

# Validation of Hepatitis C viral count using R-computing

This is a test

```
## Loading required package: nlme
```

```
## Warning: running command '/usr/bin/otool' -L '/Library/Frameworks/
```

```
## R.framework/Resources/modules/R_de.so' had status 1
```

In a summary. The lower limit of detection (LLD) at home-lab is 12 IU/ml and the LLD at the reference-lab is 20 IU/ml. So, if the result is <20IU/ml, the detected value could be anywhere between 1 and 20. Therefore, the lower limit of detection has been set for home-lab at '6 IU/ml' and '10 IU/ml' for the reference lab.

```
summary(HCV_ICL3)
```

```
##           Nr           ICL           BioS
## Min.      : 2.000   Min.    : 3467   Min.      : 620
## 1st Qu.: 4.000   1st Qu.: 21526   1st Qu.: 29402
## Median : 7.500   Median : 88628   Median : 122300
## Mean    : 7.167   Mean    :1285824   Mean    :1799998
## 3rd Qu.: 9.500   3rd Qu.:2000113   3rd Qu.:2166075
## Max.    :13.000   Max.    :4897788   Max.    :7709000
```

```
head(HCV_ICL3)
```

```
##      Nr      ICL      BioS
## 1  2 2630268 2840000
## 2  3 109648 144300
## 3  7 4897788 7709000
## 4  8   6166   5769
## 5 10  67608 100300
## 6 13   3467    620
```

To make it more easy, the set of values from Reference-lab = 'x'. The set of values from Home-lab = 'y'

Calculate the means and difference between the two sets (x and y)

```
# derive difference
```

```
mean(x)
```

```
## [1] 1285824
```

```
mean(y)
```

```
## [1] 1799998
```

```
# mean Ref_lab - mean Home_lab
```

```
mean(x)-mean(y)
```

```
## [1] -514174
```

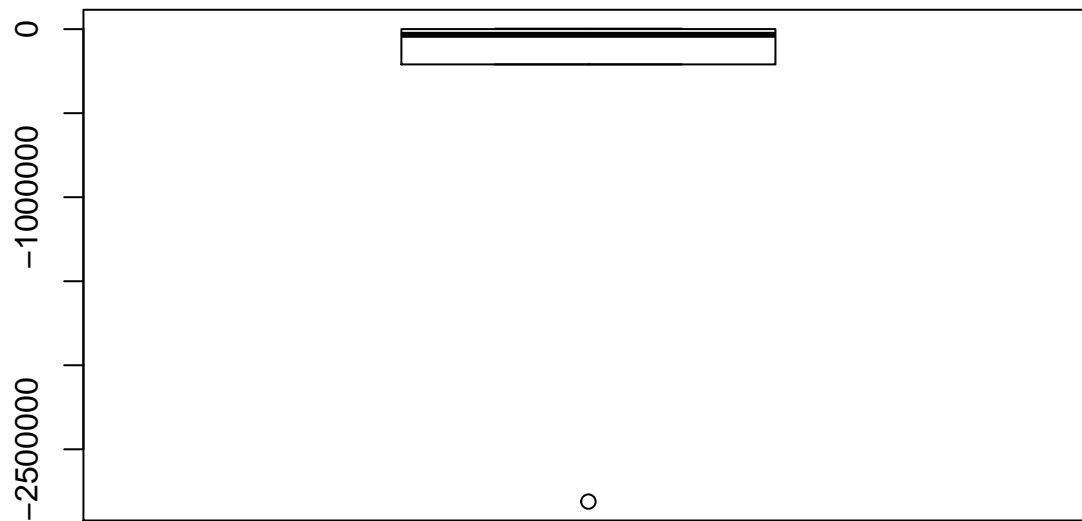
Because n=8 is small, the distribution of the differences should be approximately normal. Check using a boxplot and QQ plot. There is some skew.

```
HCV_ICL3$diff <- x-y
```

```
HCV_ICL3$diff
```

