Updated by Dylan 180505

# The Protocol has 4 parts:

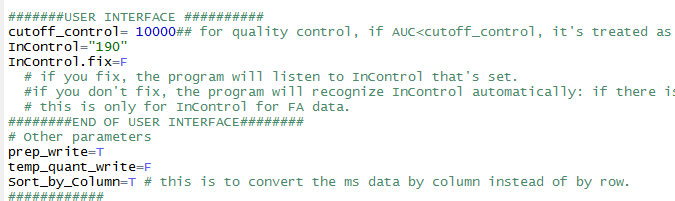
Part I: [R] formatting: from MassHunter AUC file to prepfile

Part II: [MATLAB] modeling: from “Prep…xls” to “Result…xls” file

Part III: [R] std curve and graph: from “Result..xls” , “MAP..xlsx”, “….quant.Rda” to “Summary…xlsx” and figures!

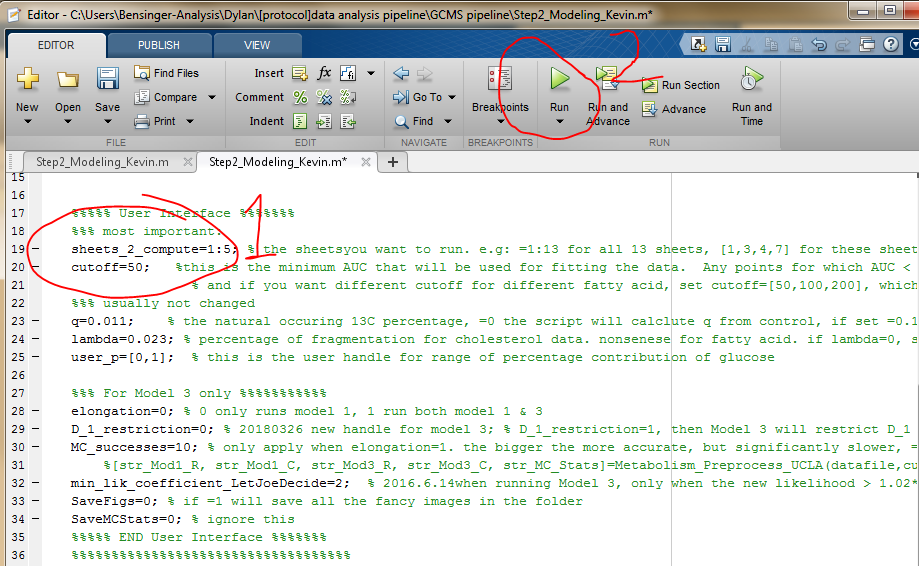
Part IV: [You] interpreting the data

**Part I: [R] formatting: from MassHunter AUC file to prepfile**

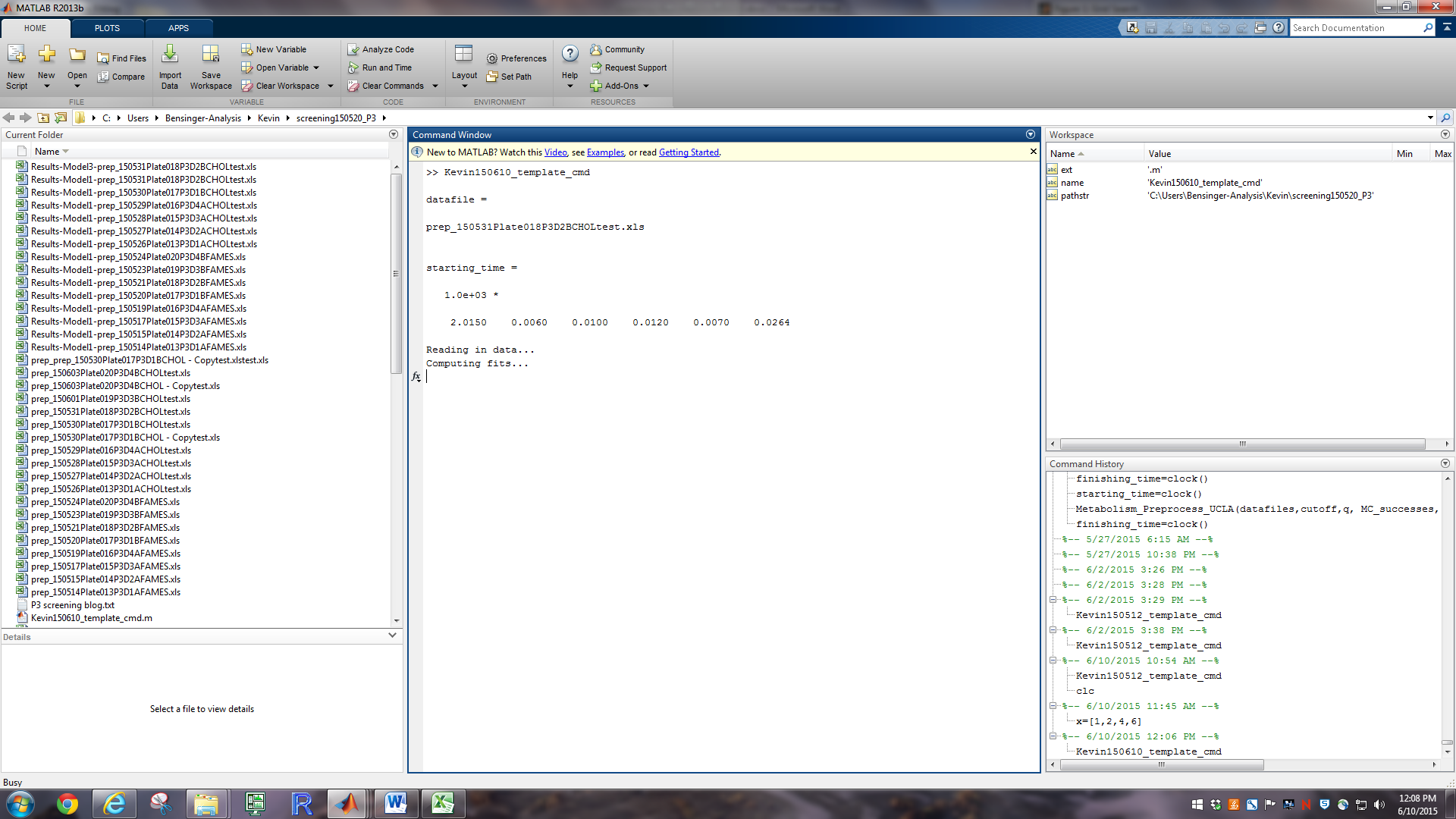
* + Have MassHunter AUC file ready, example: **“180424AdipOrianExp1\_FAMES.xlsx”**
    - Make sure inaccurate AUCs are deleted from this xlsx (because the Rda file from here will be used for downstream analysis)
    - File name must end with “**FAMES**” or “**CHOL**”
    - If you are not using first column for blank and standard, see the other protocol
  + Go to Desktop, find folder “GCMS pipeline - Shortcut”, left double click 
    - User Interface Looks like this:
    - InControl is the name for Fatty acid Internal Control. If InControl.fix=F, then the program detects Internal automatically (190 by default, if there is no 190 then 170)
  + Click “Source” button to run script
    - Follow instructions, select raw AUC data files for anaylasis. converted files will be in the same folder.
  + Spot Check **“prep\_180424AdipOrianExp1\_FAMES.xls”** in your folder
    - Check Code column: “Code” is a handle for modeling, -1 means don’t model, 1 means run the model, 0 means no label samples will be used as natural 13C control.
    - Check Internal Control tab. Remember that there was a “cutoff\_control” parameter in the script. Any sample with Internal Control < cutoff\_control will be considered as bad injection and will be trashed in quantification analysis. If lots of your samples have low internal control signal,1) delete your prep file, 2) in the script, set the cutoff\_control to lower value, 3) rerun
    - Check Bad Ions. E.g. sometimes there are giant contamination peaks in 140 and 161, if this happens 1) delete your prep file 2)back up your Masshunter file, delete the bad ions 3) rerun
  + Close Rstudio, don’t save.

