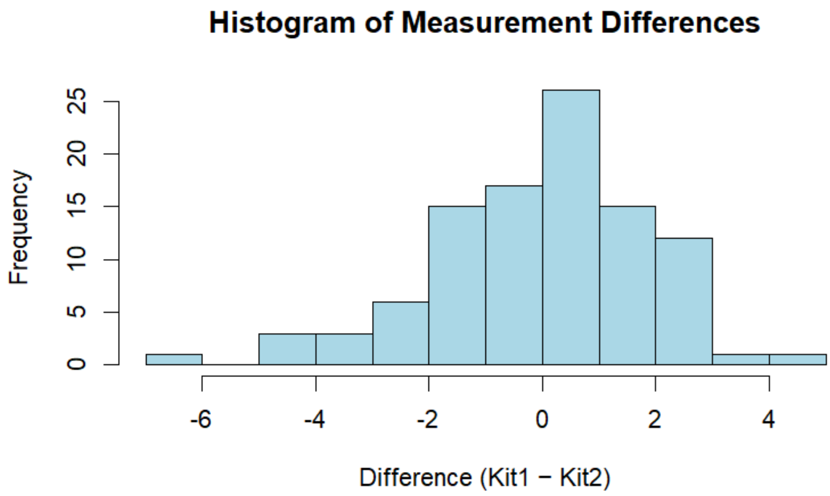
Task 2:

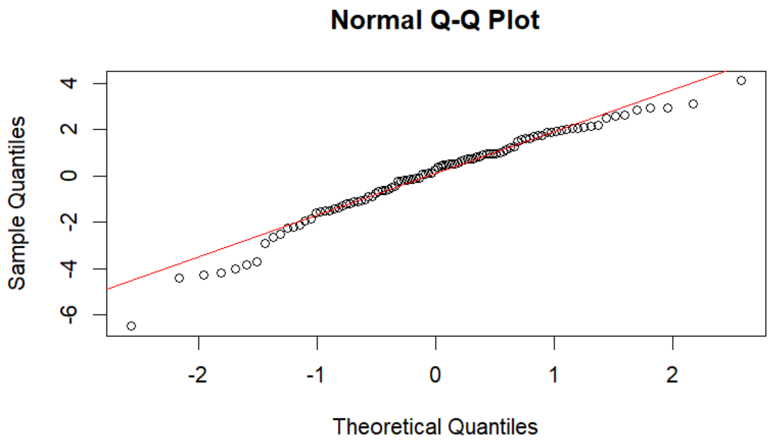
Explore methods for analysing the agreement and compare these.

We start by exploring the data set. Which includes two different measurement kits.

```
  Cyfra_Kit1        Cyfra_Kit2
 Min.   :0.469   Min.   :0.010
 1st Qu.:2.099   1st Qu.:1.640
 Median :3.181   Median :3.018
 Mean   :3.310   Mean   :3.290
 3rd Qu.:4.383   3rd Qu.:5.006
 Max.   :7.375   Max.   :8.818
```

We compute the differences for each measurement and visualize it using different plots. We continue with estimating the normality of the difference.



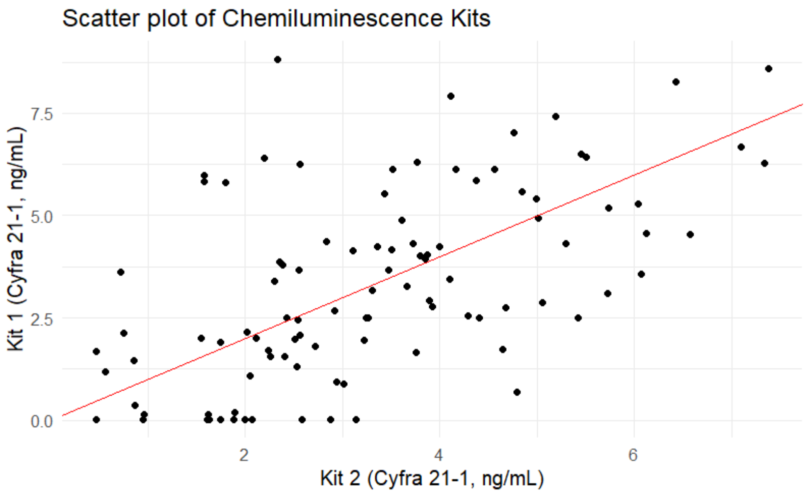


```
> shapiro.test(data$diff)

        Shapiro-Wilk normality test

data:  data$diff
W = 0.96771, p-value = 0.01481
```

According to the normality test we should reject the hypothesis that the difference is normal. However, according to both the histogram and the qq plot we still assume normality. Since approximate normality is something that the bland-altman assumes and not perfect.

