Covariates.csv

... contains the full set of variables we will be controlling for or investigating in the models. Note that missing values for any of the covariates Dx, Age, Sex, and when applicable Site is not allowed. Missings in the other covariates are allowed, but make sure you code them with "NA" without the quotes. You will need to make this file.

- In Excel, or your favorite text editor, create a series of columns with the following headers: SubjID, Dx, Age, Sex, Site1, Site2, ...
  - The IDs in the **SubjID** column should match the format of the IDs in the SubjID column in the hippo_subfields.csv. If the latter file contains the full path, reduce it to just the SubjID.
  - **Dx** should be a column for diagnosis (patients coded as 1, controls as 0).
  - **Age** is in years at time of scan.
  - **Sex** (Males=1, Females=2).
  - The **Site#** columns are optional and only required if your sample requires correction for data collection at multiple sites. If you have no site variables to include, do not include any Site# columns. If you do have sites to include as covariates in the model, we will be correcting using dummy variables where the total number of different Site# columns is equal to (*n*-1) such that *n* is the total number of sites coded as dummy variables. For example, if data were collected at four different data centers then we would include 3 columns of dummy variables to indicate site (Site1, Site2, Site3).
  - **Recur** refers to whether the MDD patient has a recurrent or first depressive episode. Code: recurrent patients=2, first episode patients=1, healthy controls=0.
  - **AD** refers to current antidepressant use. Code: antidepressant users=2, antidepressant free MDD patients=1, healthy controls=0.
  - **Rem** refers to whether the MDD patient is acutely depressed (6 month recency) or in remission. Code: acutely depressed=2, remitted patients=1, healthy controls=0.
  - **AO** is age of onset in years, i.e. age at first depressive episode. For controls or if not applicable, enter NA.
  - **Sev** is severity in number of DSM-IV MDD criteria met (on basis of DSM-IV interview ranging from 0-9). For controls or if not applicable, enter NA.
  - **BDI** is total BDI-II score. For controls or if not applicable, enter NA.
  - **HDRS** is total HDRS-17 (HAMD-17) score; for controls or if not applicable, enter NA.
  - **Epi** is total number of depressive episodes (current and previous episodes), only needs to be filled out for the MDD patients; for controls or if not applicable, enter NA.
  - **ADcur** is duration of antidepressants that are currently used (in weeks).

If variables are not applicable, i.e. you don't have this info in your study (e.g. remission or BDI), **include the variables as a column in your Covariates.csv file** but enter NA in all cells. Beware that the columns in your csv files are separated by commas (,) not by semicolons (;). In case, check by opening csv files with a texteditor and replace all ";" by "," and save file. Here is an example how the beginning of the file should look:

```
SubjID,Dx,Age,Sex,Site1,Recur,AD,Rem,AO,Sev,BDI,HDRS,Epi,ADcur
WG_001,0,24,2,1,0,0,0,NA,NA,NA,NA,NA,NA
WG_002,0,38.4,2,1,0,0,0,NA,NA,NA,NA,NA,NA
[...]
```

hippo_subfields.csv

The head of this file should look as follows - note commas as separators and no quotation markers. NA values are allowed, yet will lead to exclusion of that subject from the analysis.

```
SubjID,L_Hippocampal_tail,L_subiculum,L_CA1,L_hippocampal-
fissure,L_presubiculum,L_parasubiculum,L_molecular_layer_HP,L_GC-ML-
DG,L_CA3,L_CA4,L_fimbria,L_HATA,L_Whole_hippocampus,R_Hippocampal_tail,R_subiculum,R_CA1,R_hippocampal-
fissure,R_presubiculum,R_parasubiculum,R_molecular_layer_HP,R_GC-ML-
DG,R_CA3,R_CA4,R_fimbria,R_HATA,R_Whole_hippocampus,Lhippo,Rhippo,eTIV,Brain,TotalGM
WG_001,547.465,419.068,583.78,184.675,337.699,55.9846,552.166,272.02,180.404,229.927,71.6749,59.8195,3310.0
1,505.036,467.4,631.841,189.478,337.034,52.1946,573.191,275.76,170.063,232.029,91.4007,67.0488,3403,4266.5,
4551.6,1.49457e+06,1.08066e+06,609850
WG_002,490.829,426.746,593.825,159.854,335.216,54.7237,541.534,268.918,164.022,219.31,72.7212,55.7984,3223.
64,536.352,399.376,603.231,203.981,299.555,65.0042,507.15,245.181,157.88,201.78,86.4663,57.5482,3159.52,405
3.1,4137.5,1.32251e+06,948700,533067
[...]
```