beeswarm\_alexia

Marialexia Alfano

November 29, 2017

knitr::opts\_chunk$set(echo = TRUE, warning = FALSE, message = FALSE)

library(dplyr)  
library(readr)  
library(ggplot2)  
library(ggbeeswarm)  
library(ggthemes)  
library(tidyr)  
library(scales)

test\_data <- read\_csv(paste0("https://raw.githubusercontent.com/KatieKey/input\_output\_shiny\_group/",  
 "master/CSV\_Files/efficacy\_summary.csv"))  
#testdata <- read\_excel("DRUG\_CLASS\_I\_Mean\_Cmax\_Trough\_Efficacy\_R\_DATA\_ANALYSIS.xlsx", skip = 1, n\_max = 12, sheet = )  
summary(test\_data)

## drug dosage dose\_int level   
## Length:21 Min. : 50.00 Length:21 Length:21   
## Class :character 1st Qu.: 50.00 Class :character Class :character   
## Mode :character Median : 50.00 Mode :character Mode :character   
## Mean : 69.05   
## 3rd Qu.:100.00   
## Max. :100.00   
##   
## PLA ULU RIM OCS   
## Min. : 19 Min. : 37 Min. : 49 Min. : 117   
## 1st Qu.: 832 1st Qu.: 1674 1st Qu.: 1662 1st Qu.: 1430   
## Median : 4170 Median :10332 Median : 8760 Median : 8295   
## Mean : 8707 Mean :12661 Mean :10601 Mean :10768   
## 3rd Qu.:13452 3rd Qu.:18300 3rd Qu.:14744 3rd Qu.:12688   
## Max. :38675 Max. :49350 Max. :39100 Max. :45050   
## NA's :1 NA's :1 NA's :1   
## ICS SLU SLE ELU   
## Min. : 196.0 Min. : 59 Min. : 63 Min. :0.680   
## 1st Qu.: 780.8 1st Qu.: 2247 1st Qu.: 2021 1st Qu.:0.800   
## Median : 3725.0 Median :10900 Median : 7370 Median :1.500   
## Mean :10427.7 Mean : 9868 Mean : 9226 Mean :1.544   
## 3rd Qu.:17025.0 3rd Qu.:15150 3rd Qu.:14567 3rd Qu.:2.290   
## Max. :45200.0 Max. :26200 Max. :28200 Max. :2.500   
## NA's :1 NA's :2 NA's :2 NA's :1   
## ESP cLogP huPPB muPPB   
## Min. :-0.2000 Min. :0.1600 Min. :19.00 Min. :20.00   
## 1st Qu.: 0.4800 1st Qu.:0.3100 1st Qu.:39.00 1st Qu.:44.00   
## Median : 0.7550 Median :0.7200 Median :53.00 Median :54.00   
## Mean : 0.9875 Mean :0.9724 Mean :55.43 Mean :55.89   
## 3rd Qu.: 1.6275 3rd Qu.:1.6200 3rd Qu.:70.75 3rd Qu.:67.50   
## Max. : 2.1800 Max. :2.3800 Max. :99.00 Max. :93.00   
## NA's :1 NA's :7 NA's :2   
## MIC\_Erdman MICserumErd MIC\_Rv Caseum\_binding   
## Min. :0.0600 Min. :0.125 Min. :0.1950 Min. : 6.50   
## 1st Qu.:0.1250 1st Qu.:0.250 1st Qu.:0.3900 1st Qu.:21.00   
## Median :0.5000 Median :1.000 Median :1.0000 Median :29.00   
## Mean :0.6571 Mean :1.315 Mean :0.9544 Mean :36.69   
## 3rd Qu.:1.0000 3rd Qu.:2.000 3rd Qu.:1.5600 3rd Qu.:57.16   
## Max. :2.0000 Max. :4.000 Max. :2.0000 Max. :83.64   
## NA's :3   
## MacUptake   
## Min. : 0.640   
## 1st Qu.: 3.610   
## Median : 7.905   
## Mean :10.768   
## 3rd Qu.:18.000   
## Max. :30.000   
## NA's :1

head(test\_data)

