Supplementary to MicroRex: Assesment of the Sublingual Microcirculation with the GlycoCheck Device: Reproducibility and Examination Conditions

1 Setup

We consider the dataset `microrex` containing the variables `PBR`, `PARTICIPANT`, `Sex`, `Visit_dag`, `Intervention`, `Laterality` and `mtime` as described in the main article.

```
str(microrex)

Classes 'tbl_df', 'tbl' and 'data.frame': 3990 obs. of  7 variables:
 $ PARTICIPANT : num  99001 99001 99001 99001 99001 ...
 $ PBR         : num  2.46 2.01 2.66 2.33 2.49 ...
 $ Intervention: chr  "none" "none" "none" "none" ...
 $ Laterality  : chr  "left" "left" "left" "left" ...
 $ Visit_dag   : num  1 1 1 1 1 1 1 1 1 1 ...
 $ mtime       : Factor w/ 10 levels "b","0","20","30",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ Sex         : chr  "Mand" "Mand" "Mand" "Mand" ...
 - attr(*, "label")= chr "MICROREX_SAMLET"
```

2 Linear Modeling

The goal of the analyses in this section is to understand the variation of `PBR` in the setting described in the main article. All calculations are performed in `R`, see [3].

We fit 3 different linear models with `PBR` as outcome.

- *The Interaction Model*: A linear model with all combinations of `PARTICIPANT` and `mtime` included as factor levels.

- *The Main Effects Model*: A linear model with main effects of all relevant covariates, including `PARTICIPANT`. This model is overparametrized.
- *The Main Effects Model with a random Person-effect*: A linear model with main effects of all relevant covariates, and `PARTICIPANT` as a random effect.

The residual variation is seen to be approximately constant independently of the model. Furthermore, the residual variation is seen to be bigger than the inter-person variation in the cases where the latter is explicitely modelled.

2.1 The Interaction Model

```
##Interaction model
fit.lm_full<- lm(PBR ~ as.factor(PARTICIPANT)*as.factor(mtime),
                 data=microrex)

##residual sd
summary(fit.lm_full)$sigma

[1] 0.2681924

##degrees of freedom
##rank of model matrix, residual degrees of freedom and no. of
##coefficients in the model
summary(fit.lm_full)$df

[1]  395 2851  420

##checking that no. of coefficients minus rank of model matrix is equal to
##no. of aliased parameters
sum(summary(fit.lm_full)$aliased)

[1] 25
```

2.2 The Main Effects Model

```
## model with main effects of the classification variables
fit.lm_class <- lm(PBR ~   as.factor(Visit_dag)+
                           as.factor(Laterality) +
                           as.factor(Intervention) +
                           mtime +
                           as.factor(PARTICIPANT),
                   data=microrex)
##residual sd
summary(fit.lm_class)$sigma
```

```
[1] 0.2720161

##degrees of freedom
summary(fit.lm_class)$df

[1]   54 3192   56

##no. of aliased parameters
sum(summary(fit.lm_class)$aliased)

[1] 2

## ANOVA for corresponding fit without mtime:
anova(fit.lm_class, update(fit.lm_class, ~.-mtime))

Analysis of Variance Table

Model 1: PBR ~ as.factor(Visit_dag) + as.factor(Laterality) + as.factor(Intervention)
    mtime + as.factor(PARTICIPANT)
Model 2: PBR ~ as.factor(Visit_dag) + as.factor(Laterality) + as.factor(Intervention)
    as.factor(PARTICIPANT)
  Res.Df    RSS Df Sum of Sq      F
1   3192 236.19
2   3201 245.85 -9   -9.6651 14.514
     Pr(>F)
1
2 < 2.2e-16 ***
---
Signif. codes:
  0 '***' 0.001 '**' 0.01 '*' 0.05
  '.' 0.1 ' ' 1
```

2.3 The Main Effects Model with a Random Person-effect

We can include a random person-effect to incorporate the inter-individual variation and compare it with the residual variation. We use the R-library `lme4`, see [1].

```
####################################
## Main effects + Random effect model:
library(lme4)
fit.lmer <- lmer(PBR ~ as.factor(Visit_dag)+
                       as.factor(Laterality) +
                       as.factor(Intervention) +
                       as.factor(Sex) +
                       mtime +
                      (1|PARTICIPANT),
                 data=transform(microrex, PARTICIPANT=as.factor(PARTICIPANT)))
```

```
fixed-effect model matrix is rank deficient so dropping 1 column /
coefficient

##microrex_test <- transform(microrex, PARTICIPANT=as.factor(PARTICIPANT))


##the warning comes from the overparametrization of the model,
##and is not important wrt. the residual variation.

##interindividual sd & residual sd:
summary(fit.lmer)$varcor

 Groups      Name        Std.Dev.
 PARTICIPANT (Intercept) 0.17685
 Residual                0.27201
```

We see that in this population the residual variation is larger than the inter-person variation.