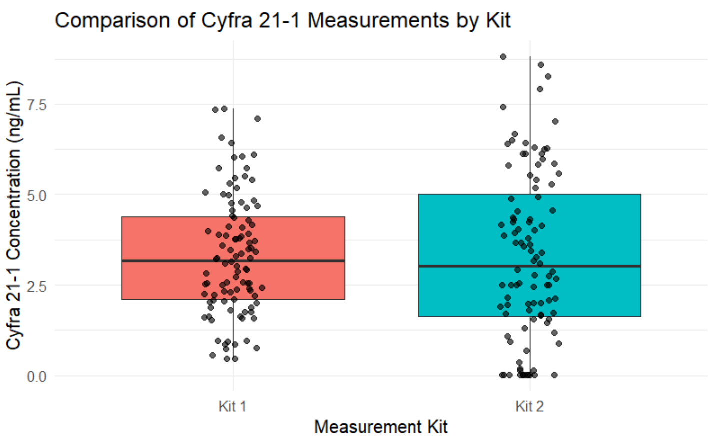
We continue with visualising the agreement of the two kits.

We start with a scatter plot of the two kits.

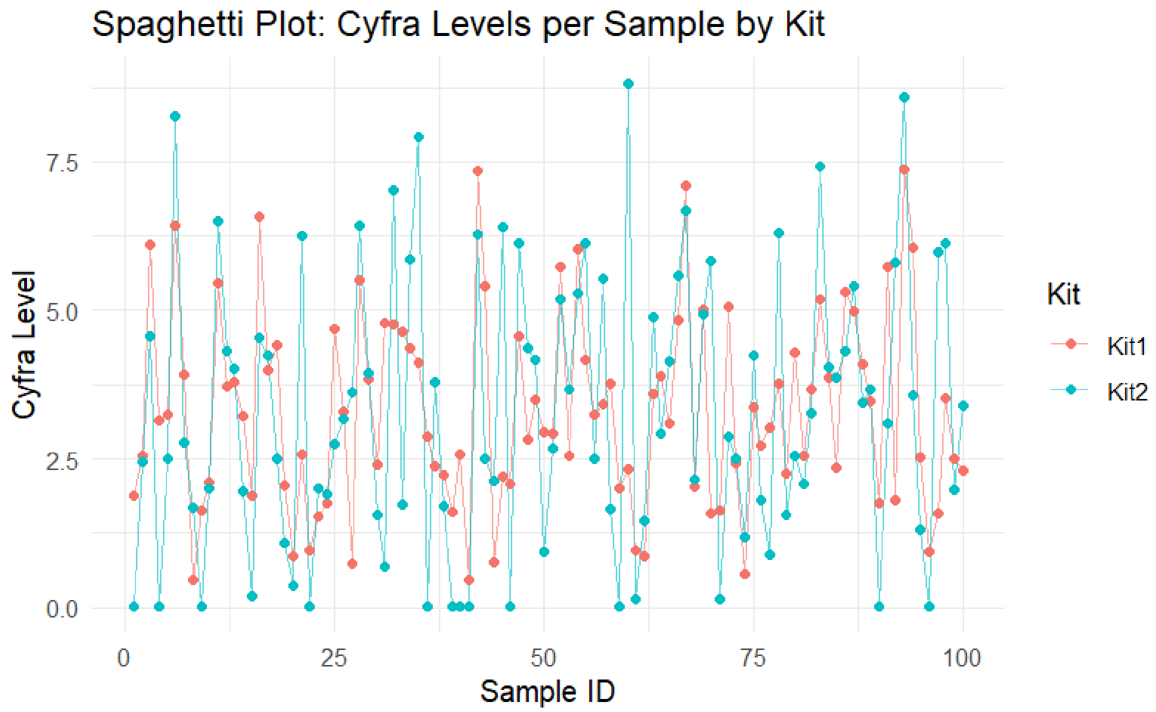


We can not determine a clear dependence between the two kits.