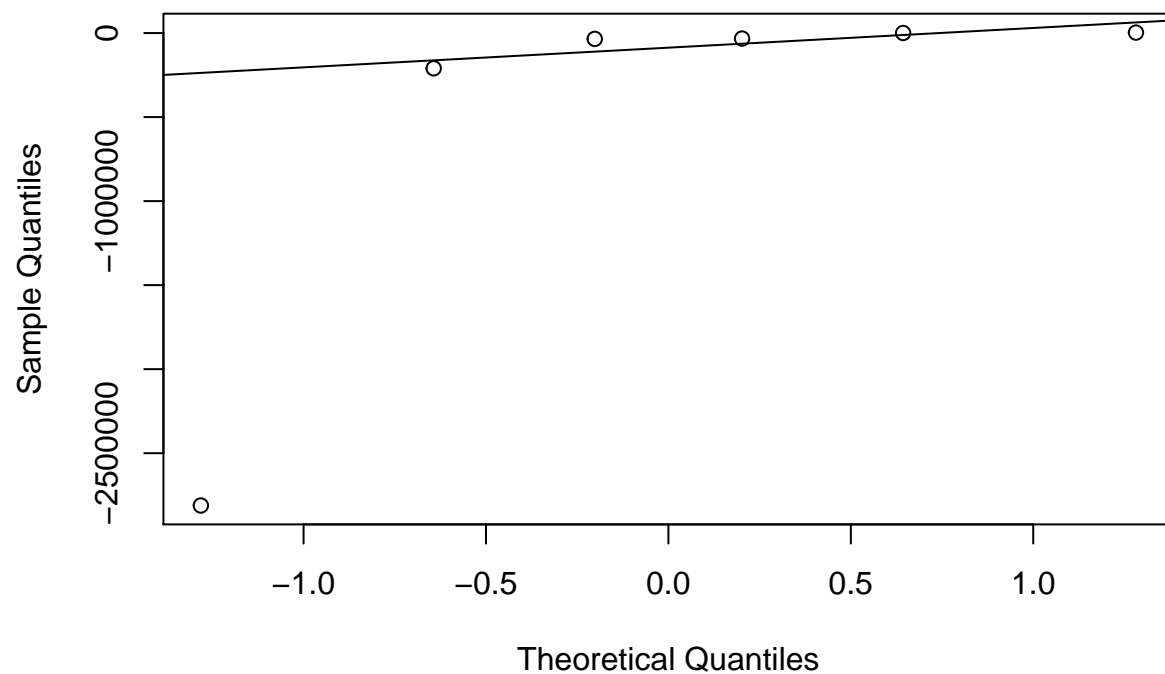
```
## [1] -209732 -34652 -2811212 397 -32692 2847
```

```
boxplot(HCV_ICL3$diff)
```



```
qqnorm(HCV_ICL3$diff)
qqline(HCV_ICL3$diff)
```

**Normal Q-Q Plot**



Shapiro test of normality.

```
shapiro.test(HCV_ICL3$diff)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  HCV_ICL3$diff
## W = 0.5487, p-value = 0.0001049
```

The normality test gives  $p < 0.001$ , which is small, so we reject the null hypothesis that the values are distributed normally.

This means that we cannot use the student t-test. Instead, use the Mann-Whitney-Wilcoxon Test, we can decide whether the population distributions are identical without assuming them to follow the normal distribution.

```
wilcox.test(x, y, paired = TRUE)
```

```
##
## Wilcoxon signed rank test
##
## data: x and y
## V = 3, p-value = 0.1563
## alternative hypothesis: true location shift is not equal to 0
```

$p > 0.05$  and therefore the  $H_0$  is NOT rejected. The two populations are identical.

Just to see what happens in the Student T-test. A paired t-test: one sample, two tests  $H_0$  = no difference;  $H_1$  = mean of 2 tests are different  $\mu$  = a number indicating the true value of the mean (or difference in means if you are performing a two sample test).

```
t.test(x, y, mu=0, paired=T, alternative="greater")
```

```
##
## Paired t-test
##
## data: x and y
## t = -1.1165, df = 5, p-value = 0.8425
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
## -1442186 Inf
## sample estimates:
## mean of the differences
## -514174
```

$p = 0.8425$ . Because  $p$  is larger than  $\alpha$ , we do NOT reject  $H_0$ . In other words, it is unlikely the observed agreements happened by chance. However, because the populations do not have a normal distribution, we can not use the outcome of this test.