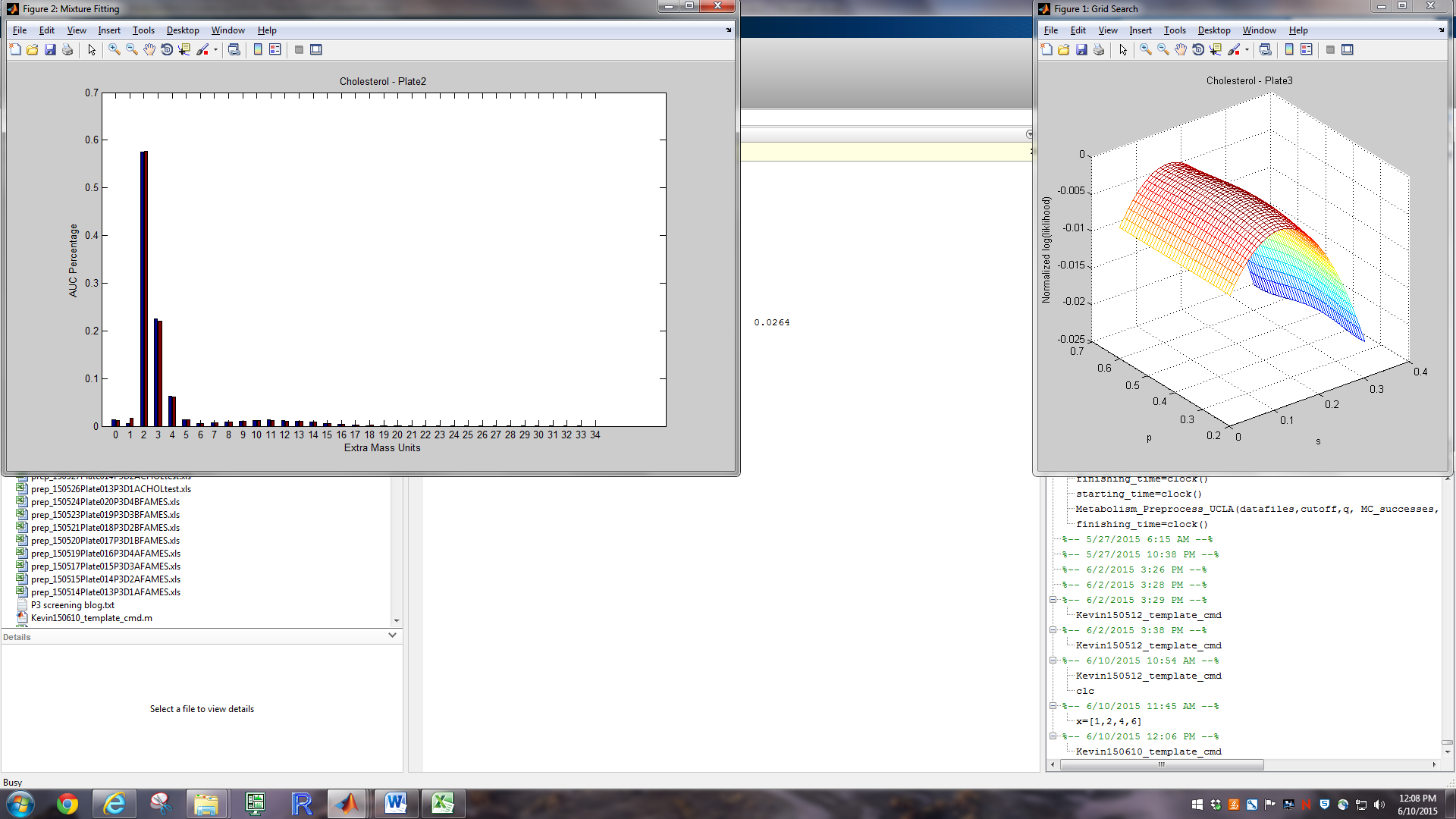
**Part II: [MATLAB] modeling: from “Prep…xls” to “Result…xls” file**