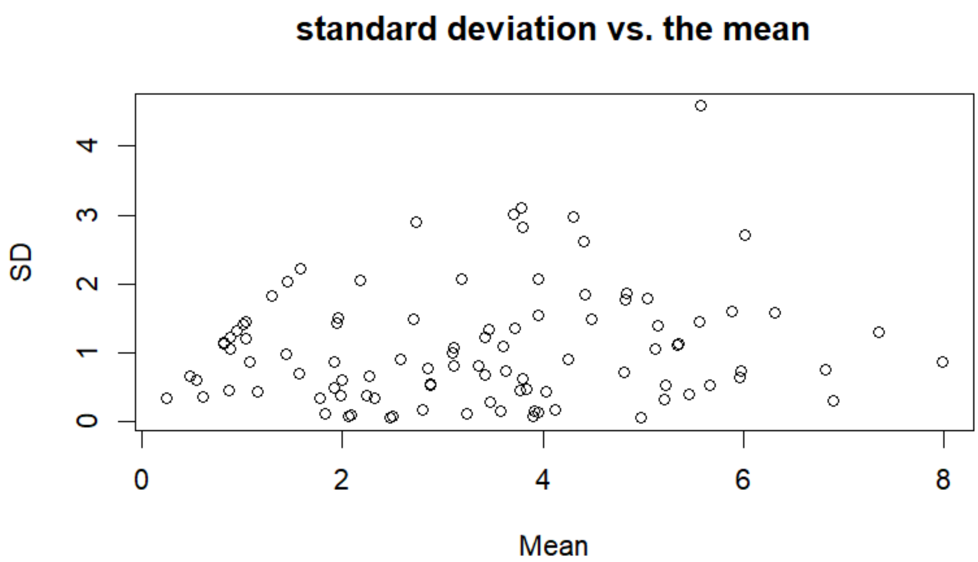
We go on to plot a side by side histogram plot which indicates that kit 2 has larger variability.



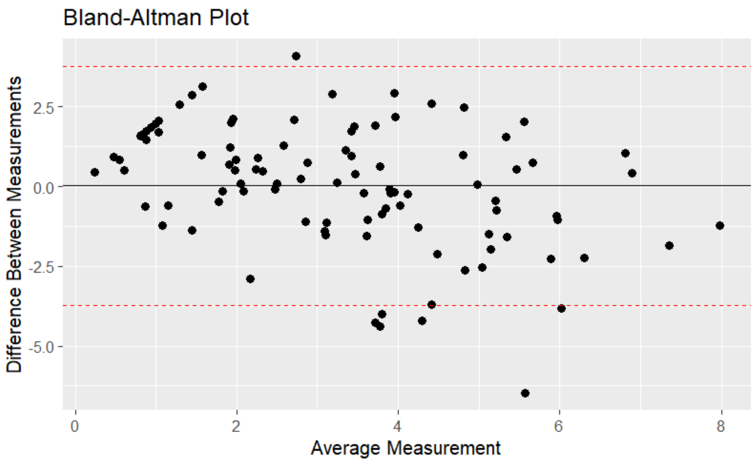
A spaghetti plot of individual samples measured by both kits shows that Kit 2 consistently produces either systematically higher or lower values for each sample, suggesting a difference in spread as seen in the boxplot.