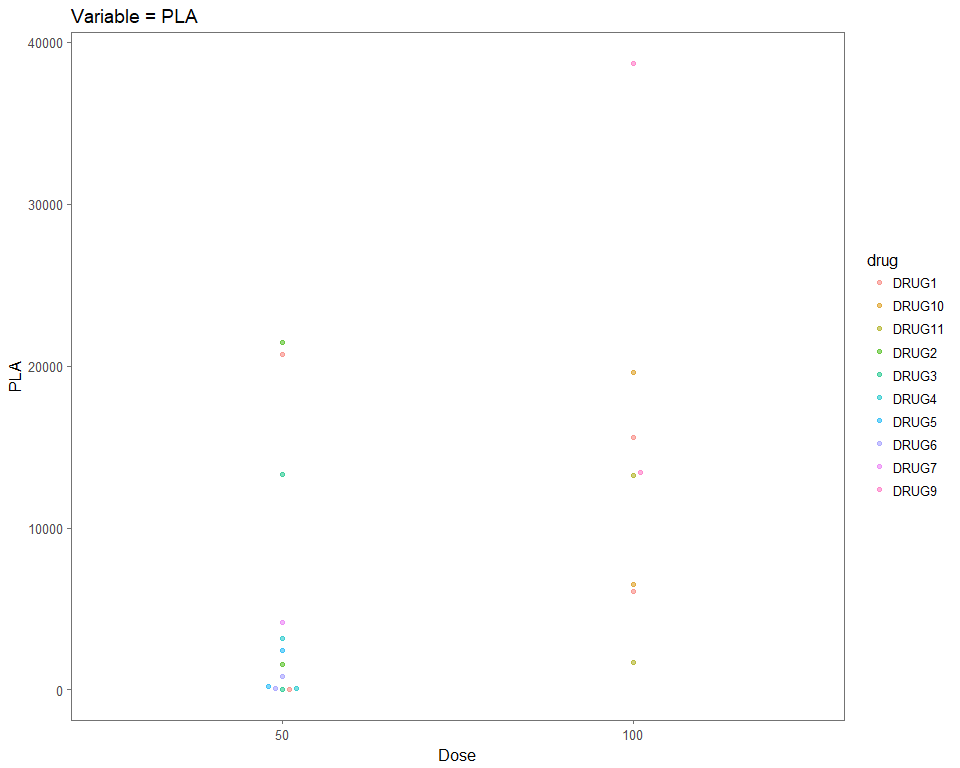
## # A tibble: 6 x 21  
## drug dosage dose\_int level PLA ULU RIM OCS ICS SLU SLE  
## <chr> <int> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 DRUG1 50 BID Cmax 20700 21900 20300 23800 23300 10900 17500  
## 2 DRUG2 50 BID Cmax 21433 36800 29900 45050 35950 20000 28200  
## 3 DRUG3 50 BID Cmax 13295 9480 10400 8760 2480 8170 7370  
## 4 DRUG4 50 BID Cmax 3203 11185 9185 1452 332 12450 5985  
## 5 DRUG5 50 BID Cmax 2420 24950 17200 6535 612 20400 15313  
## 6 DRUG6 50 BID Cmax 832 17100 11085 8080 2275 14700 9420  
## # ... with 10 more variables: ELU <dbl>, ESP <dbl>, cLogP <dbl>,  
## # huPPB <int>, muPPB <int>, MIC\_Erdman <dbl>, MICserumErd <dbl>,  
## # MIC\_Rv <dbl>, Caseum\_binding <dbl>, MacUptake <dbl>

## In-Vivo Distribution of TB Drugs

#select for in-vivo marker variables  
vivo\_sample\_test <- test\_data %>%  
 select(drug, dosage, PLA, ULU, RIM, OCS, ICS, SLU, SLE) %>%   
 mutate(dosage = factor(dosage))  
   
head(vivo\_sample\_test)

## # A tibble: 6 x 9  
## drug dosage PLA ULU RIM OCS ICS SLU SLE  
## <chr> <fctr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>  
## 1 DRUG1 50 20700 21900 20300 23800 23300 10900 17500  
## 2 DRUG2 50 21433 36800 29900 45050 35950 20000 28200  
## 3 DRUG3 50 13295 9480 10400 8760 2480 8170 7370  
## 4 DRUG4 50 3203 11185 9185 1452 332 12450 5985  
## 5 DRUG5 50 2420 24950 17200 6535 612 20400 15313  
## 6 DRUG6 50 832 17100 11085 8080 2275 14700 9420

