\epsilon^2}$.

viii. Write a summary paragraph interpreting the results for a clinical collaborator.

For linear mixed model, on average in population level, after adjusted for age group and amputation location, we found that *the change of step counts over 1000 at the 6th month* in treatment group is 1199 steps more than the control group (sd = 1.5). Hence *the change of step counts over 1000 at the 6th month* from different intervention groups is larger than 1000 with both practical and statistical significance. Therefore, we can reject the null hypothesis and claim that after adjustment for baseline step counts and stratification on

age and amputation location, treatment as a statistically significant effect on *whether the step counts over 1000 at the 6th month* over control group.

For generalized linear mixed model with Poisson regression, the conclusion is very similar to linear mixed model. However, the estimators should be interpreted as baseline rate and rate ratios.

Part 2b A secondary hypothesis of interest was whether or not the change in step counts was sustained from 6 to 18 months, specifically, interest was in testing the difference between M18-M0 and M6-M0 in the intervention group.

Carry out the appropriate hypothesis to test whether there was a sustained change and interpret the results for a clinical collaborator.

Linear mixed model with only random intercept

H_0: There is no difference between *the change of step counts over 1000 at the 6th month from baseline* and *the change of step counts over 1000 at the 18th month from baseline*

equation form: $\beta_3 + \beta_7 - \beta_2 - \beta_6 = 0$

H_1: There is a statistically significant difference between *the change of step counts over 1000 at the 6th month from baseline* and *the change of step counts over 1000 at the 18th month from baseline*

equation form: $\beta_3 + \beta_7 - \beta_2 - \beta_6 \neq 0$

Poisson generalized linear mixed model with only random intercept

H_0: There is no difference between *the change of step counts over 1000 at the 6th month from baseline* and *the change of step counts over 1000 at the 18th month from baseline*

equation form: $e^{\beta_3+\beta_7-\beta_2-\beta_6} = 1$

H_1: There is a statistically significant difference between *the change of step counts over 1000 at the 6th month from baseline* and *the change of step counts over 1000 at the 18th month from baseline*

equation form: $e^{\beta_3+\beta_7-\beta_2-\beta_6} \neq 1$

As we can see in both linear mixed model and generalized linear mixed model, with Wald tests, the Wald statistics follow χ^2. There is a change of -180.428 steps change between *the change of step counts over 1000 at the 6th month from baseline* and *the change of step counts over 1000 at the 18th month from baseline* in linear mixed model with (p-value < 0.001); *the change of step counts over 1000 at the 18th month from baseline* is 4.781% less than the *change of step counts over 1000 at the 6th month from baseline* in Poisson generalized linear mixed model with (p-value < 0.001). In both cases, the difference is statistically significant,

hence we can reject the null hypothesis and claim that *the change in step counts* was not sustained from the 6th to the 18th month.

R Code Reference

```
## import dataset data_lla
data_lla <- read_csv("data/data-lla.csv",
                     col_types = cols(
                       ## amputation location as binary variable
                       amp.above = col_factor(levels = c("0", "1")),
                       ## older age or not as binary variable
                       older = col_factor(levels = c("0", "1"))))


data_lla_diff6 <- data_lla %>%
  ## change the from long format into wide format
  pivot_wider(names_from = "time",
              values_from = "stepc") %>%
  ## calculated the different from time0 to time6
  mutate(step_diff = abs(`6` - `0`)) %>%
  ## binary outcome of whether the change of step is over 1000
  mutate(step_diff_1k = case_when(step_diff >= 1000 ~ 1,
                                  step_diff < 1000 ~ 0),
         baseline = `0`)

# View(data_lla_diff6)
```