For correlation, three methods are used: pearson, kendall and spearman at a confidence level of 95%.

```
# correlation of the two methods
cor.test(x, y,
         alternative = c("two.sided", "less", "greater"),
         method = c("pearson", "kendall", "spearman"),
         exact = NULL, conf.level = 0.95)
```

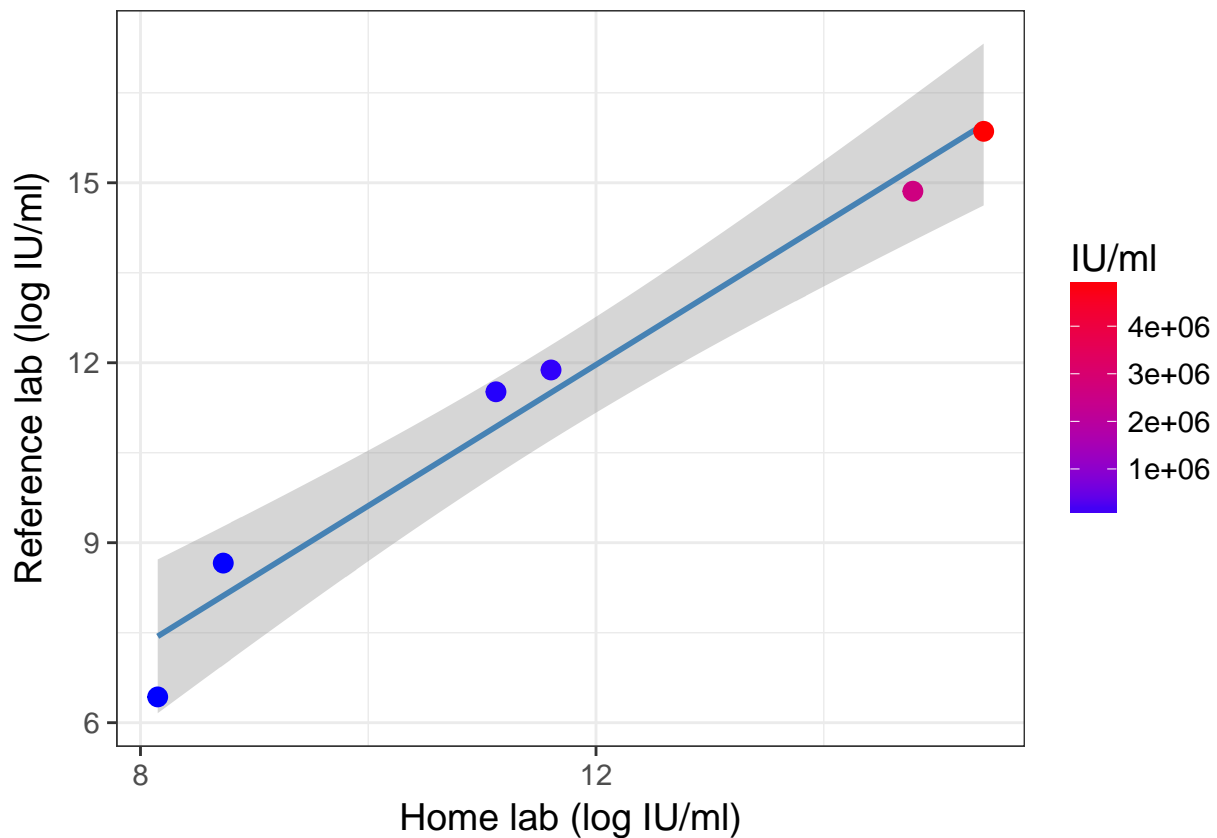
```
##
## Pearson's product-moment correlation
##
## data: x and y
## t = 12.226, df = 4, p-value = 0.000257
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.8806391 0.9986275
## sample estimates:
## cor
## 0.9868827
```

The correlation with the spearman test is 0.9868827. Almost perfect correlation.

Plotting the two methods using logarithmic scales.

```
g <- ggplot(HCV_ICL3, aes(log(ICL), log(BioS)))

# add layers
g +
  geom_smooth(method="lm", se=TRUE, col="steelblue", size = 1) +
  geom_point(size = 3, aes(colour = x)) +
  scale_colour_gradient("IU/ml", high = "red", low = "blue", space = "Lab") +
  labs(y = "Reference lab (log IU/ml)" +
  labs(x = "Home lab (log IU/ml)" +
  theme_bw(base_family = "Helvetica", base_size = 14) +
  scale_x_continuous(breaks=c(0,4,8,12))
```



Summary data on the correlation line.

```
regmod <- lm(y~x, data=HCV_ICL3)
summary(regmod)

##
## Call:
## lm(formula = y ~ x, data = HCV_ICL3)
##
## Residuals:
##      1      2      3      4      5      6
## -967570 100611 515487 116603 119386 115484
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
##
```

```
## (Intercept) -1.200e+05  2.773e+05  -0.433 0.687386
## x           1.493e+00  1.221e-01  12.226 0.000257 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 559700 on 4 degrees of freedom
## Multiple R-squared:  0.9739, Adjusted R-squared:  0.9674
## F-statistic: 149.5 on 1 and 4 DF,  p-value: 0.000257
```

