Some Assays Relevant to the Novel Coronavirus

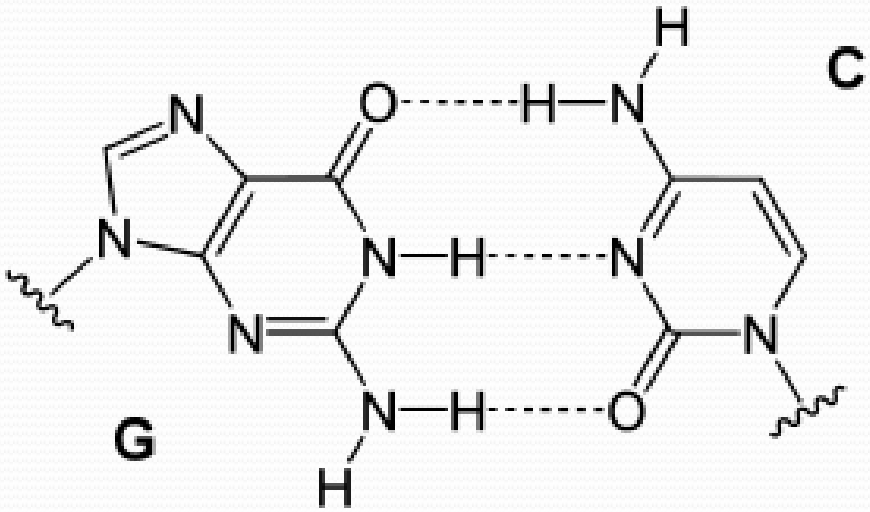
EPI 204

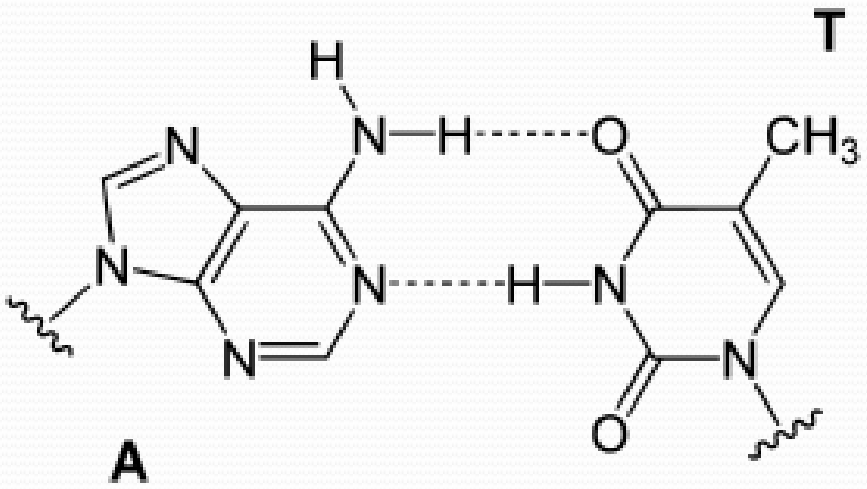
Quantitative Epidemiology III

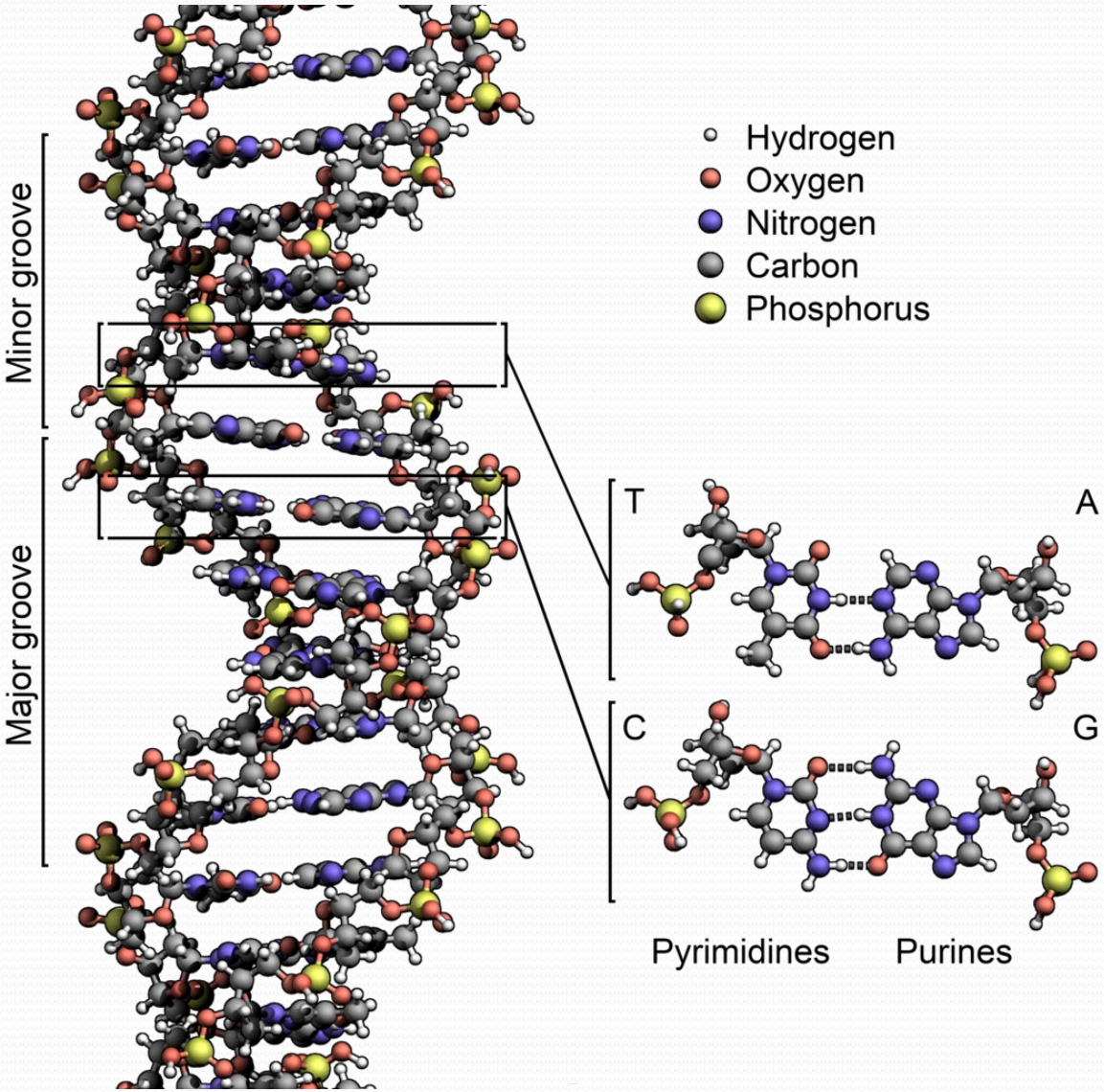
Statistical Models

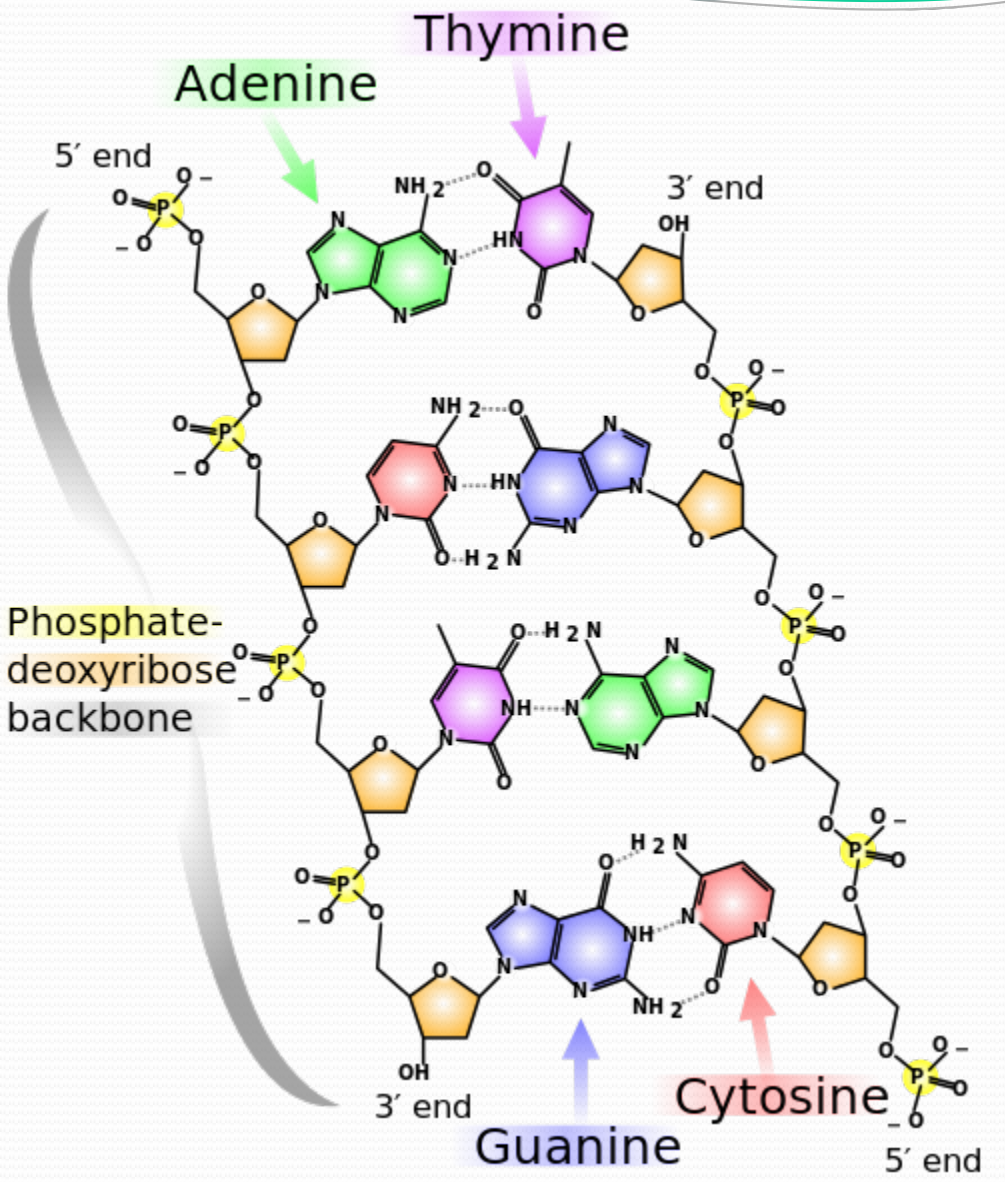
DNA and RNA

- DNA and RNA are polymers composed of four subunits.
- The four subunits in DNA are the nucleotides guanine, adenine, thymine, and cytosine (G, A, T, C) on a sugar (deoxyribose)/phosphate backbone.
- Pairs of nucleotides bind well to each other by hydrogen bonds, with GC forming one pair and AT the other.
- A and G are purines, C and T are pyrimidines. Uracil (U) is a pyrimidine that takes the place of thymine in RNA
- For a given strand of DNA the complementary strand (cDNA) has the matching nucleotides, so
  GGTCACTG matches
  CCAGTGAC


C
G


T
A


Minor groove
Major groove
Hydrogen
Oxygen
Nitrogen
Carbon
Phosphorus
T
A
C
G
Pyrimidines
Purines


Thymine
Adenine
5′ end
3′ end
Phosphate-
deoxyribose
backbone
3′ end
Guanine
Cytosine
5′ end

Transcription

- DNA can be transcribed to RNA (which has a different sugar on the backbone) using RNA polymerase
- RNA can be reverse-transcribed to DNA to allow DNA assays to be used on RNA
- DNA transcription in eukaryotes involves editing such as removal of introns.
- Reverse transcription is used in the lab for example for applying PCR to RNA.

