Statistical analysis procedure and example R code for analysis of the relationship between the movement behavior composition and children's self-reported Health-Related Quality of Life (HRQoL)

Daily movement behaviors (sleep, sedentary time, light physical activity (LPA) and moderate-to-vigorous physical activity (MVPA)) collectively make up the entire 24-h day. Each behavior can be considered as a mutually exclusive and exhaustive part of the daily behavior composition. Taken together, these parts are perfectly multicollinear (as one part changes, one or more of the remaining parts must also change to maintain the constant sum of 24 hours) and one part is superfluous, as it can always be explained by the remaining parts (it will equal 24 h – the sum of remaining parts). Consequently, raw time spent in all daily behaviors cannot be used in multivariable linear models. However, a series of isometric log-ratio coordinates can be used to express the daily behavior composition. These log ratios can be constructed so that the first coordinate contains all information regarding the first movement behavior part, relative to all remaining parts. The R function pivotCoord() in the robCompositions package performs this transformation. When these isometric log-ratio coordinates are used in the linear model, the regression parameters returned for the first log-ratio coordinate can be understood as the influence of the first part, relative to the geometric mean of the remaining parts (i.e., as the first part increases, the remaining parts are decreased equally). This procedure is described in Chastin et al [1] and Dumuid et al [2]. Example R code for the procedure is below. Note, to assist with clarity, the example code has been simplified from the code used for the analyses in the main manuscript. To account for the grouping effect of schools, nested in countries, mixed effects models (with random intercepts) such as lmer() from the lme4 package can be used in place of standard linear models lm().

```
#load required packages
library(compositions)

library(robCompositions)

#import data
data1=read.csv("example HRQOL dataset.csv")

#hrqol outcomes
hrqol=data1$hrqol

#movement behaviors
sleep=data1$sleepindex
sed=data1$SEDindexE
lpa=data1$LPAindexE
mpa=data1$MPAindexE
vpa= data1$VPAindexE
mvpa=mpa+vpa #need to amalgamate mpa and vpa to make mvpa.
```

```r
# define movement behaviours
behaviors <- cbind(sleep,sed,lpa,mvpa)

#tell R it's a compositional variable
comp <- acomp(behaviors)

#create the isometric log ratio coordinates
ilr.comp <- as.matrix(pivotCoord(comp))
#because sleep is the first part of the composition, the first coord
inate will represent sleep relative to the geometric mean of the rem
aining behaviors.

#create linear model
model <- lm(hrqol ~ ilr.comp)

#look at the regression parameters for the first ilr coordinate
(summary(model))
```

```
##
## Call:
## lm(formula = hrqol ~ ilr.comp)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.672  -6.384  -1.242   4.875  34.814
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)               50.9359     0.6620  76.938  < 2e-16 ***
## ilr.compsleep_se-lp-mv     0.6847     0.9695   0.706  0.48008
## ilr.compsed_lp-mv          1.6924     0.6179   2.739  0.00618 **
## ilr.complpa_mv            -2.6612     0.5991  -4.442 9.09e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.598 on 5851 degrees of freedom
## Multiple R-squared:  0.003441,   Adjusted R-squared:  0.00293
## F-statistic: 6.734 on 3 and 5851 DF,  p-value: 0.0001565
```

```r
#sleep:remaining is not a significant predictor of hrqol

#to consider a different behaviour, the order of the behaviors in th
e  composition must be changed so that the behaviour of interest is
first. Alternatively, the pivotCoord() function can be used to speci
fy a different part as the pivotCoord.



#to specify the second part (sedentary time), the following command
can be used:
ilr.comp.sed <- as.matrix(pivotCoord(comp, pivotvar = 2))

#now, the linear model can be run using these coordinates.
model.sed <- lm(hrqol ~ ilr.comp.sed)
```

```r
#summary of model output
summary(model.sed)
```

```