using change score model

```
fit_2_0 <- lm(step_diff ~ 1 + treat + older + amp.above,
              # control = fit_control,
              data = data_lla_diff6)

## to get the coefficients
beta_2_0 <- coef(fit_2_0) %>% round(digits = 3)
beta_2_0

## (Intercept)    treatEXP      older1  amp.above1
##      58.818    1081.156      -0.543       1.223

## to get the variance-covariance
vcov_2_0 <- vcov(fit_2_0) %>% round(digits = 3)
vcov_2_0

##             (Intercept) treatEXP older1 amp.above1
## (Intercept)       0.551   -0.349 -0.316     -0.266
## treatEXP         -0.349    0.658  0.036      0.017
## older1           -0.316    0.036  0.680      0.048
## amp.above1       -0.266    0.017  0.048      0.778

##                  [1]       [2]       [3]       [4]
## coefs           (Int)     treatEXP  old1      amp.above1
con_trt_old_abv <- c(1,        1,        1,        1)
con_trt_old_blw <- c(1,        1,        1,        0)
con_trt_yng_abv <- c(1,        1,        0,        1)
con_trt_yng_blw <- c(1,        1,        0,        0)

con_ctl_old_abv <- c(1,        0,        1,        1)
con_ctl_old_blw <- c(1,        0,        1,        0)
con_ctl_yng_abv <- c(1,        0,        0,        1)
con_ctl_yng_blw <- c(1,        0,        0,        0)

## this is the contrast of all the
contr1 <- cbind(con_trt_old_abv,
                con_trt_old_blw,
                con_trt_yng_abv,
```

```
                con_trt_yng_blw,
                con_ctl_old_abv,
                con_ctl_old_blw,
                con_ctl_yng_abv,
                con_ctl_yng_blw)

rownames(contr1) <- rownames(vcov_2_0)

mu_2_0 <- t(contr1) %*% beta_2_0 %>%
  as.data.frame() %>%
  dplyr::select("expected_mean" = 1)
## contrast variance covariance matrix
sd_2_0 <- t(contr1) %*% vcov_2_0 %*% contr1 %>%
  diag() %>% sqrt()
cbind(mu_2_0, "standard_error" = sd_2_0) %>%
  round(3)

##                 expected_mean standard_error
## con_trt_old_abv      1140.654          1.003
## con_trt_old_blw      1139.431          0.794
## con_trt_yng_abv      1141.197          0.889
## con_trt_yng_blw      1139.974          0.715
## con_ctl_old_abv        59.498          0.970
## con_ctl_old_blw        58.275          0.774
## con_ctl_yng_abv        60.041          0.893
## con_ctl_yng_blw        58.818          0.742

mu_2_01 <- t(contr1_1) %*% beta_2_0 %>%
  as.data.frame() %>%
  dplyr::select("expected_mean" = 1)
## contrast variance covariance matrix
sd_2_01 <- t(contr1_1) %*% vcov_2_0 %*% contr1_1 %>%
  diag() %>% sqrt()
cbind(mu_2_01, "standard_error" = sd_2_01) %>%
  round(3)

##   expected_mean standard_error
## 1      1081.156          0.811
## 2      1081.156          0.811
## 3      1081.156          0.811
## 4      1081.156          0.811

qnorm(0.05, 1081.156,   0.811)

## [1] 1079.822
```

