User's Guide for Conducting *A Priori* Power Analyses in R

Before we describe how to use our power programs, complete these preliminary steps:

1. Download and save *MIXREGLS.zip* from: https://www.jstatsoft.org/article/view/v052i12
2. Extract (aka, unzip) *MIXREGLS.zip* and save all extracted files to a known destination location. You will be required to use this destination location in our power programs (e.g., we extracted the files to: E:/MIXREGLS).
3. Locate the *Calculate_Values.xlsx* Excel file that will use to calculate the values to enter into the power programs.
4. Locate the *Detect.R* and *Predict.R* files.

****

Example Scenario

Say we are interested in quantifying and predicting individual differences in the intra-individual variability of positive affect in a sample of adolescents. The independent variable in this study will be the average hourly moderate-to-vigorous physical activity (MVPA) measured across the study period (a level-2, individual-level predictor). Said another way, we are interested in whether individual differences exist in how much adolescents vary around their own average amount of positive affect and whether these differences can be explained by their average amount of MVPA. Based on evidence observed in the literature, we expect the unconditional intra-class correlation (ICC) for positive affect to be 0.50, assuming that both the outcome-scale random intercept variance ($\sigma^2_{b_0}$) and the residual variance (σ^2_e) will equal 2. We expect scale-model random intercept variance ($\sigma^2_{t_0}$) to be 0.10 and expect the correlation between the location- and scale-model random intercepts to be 0.30. Finally, we expect the average amount of MVPA ($M_{\text{centered}} = 0$, $SD = 0.75$) to explain 10% of scale-model random intercept variance.

The Power to Detect Individual Differences in Intra-Individual Variability

Estimating the empirical power to detect individual differences in intra-individual variability is a two-step process: 1) calculate values using the *Calculate_Values.xlsx* Excel file and then 2) enter these values into and run the *Detect.R* power program. Open the *Detect* tab of the *Calculate_Values* Excel file. In this file, we can enter values in any yellow-shaded cell. Further, gray-shaded cells are calculated values for information purposes and are not used in the power program, whereas blue-shaded cells are values that will be used in the power programs.

Based on our example scenario, we enter *2* for both the location-model random intercept variance and residual variance in cells B13 and C13, respectively, enter 0.10 for the scale-model random intercept variance in cell D13, and enter 0.30 for the correlation between the location- and scale-model random intercepts. As shown in Figure 1, the unconditional ICC is calculated automatically for us using two different formulas that yield equivalent results—the ICC in cell E13 does not consider scale-model random intercept variance (cell E13), whereas the ICC in cell F13 does consider scale-model random intercept variance.

***A Priori* Variance Estimates**					
Location-Model Random Intercept Variance $\sigma^2_{b_0}$ (Outcome Scale)	Total Residual Variance σ^2_e (Outcome Scale)	Scale-Model Random Intercept Variance $\sigma^2_{t_0}$	Correlation of Location- and Scale-Model Random Intercepts ρ_{b_0,t_0}	Unconditional Intra-Class Correlation (ICC)	Unconditional Intra-Class Correlation (ICC) (Hedeker et al., 2008; Equation 10)
2.000	**2.000**	**0.100**	**0.300**	0.500	0.500

Figure 1. Entered values based on the example scenario

Based on the values provided in Figure 1, the Excel file automatically calculates the values to enter into our power program as shown in Figure 2.

Values to Enter into the DetectSMRI Power Program			
Location-Model Random Intercept Variance $\sigma^2_{b_0}$ (b0var)	Scale-Model Fixed Intercept τ_0 (tau0)	Scale-Model Random Intercept Variance $\sigma^2_{t_0}$ (t0var)	Correlation of Location- and Scale-Model Random Intercepts ρ_{b_0,t_0} (b0t0corr)
2.000	0.643	0.100	0.300

Figure 2. Values to enter into the *Detect* power program

These values include the location-model random intercept variance on the outcome scale (*b0var* in cell C24; $\sigma^2_{b_0}$), the scale-model fixed intercept on the log scale (*tau0* in cell D24; τ_0), the scale-model random intercept variance (*t0var* in cell E24; $\sigma^2_{t_0}$), and the correlation between the location- and scale-model random intercepts (*b0t0corr* in cell F24; ρ_{b_0,t_0}).

