R code to re-create a simulation for the exchangeable structure for equal cluster sizes

```r
# Packages and source programs
library(geepack)
source('FUNCTIONS.R')

# Design
I <- 20        # number of clusters
n <- 30        # number of participants per cluster per period
T <- 5         # number of time periods
q <- I/(T-1)   # number of clusters randomized at each step
sig <- 0.05    # significance level
p0 <- 0.3      # baseline prevalence under control
beta <- c(log(p0/(1-p0)),-0.87,-0.89,-0.91,-0.93) # gently decreasing time-effect
theta <- 0.4592 # intervention effect on the log odds ratio scale
rho <- 0.05    # inter-cluster correlation
R <- (1-rho)*diag(n*T) + rho*matrix(1,nrow=n*T,ncol=n*T)  # exchangeable correlation matrix
trtSeq <- matrix(0,T-1,T)
trtSeq[upper.tri(trtSeq)] <- 1  # treatment sequences

# Simulate data
set.seed(20210513)
B <- create.B(I,T,q,beta,theta,trtSeq,n,R) # Create b's as in Qaqish equation 3
y <- create.response(I,T,q,beta,theta,trtSeq,n,B) # Simulate binary outcomes (Qaqish equation 3)

# Fit the model with the true correlation structure (working exchangeable)
cluster <- rep(1:I,each=T*n)
period <- rep(rep(1:T,each=n),I)
X.ij <- kronecker(trtSeq,rep(1,q))
treatment <- rep(c(t(X.ij)),each=n)
simdata_bin <- data.frame(cbind(y,cluster,period,treatment))
fit_exch <- geeglm(y~treatment+factor(period),id=cluster,data=simdata_bin,
              family=binomial(link = "logit"), scale.fix=TRUE, corstr="exchangeable")
summary(fit_exch)  # mean model, sandwich SEs and correlation
```

```
##
## Call:
## geeglm(formula = y ~ treatment + factor(period), family = binomial(link = "logit"),
##     data = simdata_bin, id = cluster, corstr = "exchangeable",
##     scale.fix = TRUE)
##
##  Coefficients:
##                 Estimate Std.err   Wald Pr(>|W|)
## (Intercept)      -0.8007  0.1326 36.440 1.57e-09 ***
## treatment         0.5933  0.1373 18.657 1.56e-05 ***
## factor(period)2  -0.2789  0.1301  4.596  0.03205 *
## factor(period)3  -0.3608  0.1292  7.804  0.00521 **
## factor(period)4  -0.3823  0.1625  5.532  0.01867 *
## factor(period)5  -0.5089  0.1552 10.747  0.00104 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = exchangeable
```

```
## Scale is fixed.
##
##   Link = identity
##
## Estimated Correlation Parameters:
##       Estimate Std.err
## alpha  0.04278 0.01509
## Number of clusters:   20  Maximum cluster size: 150
```

```r
sqrt(fit_exch$geese$vbeta.naiv[2,2]) # model-based SE for the treatment effect
```

```
## [1] 0.1387
```

```r
# Get Mancl DeRouen sandwich variance
alpha <- summary(fit_exch)$corr[1,1]
var <- (1-alpha)*diag(n*T) + alpha*matrix(1,nrow=n*T,ncol=n*T)
md_var <- md.var(fit_exch,simdata_bin,var)
sqrt(md_var) # Mancl DeRouen sandwich SE for the treatment effect
```

```
## [1] 0.1531
```

```r
# Fit the model with the working independence correlation structure
fit_indep <- geeglm(y~treatment+factor(period),id=cluster,data=simdata_bin,
                family=binomial(link = "logit"), scale.fix=TRUE, corstr="independence")
summary(fit_indep) # mean model and sandwich SEs
```

```
##
## Call:
## geeglm(formula = y ~ treatment + factor(period), family = binomial(link = "logit"),
##      data = simdata_bin, id = cluster, corstr = "independence",
##      scale.fix = TRUE)
##
##  Coefficients:
##                 Estimate Std.err  Wald Pr(>|W|)
## (Intercept)       -0.800   0.133 36.34  1.7e-09 ***
## treatment          0.470   0.204  5.33    0.021 *
## factor(period)2   -0.246   0.138  3.17    0.075 .
## factor(period)3   -0.293   0.153  3.69    0.055 .
## factor(period)4   -0.291   0.203  2.06    0.151
## factor(period)5   -0.386   0.223  3.01    0.083 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Correlation structure = independence
## Scale is fixed.
##
## Number of clusters:   20  Maximum cluster size: 150
```

```r
sqrt(fit_indep$geese$vbeta.naiv[2,2]) # model-based SE for the treatment effect
```

```
## [1] 0.113
```

```r
# Get Mancl DeRouen sandwich variance
var <- diag(n*T)
md_var <- md.var(fit_indep,simdata_bin,var)
sqrt(md_var) # Mancl DeRouen sandwich SE for the treatment effect
```

```
## [1] 0.23
```