* + Copy the most updated .m file  to your folder for analysis, and rename it.
  + Double-click open this .m file e.g. (you will have to wait for 2 windows in MATLAB to pop-up, 1st is workspace, 2nd is your .m file)



* + - **sheets\_2\_compute**: sheets you want to process in “prep..xls” file :
      * sheets\_2\_compute=1 🡪 only compute 1st sheet
      * sheets\_2\_compute=1:5 🡪 compute sheet #1~5
      * sheets\_2\_compute=[1,3,4,6] 🡪 compute sheet # 1,3,4,6
    - **cutoff**: any value below will be trashed in the modeling, set it lower if your signal is low
    - change **q and lambda** value(0 means let MATLAB calculate from your non-labeled sample, or you can force it to use designated value. If there is no non-label nor designated value, model will use default q=0.011, lambda=0.023),
  + hit “Run”
  + go to the other window in MATLAB: should be reading and computing data now….
  + If you want to terminate modeling, press “Ctrl+C”



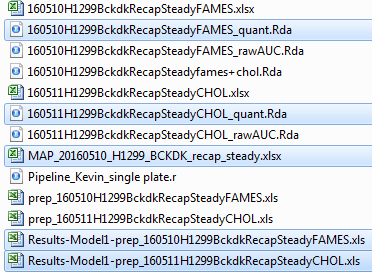
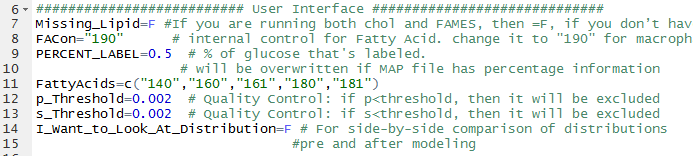


Time:

Model 1: 140-160-161-180-181 for 88 samples take 1 hour, chol for 88 samples take 40min

Model 3: 160-180-181 for 88 samples take 12 hours

**Part III: [R] std curve and graph: from “Result..xls” , “MAP..xlsx”, “….quant.Rda” to “Summary…xlsx” and figures**

* Go to Desktop, find folder “GCMS pipeline - Shortcut”, copy this to your folder 
* Before you run the pipeline script, your data folder should look like these, the highlighted files are required for the pipeline  
  
* **Left double click** it look like this:  
  
  + follow the green notes in the script to set parameters
* **Run the scrip by click Source button on top right as highlighted**
* There will be sweet instructions in the Console. Follow instructions starting with “**please**”. Others are just progress notes. (see Sup. A)

**Part IV: [You] interpreting the data**

**Quality Control For modeling:**

* Check if p and s is real.
* If p isn’t real, then s isn’t real. P.real>1 is definitely not real. If still within range, than you should check the original data with the modeled distribution

**===================================Supplement Info=====================================**

**Sup. A**: A typical processing notes look like this:

> source('C:/Users/Bensinger-Analysis/Dylan/[protocol]data analysis pipeline/GCMS pipeline/Data Template-single plate-TEST/Pipeline\_Kevin\_single plate.r')

Loading required package: rJava

Loading required package: xlsxjars

[1] "please select all files: Result\_Model1\_.....xls"

[1] "Merging results for file: 160510H1299BckdkRecapSteadyFAMES"

[1] "end of Merging file: 160510H1299BckdkRecapSteadyFAMES"

[1] "Merging results for file: 160511H1299BckdkRecapSteadyCHOL"

[1] "end of Merging file: 160511H1299BckdkRecapSteadyCHOL"

[1] "Now the .Rda files are updated with p and s info:)"

[1] "I'm going to combine FAMES\_quant.Rda with chol\_quant.Rda file"

[1] "Please select a FAMES\_quant.Rda file"

[1] "160510H1299BckdkRecapSteadyFAMES\_quant.Rda"

[1] "Please select a CHOL\_quant.Rda file"

[1] "160511H1299BckdkRecapSteadyCHOL\_quant.Rda"

[1] "Output: \_fames+chol.Rda"

[1] "please select your Map file. the map file starts with [ Map\_xxxx.xlsx ]"