After calculating the necessary values in the Excel file, we turn our attention to the *Detect.R* power program used to detect scale-model random intercept variance. However, before we begin, it is critically important to note that MixRegLS software (mixregls.dll to be more specific) is compiled to a 32-bit target; thus, **you need to use a 32-bit version of R**. Further, we need to load R packages *Formula* to run MixRegLS as well as *LaplacesDemon* to sample from the Bernoulli distribution (this is completed in both R scripts using lines 9-11). In the *Detect.R* file, we only need to change the values associated with the objects defined in lines 13-35; we define values by changing quantities after the equal sign. Below, we provide a description of what effect each line specifies as well as the value we use based on our example scenario.

Line 19 – Enter the file location where we extracted (or unzipped) the MIXREGLS.zip file –

here, “E:/MIXREGLS/”. Note that the quotation marks as well as the forward slash after the file name are critically important.

Line 21 – Enter the number of individuals (*N*) – here, 50.

Line 23 – Enter the number of repeated occasions (n_i) – here, 15.

Line 25 – Enter the fixed intercept for the location-model (*beta0*; β_0) – here, 0. This is an outcome-scale value that has no effect on estimated power.

Line 27 – Enter the location-model random intercept variance on the outcome scale (*b0var*; $\sigma^2_{b_0}$) – here, 2.

Line 29 – Enter the fixed intercept for the scale-model for the residual variance (*tau0*; τ_0) – here, 0.643.

Line 31 – Enter the scale-model random intercept variance (*t0var*; $\sigma^2_{t_0}$) – here, 0.10.

Line 33 – Enter the correlation between the location- and scale-model random intercepts (*b0t0corr*; ρ_{b_0,t_0}) – here, 0.30. It should be noted that MixRegLS does *not* estimate this correlation directly, but can allow the location-model random intercept to influence the scale-model random intercept using fixed linear or quadratic effects (see Hedeker & Nordgren, 2013, for complete details). We calculate this correlation after the model has been estimated (see line 155).

Line 35 – Enter the number of replications we want to use when estimating power. These power programs can take a long time to run especially as the number of individuals and/or repeated measures increase, so although we use 1,000 replications for this example, we recommend starting with 10 replications to get a feel for the power. In general, we recommend at least 100 replications when determining the final power estimate.

```
13 ###################################################################
14 ###             Define MACRO Variable Values Below              ###
15 ###       Use Calculate_Values.xlsx to Determine Values         ###
16 ###################################################################
17
18 #Folder Directory for (unzipped) MixRegLS (note the / before the second ")
19   MIXREGLS = "E:/MIXREGLS/"
20 #Individuals
21   N = 50
22 #Repeated Occasions
23   ni = 15
24 #Location-Model Fixed Intercept (Outcome Scale)
25   beta0 = 0
26 #Location-Model Random Intercept Variance (Outcome Scale)
27   b0var = 2
28 #Scale-Model Fixed Intercept (Log Scale)
29   tau0 = 0.643
30 #Scale-Model Random Intercept Variance
31   t0var = 0.10
32 #Correlation between Location- and Scale-Model Random Intercept
33   b0t0corr = 0.30
34 #Number of Replications for Power Analysis
35   Nreps = 1000
```