DNA, RNA, and Retro Viruses

- These all have a core of DNA or RNA
- Retroviruses (HIV) reverse transcribe the RNA core to DNA after infection
- If the virus has an RNA core, reverse transcription can create DNA particles
- In either case, the assay will try to measure DNA concentration

PCR for Measurement

- Polymerase Chain Reaction (PCR) is a method of measuring the copy number of a particular DNA sequence in a sample.
- It uses an enzyme (DNA polymerase) that copies one strand of DNA to its complement, and this is done with both strands, if double-stranded.
- Primers bind to a specific sequence, so only DNA with that sequence is amplified.
- A cycle of temperature changes should result in approximately doubling the copy number.
- The read-out is obtained by a fluorescent dye

- In 20 cycles, the amount of the analyte should be increased by a factor of 2^{20} = 1 million.
- In 40 cycles, the increase is about 2^{40} = 1 trillion.
- If we perform a fixed number of cycles, then the range of measurement is rather narrow.
- Quantitative PCR = Real-Time PCR establishes a threshold of brightness, and the cycle at which it passes that threshold is the measurement. This is usually interpolated between the last cycle dimmer than the threshold and the first cycle brighter than the threshold.

qRT-PCR

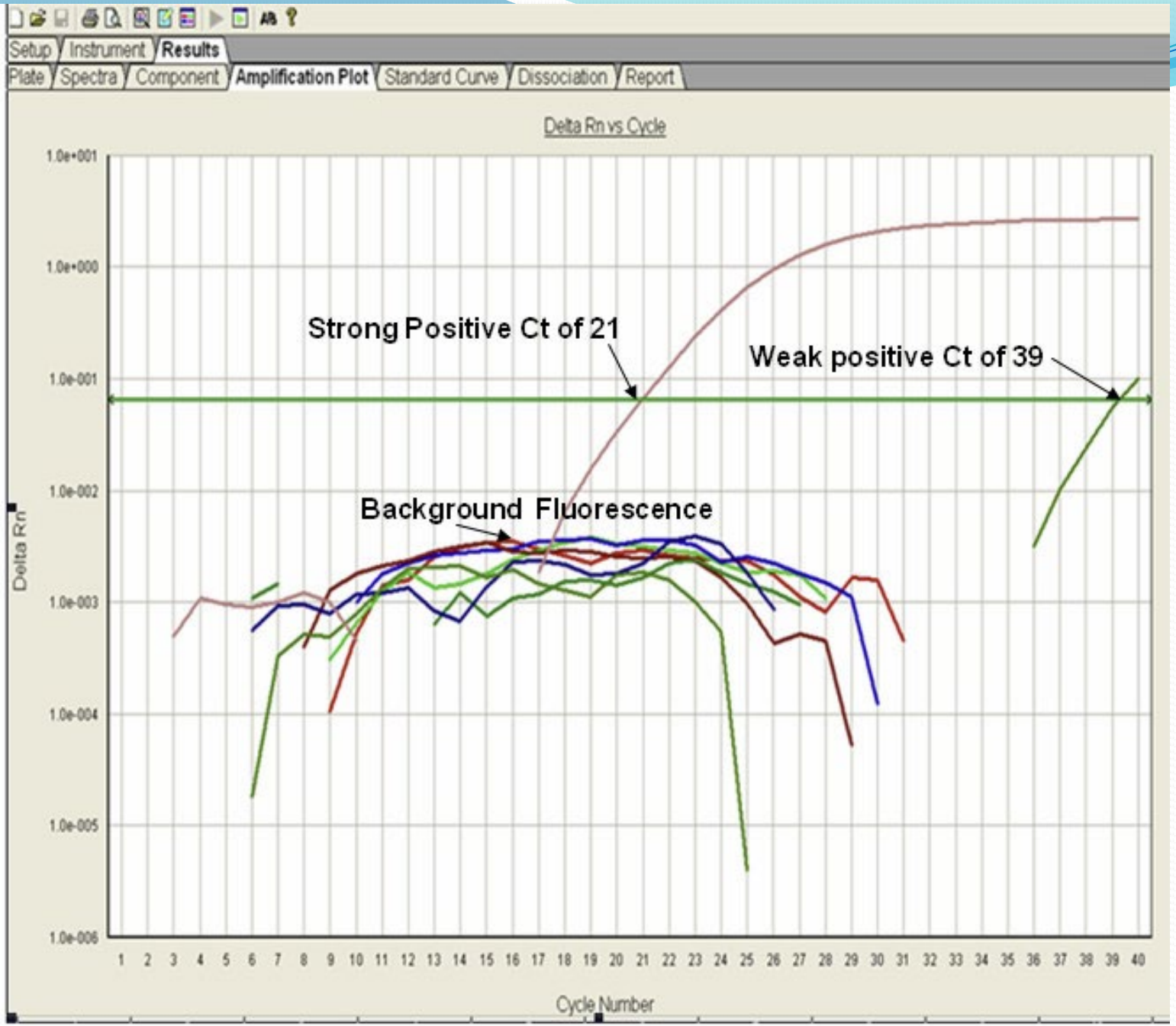
- Reverse Transcription (RT) PCR quantifies RNA by using a reverse transcriptase to create the equivalent DNA sequences.
- RT-PCR should not be used to mean “Real-Time” PCR because it can be confused with “Reverse Transcription” PCR.
- qRT-PCR uses the cycle threshold method after reverse transcription