using linear mixed model

```
data_lla_tf <- data_lla %>%
  mutate(time = as.factor(time))


fit_2_1 <- lmer(stepc ~ 1 + treat * time + older + amp.above +
                  + (1|id),
                # control = fit_control,
                data = data_lla_tf)

## to get the coefficients
beta_2_1 <- (summary(fit_2_1)$coefficients[, 1]) %>% round(digits = 3)
# beta_2_1
## to get the variance-covariance
vcov_2_1 <- vcov(fit_2_1)
# vcov_2_1

## lmm contrast-----------------------------------------------------------------------
##                     [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs              (Int)  treatEXP  time6 time18  older1  amp.abv1  trtEXP:time6  trtEXP:time18
```

```
con_trt_t0 <-   c(1,    1,        0,    0,      1,      1,        0,            0)
con_trt_t6 <-   c(1,    1,        1,    0,      1,      1,        1 * 1,        0)
con_trt_t18 <-  c(1,    1,        0,    1,      1,      1,        0,            1 * 1)
con_ctl_t0 <-   c(1,    0,        0,    0,      1,      1,        0,            0)
con_ctl_t6 <-   c(1,    0,        1,    0,      1,      1,        0,            0)
con_ctl_t18 <-  c(1,    0,        0,    1,      1,      1,        0,            0)
con_trt_t6_t0 <- con_trt_t6 - con_trt_t0
con_ctl_t6_t0 <- con_ctl_t6 - con_ctl_t0

contrast0 <- cbind(con_trt_t6_t0, con_ctl_t6_t0)
mu_lm <- t(contrast0) %*% beta_2_1
## standard error for each group
sd_lm <- t(contrast0) %*% vcov_2_1 %*% contrast0 %>%
  diag() %>% sqrt()

cbind(mu_lm, sd_lm) %>%
  as.data.frame() %>%
  dplyr::select("expected mean" = 1,
                "standard_error" = 2) %>%
  round(3)

##               expected mean standard_error
## con_trt_t6_t0      1140.121          1.062
## con_ctl_t6_t0       -58.962          1.062

t <- (1140.121 - (-58.962))/ 1.062; t

## [1] 1129.08

con_trt_t6_t0 <- con_trt_t6 - con_trt_t0
con_ctl_t6_t0 <- con_ctl_t6 - con_ctl_t0

con_tvc_t6_t0 <- con_trt_t6_t0 - con_ctl_t6_t0
mu_2_3 <- t(con_tvc_t6_t0 ) %*% beta_2_1 %>%
  as.data.frame() %>%
  dplyr::select("expected_mean" = 1)
## contrast variance covariance matrix
sd_2_3 <- t(con_tvc_t6_t0 ) %*% vcov_2_1 %*% con_tvc_t6_t0  %>%
  diag() %>% sqrt()

cbind(mu_2_3, "standard_error" = sd_2_3) %>%
  round(3)

##   expected_mean standard_error
## 1      1199.083          1.501

qnorm(0.05, mean = 1199.083, sd = 1.501)

## [1] 1196.614

pnorm(1000, mean = 1199.083, sd = 1.501)

## [1] 0
```