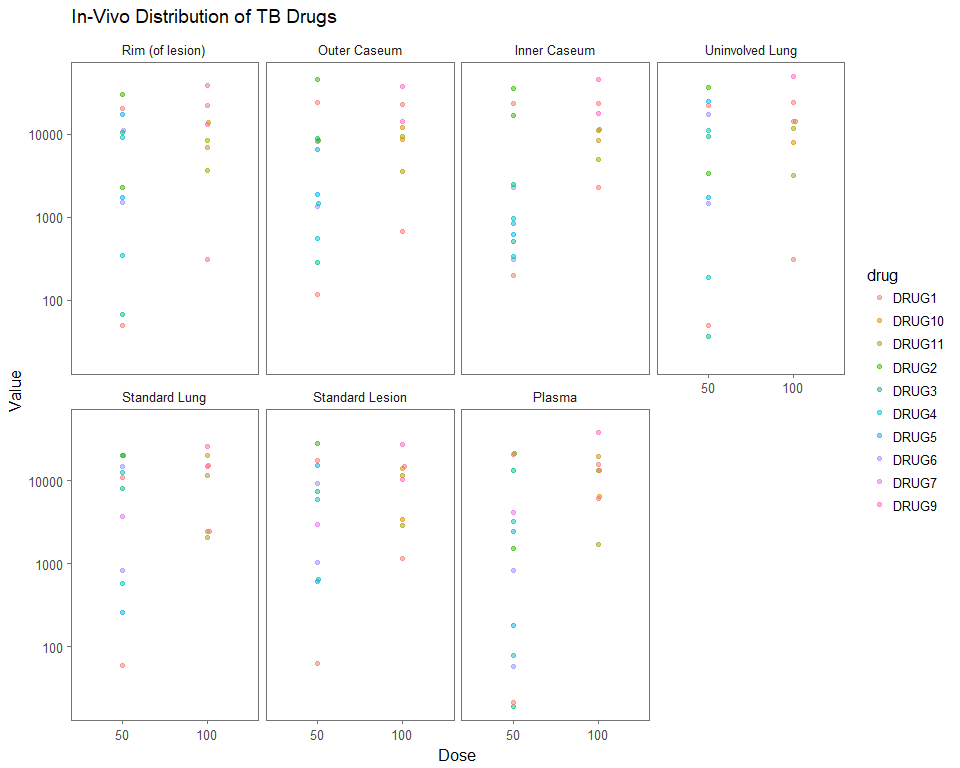
#single variable plot for in-vivo marker  
test\_plot <- vivo\_sample\_test %>%   
 ggplot(aes(dosage, PLA, color = drug))+  
 geom\_beeswarm(alpha=.5,size=1.5)+  
 ggtitle('Variable = PLA') + labs(x='Dose')+  
 theme\_few()  
  
test\_plot



#gather for small multiples in-vivo markers  
vivo\_sample\_forSM <- vivo\_sample\_test %>%   
 gather(key = variable, value = value, -drug, -dosage) %>%   
 mutate(variable = factor(variable, levels = c("RIM", "OCS","ICS","ULU","SLU","SLE","PLA"),  
 labels = c("Rim (of lesion)","Outer Caseum","Inner Caseum","Uninvolved Lung",  
 "Standard Lung", "Standard Lesion", "Plasma")))  
  
  
head(vivo\_sample\_forSM)

## # A tibble: 6 x 4  
## drug dosage variable value  
## <chr> <fctr> <fctr> <dbl>  
## 1 DRUG1 50 Plasma 20700  
## 2 DRUG2 50 Plasma 21433  
## 3 DRUG3 50 Plasma 13295  
## 4 DRUG4 50 Plasma 3203  
## 5 DRUG5 50 Plasma 2420  
## 6 DRUG6 50 Plasma 832

#plot small multiples in-vivo markers   
vivo\_sample\_SMplot <- vivo\_sample\_forSM %>%   
 ggplot(aes(x = dosage, y = value, color = drug))+  
 geom\_beeswarm(alpha = 0.5, size = 1.5)+  
 scale\_y\_log10()+  
 labs(x = 'Dose', y = 'Value')+  
 ggtitle('In-Vivo Distribution of TB Drugs')+  
 theme\_few()+  
 facet\_wrap(~ variable, ncol = 4)  
  