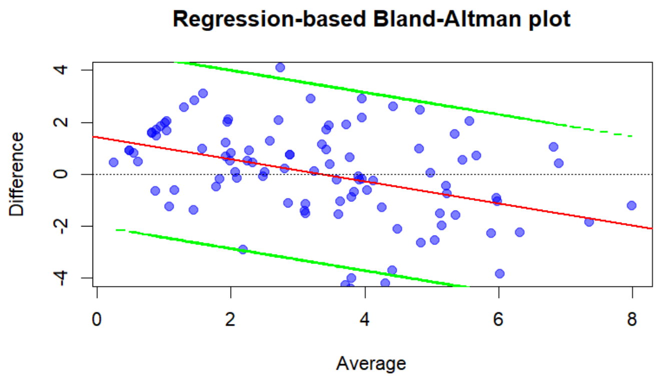
Plotting the standard deviation vs. the mean reveals no strong evidence of proportional heteroscedasticity, but we observe slightly wider variability at higher concentrations. This hints that a regression-based agreement model may be more appropriate.



The standard BA plot shows a downward trend in differences across the concentration range. This indicates: Proportional bias and violation of the constant difference assumption which motivates trying a regression-based method.



We try a regression-based Bland-Altman plot, and we clearly see that this is more appropriate.



CCC tells you whether two methods truly agree, while ICC tells you whether measurements are consistent — you can have a high ICC and still have poor agreement. This is why we compute Lin´s CCC.

```
> # Lin's CCC from epiR
> epi.ccc(data$Cyfra_Kit1, data$Cyfra_Kit2)
$rho.c
        est     lower     upper
1 0.5518658 0.4130112 0.6656461
```

0.5518 -> This value indicates poor to moderate agreement. CCC values closer to 1.0 indicate perfect agreement.

The CL interval is far from 1.0, confirming the poor agreement is statistically significant

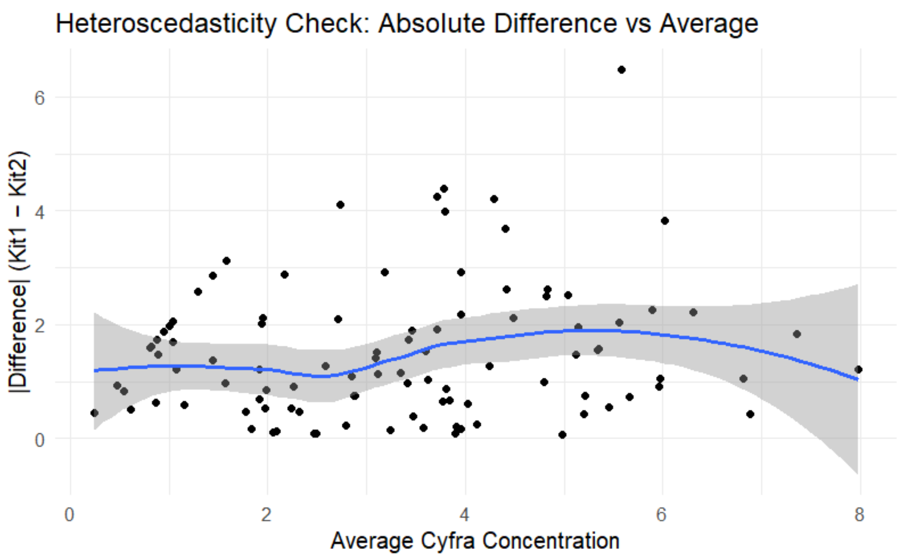
```
$s.shift
[1] 1.408829

$l.shift
[1] -0.01027646
```

1.41 -> relates to the ratio of the Standard Deviations of the two methods SD_Kit2 /SD_Kit1. A value >1 means Kit2 has a wider spread higher SD than Kit1.

-0.01 -> This relates to the difference between the means of the two methods Mean_Kit2 - Mean_Kti1. A small value means the means are very similar. Despite the near-zero mean difference, there is large random error and poor precision.

Plotting the difference against measurement magnitude shows:

- Larger differences at higher concentration
- No major curvature, but visible proportional bias
- Supports the regression-based BA method



Which method is most suitable given the data?

Given the observations:

- Non-constant variability
- Downward trend in the standard BA plot
- Evidence of proportional bias
- Larger variability in Kit 2
- CCC confirming poor precision

The regression-based Bland–Altman method is the most appropriate for this dataset.

It properly accounts for:

- Proportional bias
- Heteroscedasticity
- Non-constant limits of agreement

The classical Bland–Altman would misrepresent the agreement because it assumes constant variability.