Sensitivity of PCR

- Very specific and very sensitive.
- A mere 25 copies of a transcript will be amplified to detection in qRT-PCR.
- Suppose we need 70 ng of RNA per run, and suppose we have a sample of RNA with an average copy number of 50 copies per 70 ng.
- The actual copy number is Poisson with a mean of 50, and (therefore) a standard deviation of $\sqrt{50} = 7.07$. Almost all 70 ng aliquots will have a copy number of $50 \pm (3)(7.07) = 50 \pm 21$ or between 29 and 71 copies, and thus will be detected.
- So the detection limit is about 25 copies and the minimum detectable value is about 50 copies.

Quantitative PCR

- We can get relative quantitation because the fluorescence is proportional to the quantity present at the end of the cycle.
- The copy number present at the end of cycle k is roughly 2^k times the copy number at the start.
- To obtain absolute copy number estimates, when those are needed, we can make a calibration curve or use standards along with each sample.


Setup
Instrument
Results
Plate
Spectra
Component
Amplification Plot
Standard Curve
Dissociation
Report
Delta Rn vs Cycle
Strong Positive Ct of 21
Weak positive Ct of 39
Background Fluorescence
Delta Rn
Cycle Number

Interpreting qRT-PCR

- A threshold level is set that is enough above background that it would not happen by chance
- The response Ct is the cycle at which the signal exceeds the threshold
- We usually interpolate. If the threshold is 100, and the signal at cycle 20 is 88, and the signal at cycle 21 is 169, then we could interpolate linearly, but interpolation on the log scale is better, because the signal rises exponentially.