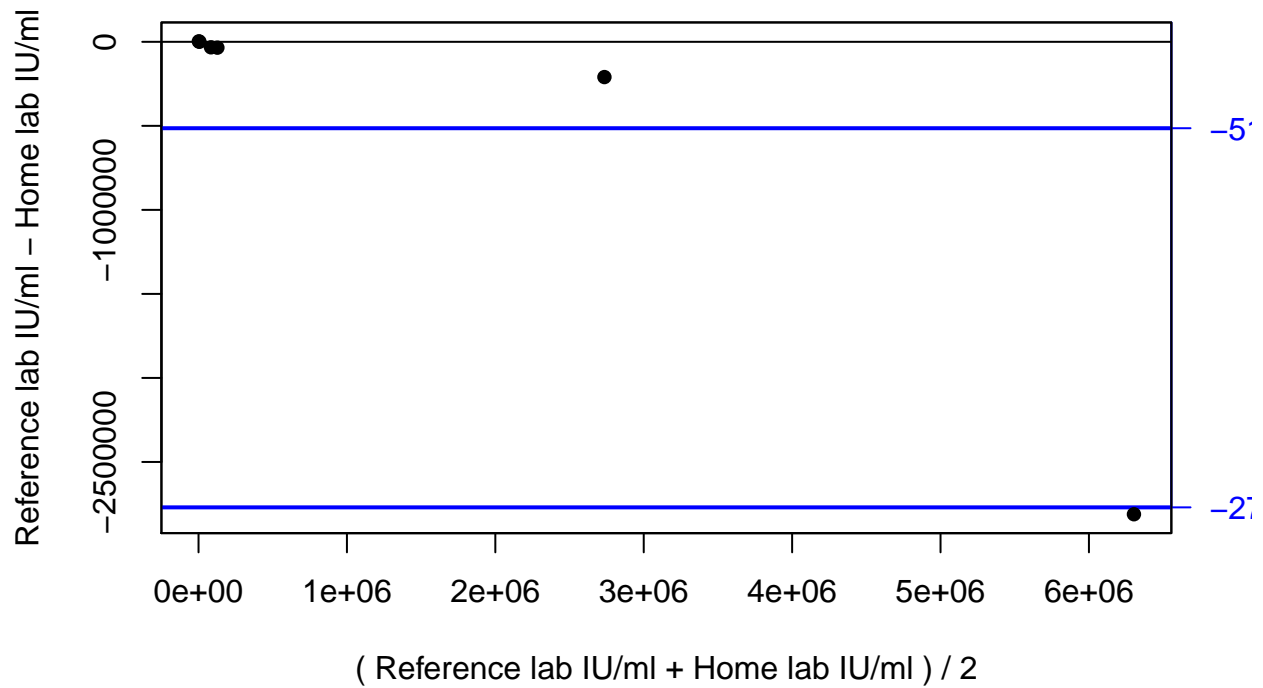
The Bland-Altman Analysis. To check if there is a bias.

```
##      ICL      BioS      diff
## 1 2630268 2840000 -209732
## 2  109648  144300 -34652
## 3 4897788 7709000 -2811212
## 4    6166    5769    397
## 5   67608   100300 -32692
## 6    3467     620    2847
```

```
BlandAltman(x, y,
  x.name = "Reference lab IU/ml",
  y.name = "Home lab IU/ml",
  maintit = "Bland-Altman plot for HBV count",
  cex = 1,
  pch = 16,
  col.points = "black",
  col.lines = "blue",
  limx = NULL,
  limy = NULL,
  ymax = NULL,
  eqax = FALSE,
  xlab = NULL,
  ylab = NULL,
  print = TRUE,
  reg.line = FALSE,
  digits = 2,
  mult = FALSE)
```

```
## NOTE:
## 'AB.plot' and 'BlandAltman' are deprecated,
## and likely to disappear in a not too distant future,
## use 'BA.plot' instead.
```

## Bland-Altman plot for HBV count



```
##
## Limits of agreement:
## Reference lab IU/ml - Home lab IU/ml          2.5% limit
##                                     -514174      -2770353
##                                     97.5% limit      SD(diff)
##                                     1742005      1128089
```

When the dots are around 0, the two test could be interchanged for a patient. So, the two test can be interchanged. There are, however, some outliers: large difference of viral count between the two labs.