Figure 3. Specified variable values in the *Detect* power program in R

After entering all values, lines 13-35 in the *Detect* R program should look like Figure 3 below. Highlight the entire R program (Ctrl + A) and click Run. In R, MixRegLS runs in the background, which is nice. After the program completes, results will be outputted to the console window presenting both the parameter recovery estimates as shown in Figure 4, as well as the empirical power estimate as shown in Figure 5.

```
> ParamRecov
                                                    Obs  Mean Lower95CI Upper95CI Pctl25 Pctl50 Pctl75
b0var: LM RI Variance (Outcome Scale) - Yes SMRI   1000 1.953     1.926     1.979   1.63  1.921   2.24
evar: Total Residual Variance (Outcome Scale) - Yes SMRI 1000 2.004 1.994   2.013   1.90  1.999   2.10
ICC: Unconditional Intra-Class Correlation - No SMRI  1000 0.488  0.485     0.492   0.45  0.491   0.53
ICC: Unconditional Intra-Class Correlation - Yes SMRI 1000 0.488  0.485     0.492   0.45  0.491   0.53
evar: Residual Variance (Log Scale) - Yes SMRI     1000 0.692     0.688     0.697   0.64  0.693   0.74
b0var: LM RI Variance (Log Scale) - Yes SMRI       1000 0.645     0.631     0.659   0.49  0.653   0.80
tau0: SM Fixed Intercept (Log Scale) - Yes SMRI    1000 0.645     0.640     0.649   0.60  0.644   0.69
t0var: SM RI Variance (Log Scale)                  1000 0.095     0.092     0.098   0.06  0.092   0.13
b0t0corr: LM-SM RI Correlation                     1000 0.331     0.314     0.348   0.17  0.336   0.49
```

Figure 4. Parameter Recovery Estimates from the *Detect* power program in R

```
> Power
  Power Lower95CI Upper95CI
1 0.785     0.758      0.81
```

Figure 5. Power Estimates from the *Detect* power program in R

On occasion, estimation problems might occur when the scale-model random intercept variance is sampled to be ~0 and/or when location-model random intercept variance or residual variance are large. In general, these types of errors tend to go by the wayside with large samples. The results in Figure 5 indicate that based on the model parameters we entered, we would achieve approximately 79% power to detect significant individual differences in intra-individual variability (i.e., scale-model random intercept variance) with 50 adolescents who each have 15 repeated occasions. The 95% Clopper-Pearson confidence interval for power was estimated to be 76% to 81%.

The Power to Predict Individual Differences in Intra-Individual Variability

Once we identify the amount or proportion of scale-model random intercept variance we want to detect, we turn our attention to prediction of this variance component. Similar to the procedure to detect scale-model random intercept variance, estimating the empirical power to predict individual differences in intra-individual variability (i.e., scale-model random intercept variance) is a two-step process. Based on our example scenario, we open the *Predict* tab of the *Calculate_Values* Excel file and enter *2* for both the location-model random intercept variance and total residual variance in cells B13 and C13, respectively, enter 0.10 for the scale-model random intercept variance in cell D13, and enter 0.30 for the correlation between the location- and scale-model random intercepts as previously shown in Figure 1. We then provide information about the MVPA predictor variable (called W_i) as shown in Figure 6.

Information about the Predictor Variable W_i					
		Continuous Predictor (Assumed Normally Distributed)		Binary Predictor	
Is Predictor W_i Continuous? (0=No; 1=Yes)	Proportion of Scale-Model Random Intercept Variance Explained (Pseudo-R^2)	Mean of Predictor W_i (Grand-Mean-Centered with a Mean of 0)	Variance of Predictor W_i	Mean of Predictor Wi	Variance of Predictor W_i
1	0.100	0	0.563	0.500	0.250

Figure 6. Information about scale-model predictor X_i

Specifically, we enter a *1* into cell B26 to indicate that predictor W_i is continuous and enter *0.10* into cell C26 for pseudo-R^2 because we expect this predictor to explain 10% of scale-model random intercept variance. Our power program assumes continuous predictors are grand-mean-centered such that they have a mean of 0. We enter *0.563* into cell D13 to indicate the predictor's variance (i.e., 0.75*0.75). For continuous predictors, we can ignore cells F26 and G26 because they apply only to a binary predictor. Had our predictor been binary, however, we would ignore cells D26 and E26 and instead enter the predictor's mean (i.e., the proportion of 1s) into cell F26 upon which its variance would be automatically calculated for us in cell G26. Based on the values provided in Figures 1 and 6, the Excel file automatically calculates the values to enter into our power program as shown in Figure 7.