In all 3 cases, the residual std.dev. is around 0.25 μm.

3 Calculation of the Repeatability Coefficient

We wish to estimate the expected difference between two independent measurements of the outcome on the same person in identical circumstances. To this end, we assume a residual deviance of $\sigma = 0.25$ μm, justified by the analyses in section 2.

For observations (Y_1, Y_2), with the same values of covariates, we have, under any of the above models, that $Y_1 - Y_2$ is an observation from a $N(0, 2\sigma^2)$ distribution. A 95%s prediction area is then given by:

$$\begin{aligned} 0.95 &= P(|Y_1 - Y_2| < m) \\ &= P((Y_1 - Y_2) \in (-m, m)), \end{aligned}$$

which implies, that $m \approx 2\sqrt{2\sigma^2} = 2 \times \sqrt{2} \times 0.25 \approx 0.707$.

This number is normally termed the repeatability coefficient. It is the upper limit for a prediction interval for the absolute difference between two measurements by the same method on the same item under identical circumstances, see [2] ch. 9.

3.1 Grouping of Observations

A clinically relevant estimate of the outcome is required to have a repeatability coefficient below 0.2μm.

One can apply the strategy of using the means of groups of repeated measurements as estimates. Since the standard deviation on the group means decreases with the square root of the size of the groups, a sufficiently big group size ensures a clinically relevant repeatability coefficient.

We will consider only groupings where all groups are of the same size.

Let $\bar{Y}_{1k}, \bar{Y}_{2k}$ be group means, each based on k independent observations, $Y_{11}, Y_{12}, \ldots, Y_{1k}$ and $Y_{21}, Y_{22}, \ldots, Y_{2k}$. $\bar{Y}_{1k} - \bar{Y}_{2k}$ follows a normal distribution with zero mean and variance given by:

$$\mathrm{Var}(\bar{Y}_{1k} - \bar{Y}_{2k}) = \mathrm{Var}(\frac{1}{k}\sum_{i=1}^{k}(Y_{1i} - Y_{2i})) = \frac{1}{k^2}\sum_{i=1}^{k}\mathrm{Var}(Y_{1i} - Y_{2i}) = \frac{1}{k}2\sigma^2,$$

where σ is the standard deviation for the single, ungrouped observations (the 0.25 above). The standard deviation for the distribution of $\bar{Y}_{1k} - \bar{Y}_{2k}$ is then $\sqrt{\frac{2}{k}}\sigma$.

We can now repeat the calculation above to find a repeatability coefficient, depending on the value of k:

$$m = 2\sqrt{\frac{2}{k}}\sigma.$$

```
##size of group, 1 to k
k <- 20
## for collecting results
A <- data.frame(1:k, NA)
colnames(A) <- c("Group size", "m")
##Group size and repeatability coefficient for the chosen size
A[,2] <-  2*(sqrt(2/(1:k))*0.25)
print(A, row.names=FALSE)
```

```
 Group size         m
          1 0.7071068
          2 0.5000000
          3 0.4082483
          4 0.3535534
          5 0.3162278
          6 0.2886751
          7 0.2672612
          8 0.2500000
          9 0.2357023
         10 0.2236068
         11 0.2132007
         12 0.2041241
         13 0.1961161
         14 0.1889822
         15 0.1825742
         16 0.1767767
         17 0.1714986
         18 0.1666667
         19 0.1622214
         20 0.1581139
```

Thus we see that to attain a repeatability coefficient below 0.2μm, we need groups of size at least 13 observations.

References

[1] Douglas Bates, Martin Mächler, Ben Bolker, and Steve Walker. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48, 2015.

[2] B. Carstensen. *Comparing Clinical Measurement Methods: A Practical Guide*. Statistics in Practice. Wiley, 2010.

[3] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.