##
## Call:
## lm(formula = hrqol ~ ilr.comp.sed)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.672  -6.384  -1.242   4.875  34.814
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 50.9359     0.6620  76.938  < 2e-16 ***
## ilr.comp.sedsed_sl-lp-mv    1.3674     0.7731   1.769    0.077 .
## ilr.comp.sedsleep_lp-mv     1.2097     0.8509   1.422    0.155
## ilr.comp.sedlpa_mv         -2.6612     0.5991  -4.442 9.09e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.598 on 5851 degrees of freedom
## Multiple R-squared:  0.003441,   Adjusted R-squared:  0.00293
## F-statistic: 6.734 on 3 and 5851 DF,  p-value: 0.0001565
```

```r
#sedenatry time relative to remaining behaviors is not significant i
n this example model
```

```r
#for LPA
ilr.comp.lpa <- as.matrix(pivotCoord(comp, pivotvar = 3))

#now, the linear model can be run using these coordinates.
model.lpa <- lm(hrqol ~ ilr.comp.lpa)

#summary of model output

summary(model.lpa)
```

```
##
## Call:
## lm(formula = hrqol ~ ilr.comp.lpa)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.672  -6.384  -1.242   4.875  34.814
##
## Coefficients:
##                            Estimate Std. Error t value Pr(>|t|)
## (Intercept)                 50.9359     0.6620  76.938  < 2e-16 ***
## ilr.comp.lpalpa_sl-se-mv   -3.1988     0.7515  -4.257 2.11e-05 ***
## ilr.comp.lpasleep_se-mv    -0.4047     0.9457  -0.428    0.669
## ilr.comp.lpased_mv          0.1351     0.4709   0.287    0.774
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.598 on 5851 degrees of freedom
## Multiple R-squared:  0.003441,   Adjusted R-squared:  0.00293
## F-statistic: 6.734 on 3 and 5851 DF,  p-value: 0.0001565


#LPA relative to remaining behaviours is significant


#CREATE ESTIMATES USING THE LINEAR MODELS

#estimate hrqol for mean movement behavior composition

#find the mean behavior composition
(m <-mean(comp))

##      sleep        sed        lpa       mvpa
## 0.37392292 0.36449617 0.22222365 0.03935726
## attr(,"class")
## [1] acomp

#note this determines the compositional mean (the geometric mean, cl
osed to sum to 1)

#make another composition, where sleep has been increased by 30 minu
tes.
#to compensate, the geometric mean of remaining behaviours must be d
ecreased
#to do this, let r be the increase in sleep (30 minutes, but must be
 expressed as a proportion).
r <- 30/1440/m[1]

#to compensate, the remaining behaviours are decreased equally by s
(to maintain a total of 1440).
s <- r* m[1]/(1-m[1])

#this is the new composition
new <- cbind(m[1]*(1+r),m[2]*(1-s),m[3]*(1-s),m[4]*(1-s))

# check what the new composition looks like in minutes
clo(new, total=1440)

##          [,1]     [,2]     [,3]     [,4]
## sleep 568.449 507.4088 309.3537 54.78856

#compared to the mean composition (in minutes)
clo(m, total=1440)

##     sleep       sed       lpa      mvpa
## 538.44900 524.87448 320.00205  56.67446

#can see that sleep has been increased by 30 min (a factor of 1+r),
whilst the remainingbehaviours have been decreased by the factor (1-
```

```
s)

df <- as.matrix(rbind(m,new))

#express as the two compositions as log-ratio coordinates
(ilr.set <- as.matrix(pivotCoord(df)))
```

```
##      sleep_se-lp-mv sed_lp-mv   lpa_mv
## [1,]      0.8075018  1.110708 1.224004
## [2,]      0.8837649  1.110708 1.224004
```

```
#1 is mean ilr, 2 is new ilr.

#estimate hrqol for the 2 ilr (the 2 compositions)
(predictions <- predict(model, newdata=list(ilr.comp=ilr.set)))
```

```
##        1        2
## 50.11122 50.16344
```

```
#this gives you predicted hrqol for (1) the mean composition and (2)
 the new composition.

#the change in hrqol is
predictions[2]-predictions[1]
```

```
##          2
## 0.05221755
```

References

1. Chastin, S. F., Palarea-Albaladejo, J., Dontje, M. L., & Skelton, D. A. (2015). Combined effects of time spent in physical activity, sedentary behaviors and sleep on obesity and cardio-metabolic health markers: A novel compositional data analysis approach. *PLoS ONE, 10*(10), e0139984.

2. Dumuid, D., Stanford, T., Martín-Fernández, J., Pedišić, Ž., Maher, C., Lewis, L., et al. (2017). Compositional data analysis for physical activity, sedentary time and sleep research. *Statistical Methods in Medical Research*, in press.