Some information about the issue of comparing groups with a difference in severity. The following is from the Supplementary Information to: Van Borkulo et al., 2015. Association of symptom network structure with the course of depression. JAMA Psychiatry, 72(12):1219-1226. doi:10.1001/jamapsychiatry.2015.2079

```
# If you want to compare groups that differ in severity (e.g., healthy individuals and patients), it is possible that differences in network structures are due to a confound. Not because of difference in mean itself. If severity in itself were a confound with respect to network connectivity, then it should be the case that if a group has higher means on a set of variables, that group should also have a more connected network. However, the means of the variables do not play a role in the construction of the network; only the covariances do. In fact, one may standardize the variables without loss of generality: this will lead to exactly the same network, even though all variables in all groups would then have a mean of zero. Thus, mean level of the variables in itself cannot be a confound.
# Although mean level of the variables in itself cannot be a confound, it is possible that something associated with severity, and which does influence network connectivity, plays the role of confound. An important candidate in this respect is variance. If, due to a methodological artifact, the variance in the individual item scores is lowered in the less severe group, so that it is associated with the mean levels of the variables in the network, then that could lead to a lower network connectivity due to restriction of range. A plausible mechanism that could produce this situation is the existence of floor and/or ceiling effects. If the group with low connectivity shows symptom score distributions with floor and/or ceiling effects while the group with high connectivity does not, the floor and/or ceiling effects might be a confound with respect to network connectivity.

# Another possible mechanism that could lead to increased network connectivity in the more severe group is the presence of unmodelled latent variables. That is, if symptoms in the persister group were influenced more strongly by a latent variable (which would have to also be related to severity systematically), then the connectivity of persisters' network would be higher as a result. If, after controlling for such a latent variable, differences in connectivity disappear, the original difference was due to the latent variable. Conversely, if a difference in connectivity sustains, the latent variable cannot explain the difference.


# In this example code, we take whodas sum score as covariate (partialling out external measure of severity)
nvar <- ncol(x1) # the number of variables in your data
res_partout1 <- matrix(NA,nrow(x1),nvar)
res_partout2 <- matrix(NA,nrow(x2),nvar)
covariate1 <- xxx # a vector with WHODAS sum scores matching the patients in x1
covariate2 <- yyy # a vector with WHODAS sum scores matching the patients in x2

for(i in 1:nvar){
  lm1 <- lm(x1[,i] ~ covariate1)
  lm1.resid <- resid(lm1)
  lm2 <- lm(x2[,i] ~ covariate2)
  lm2.resid <- resid(lm2)
  res_partout1[,i] <- lm1.resid
  res_partout2[,i] <- lm2.resid
}
save(res_partout1,file='res_partout1.rdata') # now you can load the data next time with load('res_partout1.rdata')
save(res_partout2,file='res_partout2.rdata') # now you can load the data next time with load('res_partout2.rdata')
```