$$\text{Threshold} = 100$$

$$S_{20} = 88$$

$$S_{21} = 169$$

$$C_t = 20 + (100 - 88) / (169 - 88) = 20.15$$

$$\ln(S_{20}) = \ln(88) = 4.417$$

$$\ln(S_{21}) = \ln(169) = 5.130$$

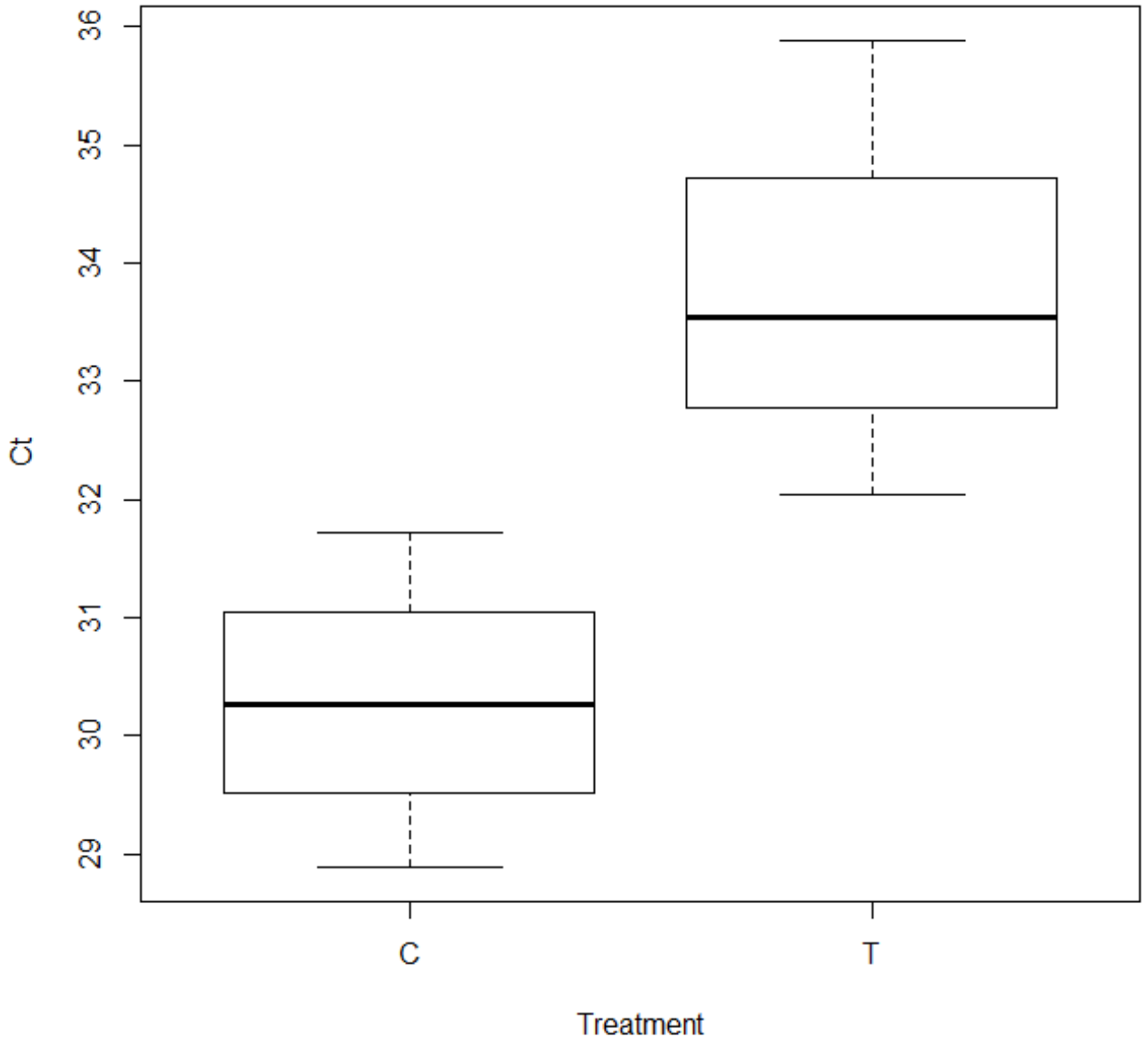
$$\ln(100) = 4.605$$

$$C_t = 20 + (4.605 - 4.417) / (5.130 - 4.417) = 20.20$$

- The higher the Ct value, the lower the original copy number
- Ct is on the log scale so this is often a good scale for ANOVA and regression
- If there is a control in each well/tube, we then can analyze Ct(sample) – Ct(control)
- If we need to know the actual copy number (as for HIV), we need a calibration curve, which may be run on a periodic basis.

Example

- Eight HIV patients are assayed, four with a new treatment and four controls with current standard of care.
- The Ct values for the treated patients are 32.03, 35.89, 33.57, and 33.51
- The Ct values for the control patients are 30.14, 31.72, 28.88, and 30.38.
- Lower Ct values mean higher copy number in the original sample, which is more copies of the virus.


