

Example R Script for Analysis of Longitudinal, Single-Well Luminex® Data Sets

```

library(lattice)
library(nlme)
Directory <- "C:\\MyDirectory"
DF <- read.csv(file = paste0(Directory, "\\Project Data.csv"), header = T)
for (i in 5:(dim(DF)[2])) {
  TR <- data.frame(DF[,1:4], MFI = DF[, i])
  png(filename = paste0(Directory, "\\ ", names(DF)[i], " Log MFI.png"), width = 4, height = 4, units = "in", bg = "white",
res = 1100)
  print(xyplot(log(MFI) ~ Time, TR, type = "o",
              groups = Participant, lty = 1, pch = 1, cex = 0.2,
              col.line = "darkgrey", col.symbol = "black",
              xlab = "Time",
              ylab = "Logarithm MFI"))
  dev.off()
  Prot.lme <- lme(I(log(MFI)) ~ Time + I(Time ^ 2) + factor(Plate) + I(log(CHEX4)), data = TR,
                random = ~ 1 | Participant)
  cat("      ", "\n")
  cat("      ", "\n")
  cat(paste("Protein = ", names(DF)[i], "\n"))
  print(summary(Prot.lme))
}

```

First 2 lines load R packages.

Third line gives directory location for input data and output figure. Update this to your local directory name.

Fourth line reads in CSV file containing data set. (You can supply a different input file name in place of `Project Data.csv`.) Data set should be structured as follows.

Plate	Participant	Time	CHEX4	Prot1	Prot2	Prot3	Prot4
1	1	0.926698	3674	71	389	577	642
1	1	1.771945	297	375	74	303	765
1	1	3.205812	14061	976	507	595	3218
1	1	4.207647	2576	1231	233	756	2618
1	1	4.769608	5981	1970	113	895	1372
1	1	5.984521	640	3172	338	1357	2242
1	1	7.053555	550	2470	377	966	20803
1	1	7.867472	2088	6138	277	2303	9359
1	1	9.120917	4279	5451	125	1666	34355
1	1	9.913232	1632	10168	71	611	30173
1	2	0.873971	978	549	35	677	371
1	2	1.865528	1676	822	185	329	2210
1	2	3.04307	1756	1323	382	139	1134
1	2	3.952711	91	342	798	367	814
1	2	4.934261	925	1082	458	301	2448
1	2	5.806275	201235	3670	212	1385	14459

1	2	7.127764	73	3479	189	1347	5305
...
4	30	0.90224	579	451	292	302	664
4	30	2.09797	17770	965	493	293	1670
4	30	2.89757	2030	2929	479	228	1034
4	30	4.109301	2604	6830	334	323	1384
4	30	5.058789	497	1740	327	193	7773
4	30	6.085261	387	10282	277	658	14532
4	30	6.753771	967	11066	72	1068	11718
4	30	7.887619	12104	5317	537	3153	14028
4	30	8.915725	5752	9699	318	5187	251613
4	30	10.24847	5115	20051	673	2678	159254

Plates and participants are each numbered sequentially. Time points are ordered sequentially within each participant. For example, a data set on 10 participants at three time points each will have $10 \times 3 = 30$ rows. You may use names of proteins (e.g., IL2, etc.) instead of "Prot1", etc. CHEX4 is MFI for nonspecific binding. Data set can be wider than four proteins.

Seventh through fourteenth lines plot logarithm of MFI data versus time for participants for each protein. Note that height and width of plot (in inches) can be specified as can axis labels.

Fifteenth and sixteenth lines fit longitudinal regression model with time, plate, and nonspecific binding as predictor variables, for each protein. Model allows for linear (Time) and curving ($\text{I}(\text{Time}^2)$) trends over time. Add additional predictor variables after $+$ $\text{I}(\log(\text{CHEX4}))$, such as $+$ $\text{I}(\log(\text{CHEX4})) + \text{BMI}$. Use of logarithms assumes all MFI values exceed 0. For more information on constructing and testing assumptions of `lme` model see Pinheiro and Bates, *Mixed-Effects Models in S and S-PLUS*, New York: Springer; 2000.