**Concussion biotyping Study – Master Script**

1. **Getting the data**
   1. **Extracting mTBI cases**
      1. **Go to ABCD Data Dictionary website (**[**https://nda.nih.gov/data\_dictionary.html?source=ABCD%2BRelease%2B1.1&submission=ALL**](https://nda.nih.gov/data_dictionary.html?source=ABCD%2BRelease%2B1.1&submission=ALL)**)**
      2. **Log in with your NDA credentials**
      3. **Select ABCD Release 2.0 as Data Source**
      4. **Select ABCD Sum Scores Traumatic Brain Injury (abcd\_tbi01) and press “Add to Filter Cart”**
      5. **On the top right, a small window named “Filter Cart” will display “1 filter”. Click on “Package/Add to Study”.**
      6. **That will take you to a page where you will see the data you filtered. Press on Create Package, and give your package a name.**
      7. **Then go to your profile (click on your username) and go to the Packages section**
      8. **If you don’t have the Download Manager already, click on the option to download it.**
      9. **Open the Download Manager, enter your NDA credentials and download the file (should be called abcd\_tbi01.txt).**
      10. **Run the R script “ABCD\_isolate\_mTBIs.rmd”**
          1. **This script uses the tbi\_ss\_worst\_overall variable, which lists the worst injury a participant has had**
          2. **It selects tbi\_ss\_worst\_overall=2 (Possible mTBI: TBI without LOC) and tbi\_ss\_worst\_overall=3 (mTBI: TBI with LOC ≤30mins)**
          3. **This script also returns Gender, Age at first TBI…**
          4. **This script also excludes participants with neurological and/or psychiatric conditions** 
             1. **Note: the following variables are needed to run this part: FILL IN**
             2. **Autism**
             3. **Epilepsy**
             4. **Lead poisoning**
             5. **Multiple Sclerosis**
             6. **Cerebral Palsy**
             7. **Depression**
             8. **Bipolar disorder**
             9. **Psychosis**
             10. **Panic disorder**
             11. **Agoraphobia**
             12. **ADHD**
          5. **The resulting file will be called “mTBI\_definitions\_and\_related\_variables.csv”**
          6. **Copy the ‘subjectkey’ for all participants into a text file (note: there should be 434 subjects)**
   2. **Extracting mTBI DWI scans**
      1. **Go to the ‘Query by GUIDs’ part of the ‘Get Data’ section (**[**https://www.computecanada.ca/research-portal/national-services/globus-portal/**](https://www.computecanada.ca/research-portal/national-services/globus-portal/)**)**
      2. **Click on Upload File and select your text file with the subjectkeys, then ‘Apply Filters’**
      3. **Once the filter is applied, click on ‘Add to Workspace’, and then on an small icon on the right that will open a window where you can click on ‘Submit to Filter Cart’.**
      4. **Once the ‘Filter Cart’ window has finished loading, click on ‘Package/Add to Study’.**
      5. **In the list called ‘Data Structure by Category’, deselect all, and only select the variable ‘DTI, MRI, fMRI: Image’**
      6. **Click on Create Package (DO NOT ‘Include Associated Data Files’)**
      7. **Open Download Manager and download the text sheet which will contain all the information related to DWI, MRI, and fMRI images**
      8. **Run the “Extract\_DWIandT1Scans\_s3links.R”, which isolates DWI scans based on the ‘image\_description’ variable:**
         1. **ABCD-Diffusion-FM-PA**
         2. **ABCD-Diffusion-FM**
         3. **ABCD-DTI**
         4. **ABCD-T1**
      9. **This script writes a txt file with the S3 links for these files**
      10. **This script also writes a txt file with the GUIDs of all subjects (n=406) with one of the 4 image-types described above (referred to as ‘List#2’).**
   3. **Downloading mTBI DWI scans**
      1. **Set up temporary access credentials**
         1. **Open Download Manager**
         2. **Click on Tools -> Generate AWS Credentials**
         3. **On a command window, type vi ~/.NDATools/settings.cfg**
         4. **Enter your NDA username, password, and the Access Key and Secret Key generated in step 2**
         5. **Exit vi (press ‘:x’)**
         6. **Note: these credentials last 24 hours**
      2. **In a command window, write ‘pip install nda-tools’ (documentation available in NDAR Github:** [**https://github.com/NDAR**](https://github.com/NDAR)**)**
      3. **If you are using Python2:**
         1. **Open the script Download.py (in anaconda/lib/python2.X/site-packages/NDATools.**
         2. **Add on top of the script ‘from \_\_future\_\_ import print\_function’**
         3. **In your home directory, create a directory called ‘NDAValidationResults’**
      4. **Run the downloadcmd script**
         1. **downloadcmd /*path/of/txtfile/with/S3links* –t –d *outdir***
         2. **This script organizes data by submission**
      5. **Organize the data by participant by running the organize\_ABCD\_data.sh script:**
         1. **Note: for this script and the next to work, the subject IDs need to have the ‘NDAR\_’ removed. It can be done in Excel**
         2. **organize\_ABCD\_data.sh */path/to/rootdir/with/datafolders*** ***/path/to/txtfile/with/LIST#2* */path/to/txtfile/with/submission\_numbers***
      6. **Look for data inconsistencies by running the verify\_ABCD\_data\_inconsistencies.sh script:**
         1. **flag\_ABCD\_data\_inconsistencies.sh */path/to/rootdir/with/datafolders /path/to/txtfile/with/subjectIDs***
         2. **This script will rename directories based on the data inconsistency problem in a specific order**