Ct
36
35
34
33
32
31
30
29
C
T
Treatment

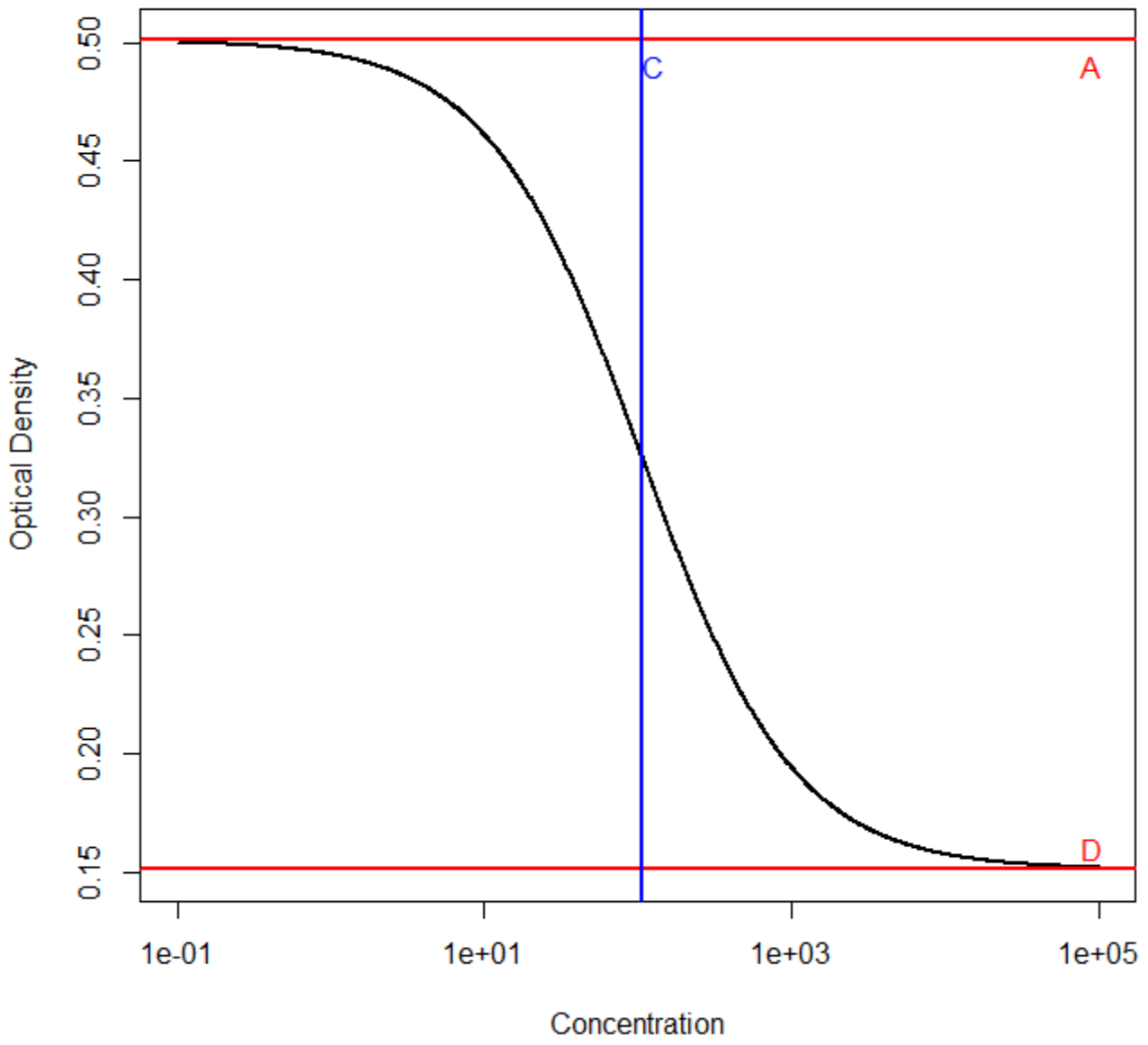
```
> anova(lm(Ct ~ Treat))
Analysis of Variance Table

Response: Ct
          Df Sum Sq Mean Sq F value  Pr(>F)
Treat      1 24.064 24.0642  12.361 0.01258 *
Residuals  6 11.681  1.9468
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

ELISA

- Enzyme-linked immuno sorbent assay (ELISA) is a test that uses antibodies and color change to identify a substance.
- Requires an antibody to the analyte
- Depending on the assay, binding results either in increased color or decreased color.
- Readout is through optical density or brightness at a particular frequency