Values to Enter into the PredictSMRI Power Program								
Location-Model Random Intercept Variance $\sigma^2_{b_0}$ (b0var)	Scale-Model Fixed Intercept τ_0 (tau0)	Scale-Model Random Intercept Variance $\sigma^2_{t_0}$ (t0var)	Correlation of Location- and Scale-Model Random Intercepts ρ_{b_0,t_0} (b0t0corr)	Unstandardized Scale-Model Fixed Effect for Predictor W_i τ_1 (tau1)	Pseudo-R^2 (R2)	Is Predictor Continuous? (Continuous)	Mean of Predictor W_i (Wmean)	Variance of Predictor W_i (Wvar)
2.000	0.643	0.100	0.300	0.133	0.100	1	0	0.563

Figure 7. Values to enter into the *Predict* power program

The values in Figure 7 include the location-model random intercept variance on the outcome scale (*b0var* in cell B37; $\sigma^2_{b_0}$), the scale-model fixed intercept (*tau0* in cell C37; τ_0), the amount scale-model random intercept variance (*t0var* in cell D37; $\sigma^2_{t_0}$), and the correlation between the location- and scale-model random intercepts (*b0t0corr* in cell E37; ρ_{b_0,t_0}). Also provided is the unstandardized fixed effect for the predictor variable (*tau1* in cell F37; τ_1) and population pseudo-R^2 (*R2* in cell G37). Finally, predictor information including an indicator of whether the predictor is continuous (*Continuous* in cell H37) as well as the predictor's mean and variance (*Wmean* in cell I37 and *Wvar* in cell J37) are also provided.

After calculating the values in the Excel file, we turn our attention to the *Predict.R* power program used to predict scale-model random intercept variance. In this R file, we only need to change the values associated with the macro variables located in lines 13-47. Below we provide a description of what each line requires as well as the value we use in our example scenario.

Line 19 – Enter the file location where we extracted (or unzipped) the MIXREGLS.zip file – here, "E:/MIXREGLS/". Note that the quotation marks as well as the forward slash after the file name are critically important.

Line 21 – Enter the number of individuals (*N*) – here, 50.

Line 23 – Enter the number of repeated occasions (n_i) – here, 20.

Line 25 – Enter the fixed intercept for the location-model (*beta0*; β_0) – here, 0. This is an outcome-scale value that has no effect on estimated power.

Line 27 – Enter the location-model random intercept variance on the outcome scale (*b0var*; $\sigma^2_{b_0}$) – here, 2.

Line 29 – Enter the fixed intercept for the scale-model for the residual variance (*tau0*; τ_0) – here, 0.643.

Line 31 – Enter the scale-model random intercept variance (*t0var*; $\sigma_{t_0}^2$) – here, 0.10.

Line 33 – Enter the correlation between the location- and scale-model random intercepts (*b0t0corr*; ρ_{b_0,t_0}) – here, 0.30. It should be noted that MixRegLS does *not* estimate this correlation directly, but can allow the location-model random intercept to influence the scale-model random intercept using fixed linear or quadratic effects (see Hedeker & Nordgren, 2013, for complete details). We calculate this correlation after the model has been estimated (see lines 190 and 216).

Line 35 – Enter the unstandardized fixed effect of the scale-model predictor X_i (*tau1*; τ_1) – here, 0.133.

Line 37 – Enter the population pseudo-R^2 value (*R2*) – here, 0.10.

Line 39 – Enter either 0 or 1 to indicate whether the scale-model predictor X_i is continuous or binary (0 = binary, 1 = continuous) – here, 1.

Line 41 – Enter the mean of the scale-model predictor W_i (*Wmean*). Note that continuous predictors are assumed to be grand-mean-centered with a mean of 0 – here, 0.

Line 45 – Enter the variance of the scale-model predictor X_i (*Xvar*). Note that, although the variance of a binary predictor is defined by its mean and this value is not used when sampling from the Bernoulli distribution, some non-null value must be entered after the equal sign on lines 36 for the R program to execute.