[1] "MAP\_20160510\_H1299\_BCKDK\_recap\_steady.xlsx"

[1] "160510H1299BckdkRecapSteadyfames+chol.Rda"

[1] "Summary\_20160510\_H1299\_BCKDK\_recap\_steady.xlsx"

[1] "p --> p.real"

[1] "Cellcounts ready"

[1] "writing cell\_count"

[1] "writing quant"

[1] "writing quant\_norm"

[1] "writing s"

[1] "writing p.real"

[1] "writing Abs\_total"

[1] "Summary file is in file [ Summary\_20160510\_H1299\_BCKDK\_recap\_steady.xlsx ]"

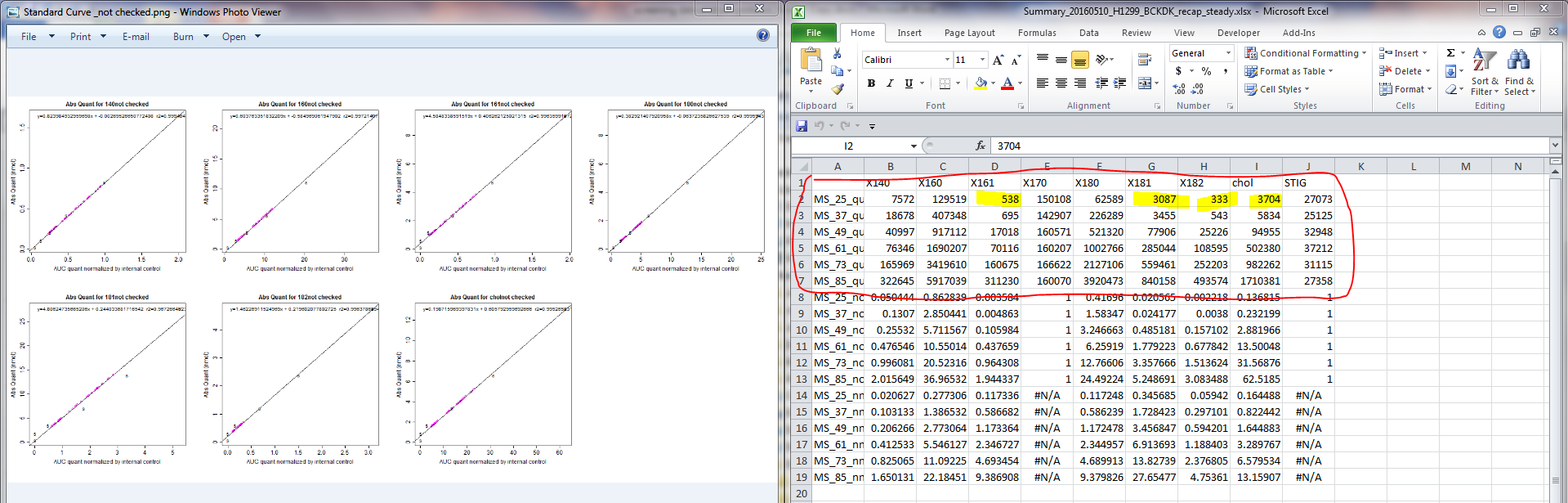
[1] "please check \" Standard Curve.png\" and delete bad data in \"Summary\_20160510\_H1299\_BCKDK\_recap\_steady.xlsx\""