using generalized linear mixed model

```
fit_2_2 <- glmer(stepc ~ 1 + treat * time + older + amp.above +
                   + (1|id),
                 family = "poisson",
                 # control = fit_control,
                 data = data_lla_tf)

## Warning in checkConv(attr(opt, "derivs"), opt$par, ctrl = control$checkConv, : Model is nearly unidenti
fiable: very large eigenvalue
##  - Rescale variables?
```

```r
## to get the coefficients
beta_2_2 <- (summary(fit_2_2)$coefficients[, 1]) %>% round(digits = 3)
# beta_2_2
## to get the variance-covariance
vcov_2_2 <- vcov(fit_2_2)
# vcov_2_2

## old amputation above ------------------------------------------------------------------------
##                      [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs              (Int)  treatEXP  time6 time18  older1  amp.abv1  trtEXP:time6  trtEXP:time18
con_old_abv_trt_t0 <-   c(1,    1,        0,    0,      1,      1,        0,            0)
con_old_abv_trt_t6 <-   c(1,    1,        1,    0,      1,      1,        1 * 1,        0)
con_old_abv_trt_t18 <-  c(1,    1,        0,    1,      1,      1,        0,            1 * 1)
(con_old_abv_trt_t6_t0 <- con_old_abv_trt_t6 - con_old_abv_trt_t0)

## [1] 0 0 1 0 0 0 1 0

(con_old_abv_trt_t18_t0 <- con_old_abv_trt_t18 - con_old_abv_trt_t0)

## [1] 0 0 0 1 0 0 0 1

con_old_abv_ctl_t0 <-   c(1,    0,        0,    0,      1,      1,        0,            0)
con_old_abv_ctl_t6 <-   c(1,    0,        1,    0,      1,      1,        0,            0)
con_old_abv_ctl_t18 <-  c(1,    0,        0,    1,      1,      1,        0,            0)
(con_old_abv_ctl_t6_t0 <- con_old_abv_ctl_t6 - con_old_abv_ctl_t0)

## [1] 0 0 1 0 0 0 0 0

(con_old_abv_ctl_t18_t0 <- con_old_abv_ctl_t18 - con_old_abv_ctl_t0)

## [1] 0 0 0 1 0 0 0 0

## young amputation above ----------------------------------------------------------------------
##                      [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs              (Int)  treatEXP  time6 time18  older1  amp.abv1  trtEXP:time6  trtEXP:time18
con_yng_abv_trt_t0 <-   c(1,    1,        0,    0,      1,      0,        0,            0)
con_yng_abv_trt_t6 <-   c(1,    1,        1,    0,      1,      0,        1 * 1,        0)
con_yng_abv_trt_t18 <-  c(1,    1,        0,    1,      1,      0,        0,            1 * 1)
con_yng_abv_trt_t6_t0 <- con_yng_abv_trt_t6 - con_yng_abv_trt_t0
con_yng_abv_trt_t18_t0 <- con_yng_abv_trt_t18 - con_yng_abv_trt_t0
con_yng_abv_ctl_t0 <-   c(1,    0,        0,    0,      1,      0,        0,            0)
con_yng_abv_ctl_t6 <-   c(1,    0,        1,    0,      1,      0,        0,            0)
con_yng_abv_ctl_t18 <-  c(1,    0,        0,    1,      1,      0,        0,            0)
con_yng_abv_ctl_t6_t0 <- con_yng_abv_ctl_t6 - con_yng_abv_ctl_t0
con_yng_abv_ctl_t18_t0 <- con_yng_abv_ctl_t18 - con_yng_abv_ctl_t0
## old amputation below ------------------------------------------------------------------------
##                      [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs              (Int)  treatEXP  time6 time18  older1  amp.abv1  trtEXP:time6  trtEXP:time18
con_old_blw_trt_t0 <-   c(1,    1,        0,    0,      0,      1,        0,            0)
con_old_blw_trt_t6 <-   c(1,    1,        1,    0,      0,      1,        1 * 1,        0)
con_old_blw_trt_t18 <-  c(1,    1,        0,    1,      0,      1,        0,            1 * 1)
con_old_blw_trt_t6_t0 <- con_old_blw_trt_t6 - con_old_blw_trt_t0
con_old_blw_trt_t18_t0 <- con_old_blw_trt_t18 - con_old_blw_trt_t0
con_old_blw_ctl_t0 <-   c(1,    0,        0,    0,      0,      1,        0,            0)
con_old_blw_ctl_t6 <-   c(1,    0,        1,    0,      0,      1,        0,            0)
con_old_blw_ctl_t18 <-  c(1,    0,        0,    1,      0,      1,        0,            0)
con_old_blw_ctl_t6_t0 <- con_old_blw_ctl_t6 - con_old_blw_ctl_t0