vivo\_sample\_SMplot

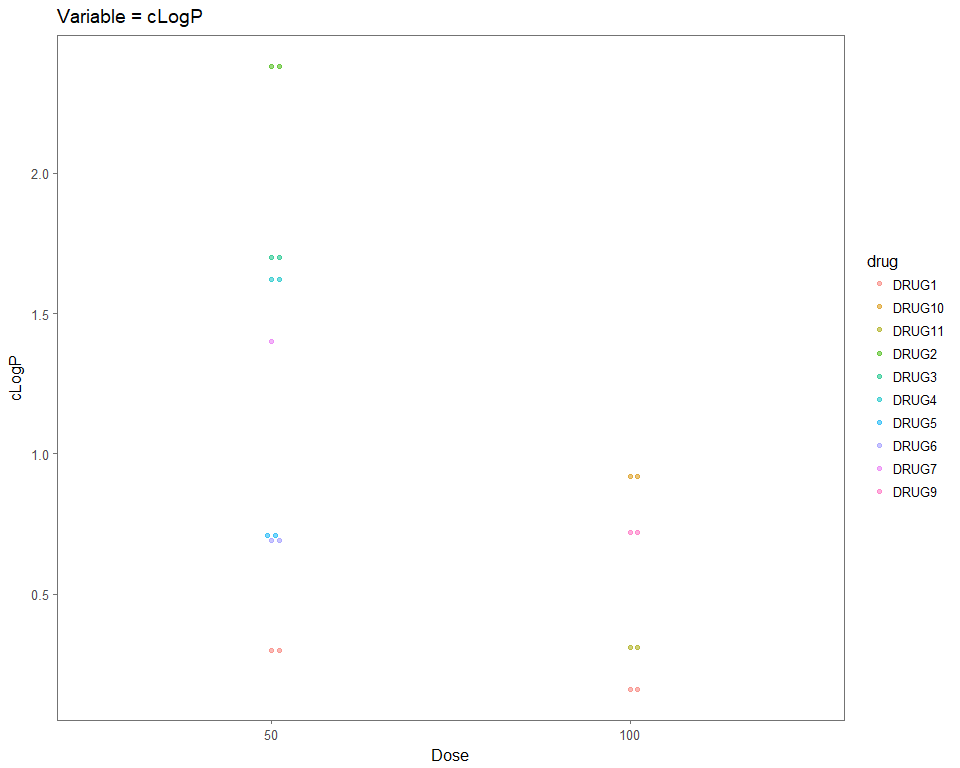


## In-Vitro Distribution of TB Drugs

#select for in-vitro marker variables  
vitro\_sample\_test <- test\_data %>%  
 select(drug, dosage, cLogP, huPPB, muPPB,   
 MIC\_Erdman, MICserumErd, MIC\_Rv, Caseum\_binding, MacUptake) %>%   
 mutate(dosage = factor(dosage))  
  
head(vitro\_sample\_test)

## # A tibble: 6 x 10  
## drug dosage cLogP huPPB muPPB MIC\_Erdman MICserumErd MIC\_Rv  
## <chr> <fctr> <dbl> <int> <int> <dbl> <dbl> <dbl>  
## 1 DRUG1 50 0.30 19 20 0.500 1.00 1.560  
## 2 DRUG2 50 2.38 61 64 0.125 0.25 0.195  
## 3 DRUG3 50 1.70 NA 71 0.125 0.25 0.390  
## 4 DRUG4 50 1.62 99 93 0.060 2.00 0.195  
## 5 DRUG5 50 0.71 74 64 0.060 0.25 0.390  
## 6 DRUG6 50 0.69 NA 54 0.500 1.00 1.560  
## # ... with 2 more variables: Caseum\_binding <dbl>, MacUptake <dbl>

#single variable plot for in-vitro marker  
test\_plot\_2 <- vitro\_sample\_test %>%   
 ggplot(aes(dosage, cLogP, color = drug))+  
 geom\_beeswarm(alpha=.5,size=1.5)+  
 ggtitle('Variable = cLogP') + labs(x='Dose')+  
 theme\_few()  
  
test\_plot\_2



#gather for small multiples in-vitro markers  
vitro\_sample\_forSM <- vitro\_sample\_test %>%   
 gather(key = variable, value = value, -drug, -dosage) %>%   
 mutate(variable = factor(variable, levels = c("Caseum\_binding", "cLogP", "huPPB", "muPPB", "MIC\_Erdman", "MICserumErd", "MIC\_Rv", "MacUptake"),  
 labels = c("Caseum \nBinding", "In-Vitro \nLipophilic \nBinding",   
 "Human \nPlasma \nBinding", "Mouse \nPlasma \nBinding",   
 "MIC Erdman \nStrain", "MIC Erdman \nStrain \nwith Serum", "MIC Rv Strain",  
 "Macrophage \nUptake (Ratio)")))   
  
head(vitro\_sample\_forSM)

## # A tibble: 6 x 4  
## drug dosage variable value  
## <chr> <fctr> <fctr> <dbl>  
## 1 DRUG1 50 "In-Vitro \nLipophilic \nBinding" 0.30  
## 2 DRUG2 50 "In-Vitro \nLipophilic \nBinding" 2.38  
## 3 DRUG3 50 "In-Vitro \nLipophilic \nBinding" 1.70  
## 4 DRUG4 50 "In-Vitro \nLipophilic \nBinding" 1.62  
## 5 DRUG5 50 "In-Vitro \nLipophilic \nBinding" 0.71  
## 6 DRUG6 50 "In-Vitro \nLipophilic \nBinding" 0.69

#plot small multiples in-vitro markers   
vitro\_sample\_SMplot <- vitro\_sample\_forSM %>%   
 ggplot(aes(x = dosage, y = value, color = drug))+  
 geom\_beeswarm(alpha = 0.5, size = 1.5)+  
 labs(x = 'Dose', y = 'Value')+  
 ggtitle('In-Vitro Distribution of TB Drugs')+  
 theme\_few()+  
 facet\_wrap(~ variable, ncol = 4, scale="free")  
  
vitro\_sample\_SMplot