After manually checking the data, Press [enter] to continue

**Now go back to your folder and open these two files:**

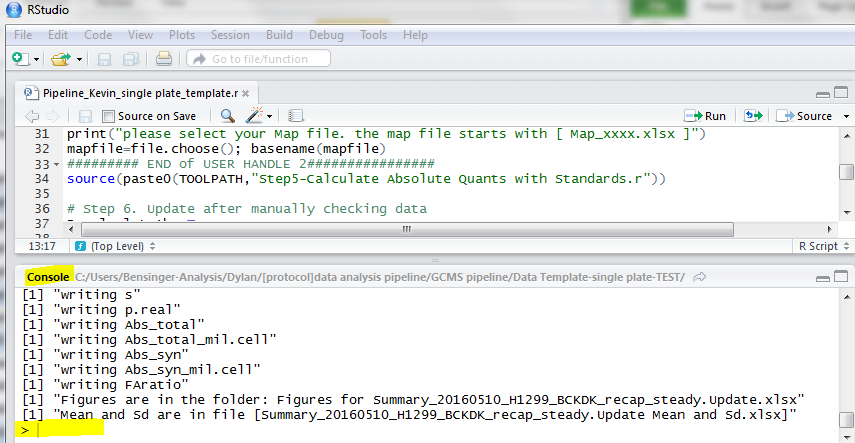
****

**and delete bad standards in the “standard” sheet. You only have to delete data in the first 6 rows.**

****

**Then you go back to Rstudio and press [Enter].**

**(note: your cursor has to be in the console section as highlighted)**

****

[1] "please select your .xlsx file with (standard,quant,s, p.real) to be updated"

[1] "updated: Abs\_total\_mil.cell "

[1] "updated: Abs\_syn\_mil.cell "

[1] "updated: Abs\_syn "

[1] "updated: Abs\_scav\_mil.cell "

[1] "updated: Fatty acid ratios from quant\_norm data"

[1] "writing standard"

[1] "writing cell\_count"

[1] "writing quant"

[1] "writing s"

[1] "writing p.real"

[1] "writing Abs\_total"

[1] "writing Abs\_total\_mil.cell"

[1] "writing Abs\_syn"

[1] "writing Abs\_syn\_mil.cell"

[1] "writing FAratio"

**Now, all the parameters are generated. You can go to your folder to check your updated file and standard curve if you want: **

**Now, let’s calculate the mean and std and plot them out!**

**But before doing so, you can go to your [Map\_xxx.xlsx] file, and delete unwanted samples in column “Exclude”, just mark them as “1”, save the xlsx, and press [Enter] in Rstudio console**

[1] "Please update your mapfile to exclude unwanted samples!"

After mark samples for exclusion in column [Exclude], save Map.xlsx,Press [enter] to continue

[1] "writing cell\_count"

[1] "writing quant"

[1] "writing s"

[1] "writing p.real"

[1] "writing Abs\_total"

[1] "writing Abs\_total\_mil.cell"

[1] "writing Abs\_syn"

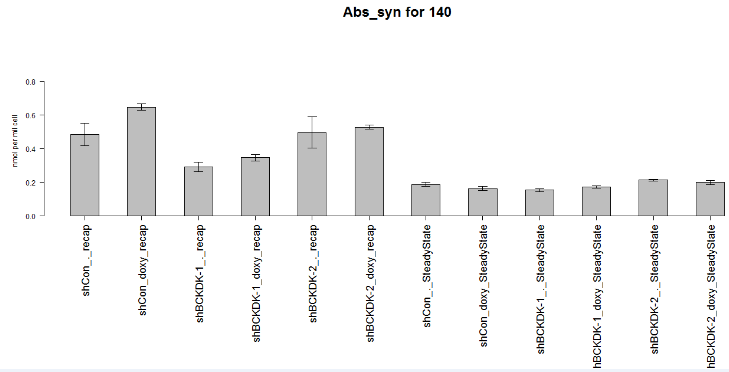
[1] "writing Abs\_syn\_mil.cell"

[1] "writing FAratio"

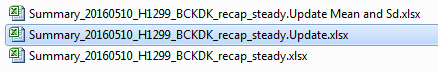
[1] "Figures are in the folder: Figures for Summary\_20160510\_H1299\_BCKDK\_recap\_steady.Update.xlsx"

[1] "Mean and Sd are in file [Summary\_20160510\_H1299\_BCKDK\_recap\_steady.Update Mean and Sd.xlsx]"

**You’ll see a bunch of figures by now:**



**(OPTIONAL): you might want to delete several data after you see the figure.**

* **Delete sample: you can exclude samples in [Map.xlsx] again and save**
* **Delete individual parameters for an sample:** 
  + **delete them in [Update.xlsx], and**
  + **open****and**
  + **hit “source”. You are set!**