con_old_blw_ctl_t18_t0 <- con_old_blw_ctl_t18 - con_old_blw_ctl_t0
## young amputation below ----------------------------------------------------------------------
##                      [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs              (Int)  treatEXP  time6 time18  ynger1  amp.abv1  trtEXP:time6  trtEXP:time18
con_yng_blw_trt_t0 <-   c(1,    1,        0,    0,      0,      0,        0,            0)
con_yng_blw_trt_t6 <-   c(1,    1,        1,    0,      0,      0,        1 * 1,        0)
con_yng_blw_trt_t18 <-  c(1,    1,        0,    1,      0,      0,        0,            1 * 1)
con_yng_blw_trt_t18_t0 <- con_yng_blw_trt_t18 - con_yng_blw_trt_t0
con_yng_blw_trt_t6_t0 <- con_yng_blw_trt_t6 - con_yng_blw_trt_t0
con_yng_blw_ctl_t0 <-   c(1,    0,        0,    0,      0,      0,        0,            0)
con_yng_blw_ctl_t6 <-   c(1,    0,        1,    0,      0,      0,        0,            0)
```

```
con_yng_blw_ctl_t18 <-  c(1,    0,        0,    1,      0,      0,        0,            0)
con_yng_blw_ctl_t18_t0 <- con_yng_blw_ctl_t18 - con_yng_blw_ctl_t0
con_yng_blw_ctl_t6_t0 <- con_yng_blw_ctl_t6 - con_yng_blw_ctl_t0

# contast_old_abv_t06 <- cbind(con_old_abv_trt_t6_t0, con_old_abv_ctl_t6_t0)
# contast_yng_abv_t06 <- cbind(con_yng_abv_trt_t6_t0, con_yng_abv_ctl_t6_t0)
# contast_old_blw_t06 <- cbind(con_old_blw_trt_t6_t0, con_old_blw_ctl_t6_t0)
# contast_yng_blw_t06 <- cbind(con_yng_blw_trt_t6_t0, con_yng_blw_ctl_t6_t0)

contrast1 <- cbind(con_old_abv_trt_t6_t0,
                   con_yng_abv_trt_t6_t0,
                   con_old_blw_trt_t6_t0,
                   con_yng_blw_trt_t6_t0,
                   con_old_abv_ctl_t6_t0,
                   con_yng_abv_ctl_t6_t0,
                   con_old_blw_ctl_t6_t0,
                   con_yng_blw_ctl_t6_t0)
rownames(contrast1) <- rownames(vcov_2_2)

con_ctl_t0 <- cbind(con_old_abv_trt_t0,
                    con_yng_abv_trt_t0,
                    con_old_blw_trt_t0,
                    con_yng_blw_trt_t0,
                    con_old_abv_ctl_t0,
                    con_yng_abv_ctl_t0,
                    con_old_blw_ctl_t0,
                    con_yng_blw_ctl_t0)

con_ctl_t6 <- cbind(con_old_abv_trt_t6,
                    con_yng_abv_trt_t6,
                    con_old_blw_trt_t6,
                    con_yng_blw_trt_t6,
                    con_old_abv_trt_t6,
                    con_yng_abv_trt_t6,
                    con_old_blw_trt_t6,
                    con_yng_blw_trt_t6)

mu_ctl_t0 <- t(con_ctl_t0) %*% beta_2_2 %>%
  exp() %>%
  as.data.frame() %>%
  dplyr::select("t0" = 1)
sd_ctl_t0 <- t(con_ctl_t0) %*% vcov_2_2 %*% con_ctl_t0 %>%
  diag() %>% sqrt()
mu_ctl_t6 <- t(con_ctl_t6) %*% beta_2_2 %>%
  exp() %>%
  as.data.frame() %>%
  dplyr::select("t6" = 1)
sd_ctl_t6 <- t(con_ctl_t6) %*% vcov_2_2 %*% con_ctl_t6 %>%
  diag() %>% sqrt()


mu_2_2 <- t(contrast1) %*% beta_2_2 %>%
  as.data.frame() %>%
  dplyr::select("log_rate" = 1) %>%
  mutate(rate = exp(log_rate))
## contrast variance covariance matrix
sd_2_2 <- t(contrast1) %*% vcov_2_2 %*% contrast1 %>%
  diag() %>% sqrt()

result2 <- cbind(mu_2_2,
      sd_log_rate = sd_2_2,
      mu_ctl_t0,
      mu_ctl_t6) %>%
  mutate(diff_t6_t0 = (rate - 1) * t0,
         sd_t6_t0 = abs(rate * sd_2_2 * t0)) %>%
  round(3)
result2
```

```
##                       log_rate  rate sd_log_rate       t0       t6 diff_t6_t0
## con_old_abv_trt_t6_t0    0.349 1.418       0.002 2657.125 3766.871   1109.746
## con_yng_abv_trt_t6_t0    0.349 1.418       0.002 2670.444 3785.753   1115.309
## con_old_blw_trt_t6_t0    0.349 1.418       0.002 2581.175 3659.201   1078.026
## con_yng_blw_trt_t6_t0    0.349 1.418       0.002 2594.113 3677.542   1083.429
## con_old_abv_ctl_t6_t0   -0.023 0.977       0.002 2232.774 3766.871    -50.768
## con_yng_abv_ctl_t6_t0   -0.023 0.977       0.002 2243.966 3785.753    -51.022
## con_old_blw_ctl_t6_t0   -0.023 0.977       0.002 2168.953 3659.201    -49.317
## con_yng_blw_ctl_t6_t0   -0.023 0.977       0.002 2179.825 3677.542    -49.564
##                       sd_t6_t0
## con_old_abv_trt_t6_t0    7.528
## con_yng_abv_trt_t6_t0    7.566
## con_old_blw_trt_t6_t0    7.313
## con_yng_blw_trt_t6_t0    7.350
## con_old_abv_ctl_t6_t0    4.921
## con_yng_abv_ctl_t6_t0    4.945
## con_old_blw_ctl_t6_t0    4.780
## con_yng_blw_ctl_t6_t0    4.804
```