Line 47 – Enter the number of replications we want to use when estimating power. These power programs can take a long time to run especially as the number of individuals and/or repeated measures increase, so although we use 1,000 replications for this example, we recommend starting with 10 replications to get a feel for the power. In general, we recommend at least 100 replications when determining the final power estimate.

```
13 ###################################################################
14 ###              Define MACRO Variable Values Below             ###
15 ###        Use Calculate_Values.xlsx to Determine Values        ###
16 ###################################################################
17
18 #Folder Directory for (unzipped) MixRegLS (don't forget the second forward slash!)
19   MIXREGLS = "E:/MIXREGLS/"
20 #Individuals
21   N = 50
22 #Repeated Occasions
23   ni = 20
24 #Location-Model Fixed Intercept (Outcome Scale)
25   beta0 = 0
26 #Location-Model Random Intercept Variance (Outcome Scale)
27   b0var = 2
28 #Scale-Model Fixed Intercept (Log Scale)
29   tau0 = 0.643
30 #Scale-Model Random Intercept Variance
31   t0var = 0.10
32 #Correlation between Location- and Scale-Model Random Intercept
33   b0t0corr = 0.30
34 #Scale-Model Fixed Predictor Effect (Outcome Scale)
35   tau1 = 0.133
36 #Pseudo-R2
37   R2 = 0.10
38 #Is the Predictor Continuous? (1=Yes, 0=No)
39   Continuous = 1
40     #Mean of Predictor W (=0 if W is continuous)
41       Wmean = 0
42     #Variance of Continuous Predictor W
43     #Does not apply to binary predictor, but must have
44     #non-null value after equal sign for program to run
45       Wvar = 0.563
46 #Number of Replications for Power Analysis
47   Nreps = 1000
```

Figure 8. Specified macro variables in the *Predict* power program in R

After entering all values, lines 13-47 in this R program should look like Figure 8 below. Highlight the entire R program (Ctrl + A) and click Run. In R, MixRegLS runs in the background, which is nice. After the program completes, results will be outputted to the console window presenting both the parameter recovery estimates as shown in Figure 9, as well as the empirical power estimate as shown in Figure 10.

```
> ParamRecov
                                                            Obs   Mean Lower95CI Upper95CI  Pctl25   Pctl50 Pctl75
beta0: LM Fixed Intercept (Outcome Scale)                  1000 0.0023   -0.0106    0.0152 -0.1421 -0.00194  0.148
b0var: LM RI Variance (Outcome Scale) - Yes Predictor      1000 1.9699    1.9452    1.9945  1.6955  1.94486  2.217
tau0: SM Fixed Intercept (Log Scale) - Yes Predictor       1000 0.6428    0.6384    0.6472  0.5944  0.64334  0.692
tau1: SM Fixed Effect of Predictor W (Log Scale)           1000 0.1389    0.1333    0.1445  0.0814  0.13825  0.201
t0var: SM RI Variance - No Predictor                       1000 0.0951    0.0922    0.0980  0.0615  0.08971  0.124
t0var: SM RI Variance - Yes Predictor                      1000 0.0799    0.0771    0.0826  0.0475  0.07365  0.106
PseudoR2: Proportion Reduction SM RI Variance from Predictor W 1000 0.1823 0.1704   0.1941  0.0424  0.11852  0.263
Correlation b/w LM and SM RI - Yes Predictor               1000 0.3524    0.3340    0.3708  0.1626  0.35152  0.536
```

Figure 9. Parameter Recovery Estimates from the *Predict* power program in R

```
> Power
  Power Lower95CI Upper95CI
1  0.35     0.286     0.345
```

Figure 10. Power Estimates from the *Predict* power program in R

The results in Figure 10 indicate that based on the model parameters we entered, we would achieve only 35% power to predict 10% of the individual differences in intra-individual variability (i.e., 10% of the scale-model random intercept variance) with 50 adolescents who each have 15 repeated occasions. The 95% Clopper-Pearson CI confidence interval was estimated to be 29% to 35%.

When considering the empirical power estimates from the *Detect* and *Predict* programs, we can see that although we had sufficient power to detect individual differences in intra-individual variability with 50 adolescents and 20 repeated occasions, power to predict 10% of this variability was woefully inadequate. As such, we would subsequently re-run the *Predict* program with increased numbers of adolescents and/or repeated occasions to ensure adequate power to predict individual differences in intra-individual variability.