```
result3a <- result2[1:4, ]
result3b <- result2[5:8, ]

result4 <- data.frame(diff_trt_ctl = result3a$diff_t6_t0 - result3b$diff_t6_t0,
                      sd_trt_ctl = sqrt(result3a$sd_t6_t0^2 + result3b$sd_t6_t0^2)) %>%
  mutate(q05 = qnorm(0.05, mean = diff_trt_ctl, sd = sd_trt_ctl)) %>%
  round(3)

result4
```

```
##   diff_trt_ctl sd_trt_ctl      q05
## 1     1160.514      8.994 1145.721
## 2     1166.331      9.039 1151.464
## 3     1127.343      8.737 1112.973
## 4     1132.993      8.781 1118.550
```

```
plot_lla <- data_lla %>%
  mutate(older = case_when(older == 0 ~ "Young",
                           older == 1 ~ "Old"),
         amp.above = case_when(amp.above == 1 ~ "Above",
                               amp.above == 0 ~ "Below")) %>%
  ggplot(aes(x = time,
             y = stepc,
             group = id,
             color = treat)) +
  geom_line() +
  facet_wrap(older ~ amp.above)

plot_lla_log <- data_lla %>%
  mutate(older = case_when(older == 0 ~ "Young",
                           older == 1 ~ "Old"),
         amp.above = case_when(amp.above == 1 ~ "Above",
                               amp.above == 0 ~ "Below")) %>%
  ggplot(aes(x = time,
             y = log(stepc),
             group = id,
             color = treat)) +
  geom_line() +
  facet_wrap(older ~ amp.above)

plot_lla + theme_bw()

plot_lla_log + theme_bw()

plot_lla2 <- data_lla %>%
  mutate(older = case_when(older == 0 ~ "Young",
                           older == 1 ~ "Old"),
         amp.above = case_when(amp.above == 1 ~ "Above",
                               amp.above == 0 ~ "Below")) %>%
```

```
  pivot_wider(names_from = time,
              names_prefix = "t",
              values_from = stepc) %>%
  mutate(diff = t6 - t0) %>%
  ggplot(aes(y = diff,
             x = treat,
             color = treat)) +
  geom_boxplot() +
  facet_wrap(older ~ amp.above)

plot_lla2 + theme_bw()

## lmm contrast-----------------------------------------------------------------------
##               [0]    [1]       [2]   [3]     [4]     [5]       [6]           [7]
## coefs        (Int)  treatEXP  time6 time18  older1  amp.abv1  trtEXP:time6  trtEXP:time18
con_trt_t0 <-   c(1,    1,        0,    0,      1,      1,        0,            0)
con_trt_t6 <-   c(1,    1,        1,    0,      1,      1,        1 * 1,        0)
con_trt_t18 <-  c(1,    1,        0,    1,      1,      1,        0,            1 * 1)

con_trt_t6_t0 <- con_trt_t6 - con_trt_t0
con_trt_t18_t0 <- con_trt_t18 - con_trt_t0
contrast4 <- con_trt_t18_t0 - con_trt_t6_t0
contrast4

## [1]  0  0 -1  1  0  0 -1  1

mu_b <- t(contrast4) %*% beta_2_1
mu_b

##          [,1]
## [1,] -180.428

## contrast variance covariance matrix
cov_b <- t(contrast4) %*% vcov_2_1 %*% contrast4 %>% as.matrix()

## with both point estimates and standard deviation
## an anova or pairwise comparison can be performed
W0 <- mu_b^2 * solve(cov_b)
pchisq(W0, df = 1, lower.tail = FALSE)

##      [,1]
## [1,]    0

contrast5 <- con_old_abv_trt_t18_t0 - con_old_abv_trt_t6_t0
contrast5

## [1]  0  0 -1  1  0  0 -1  1

mu_b2 <- t(contrast5) %*% beta_2_2
exp(mu_b2)

##           [,1]
## [1,] 0.9521811

## contrast variance covariance matrix
cov_b2 <- t(contrast5) %*% vcov_2_2 %*% contrast5 %>% as.numeric()

W1 <- mu_b2^2 / cov_b2
pchisq(W1, df = 1, lower.tail = FALSE)

##               [,1]
## [1,] 3.323379e-156
```