- This is one standard method for coronavirus antibodies (uses antibodies to the antibodies).

ELISA

- The performance of the test depends on the antibody to the antibody
- It may bind to molecules other than the intended target
- It may bind well or poorly to the intended target
- Different tests can have very different performance characteristics


Optical Density
0.15
0.20
0.25
0.30
0.35
0.40
0.45
0.50
C
A
D
1e-01
1e+01
1e+03
1e+05
Concentration

Four-Parameter (Log-)Logistic Curve

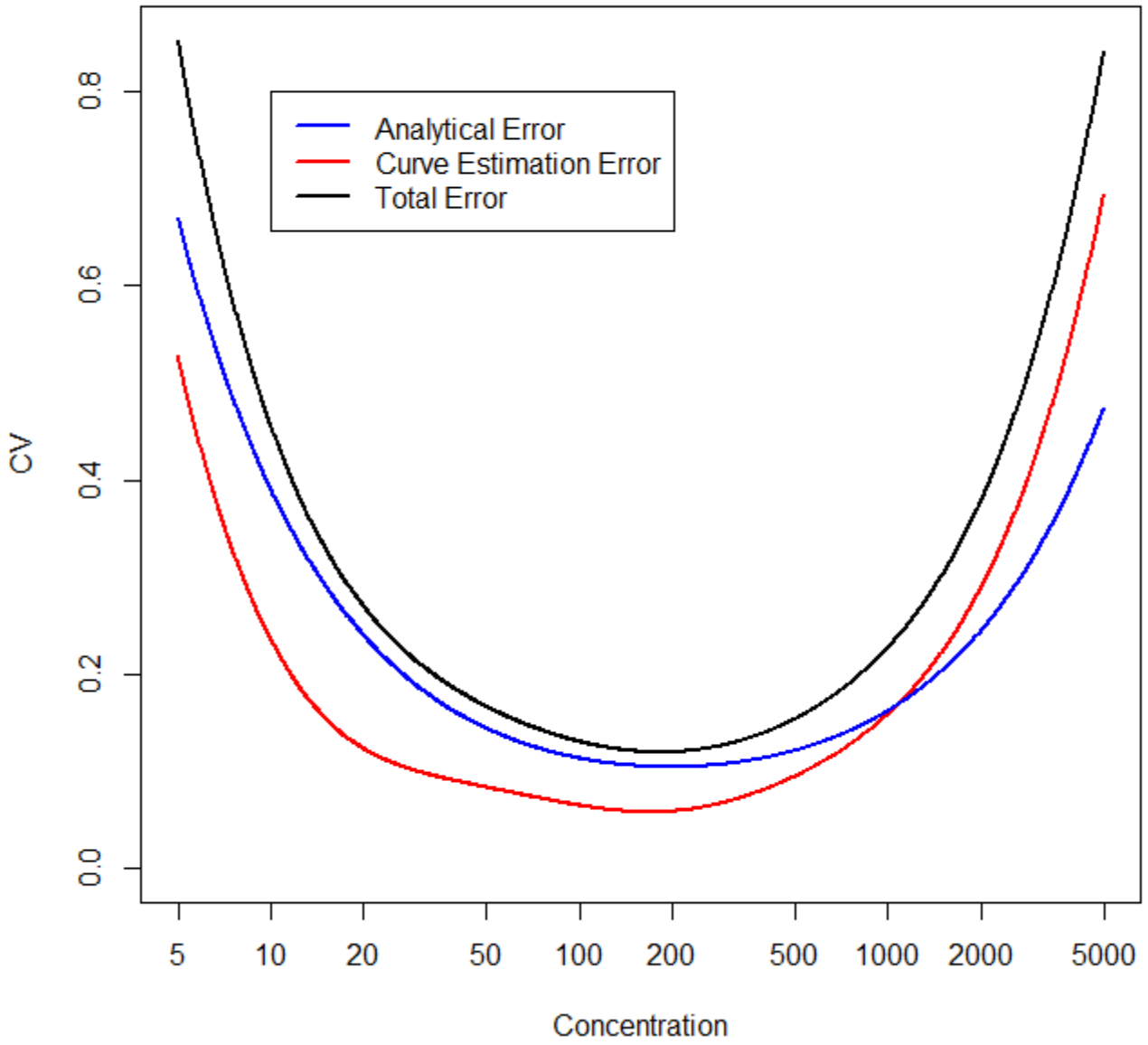
$$y = f(x) = \frac{A - D}{1 + (x / C)^B} + D$$

$$\frac{y - D}{A - D} = \frac{1}{1 + (x / C)^B}$$

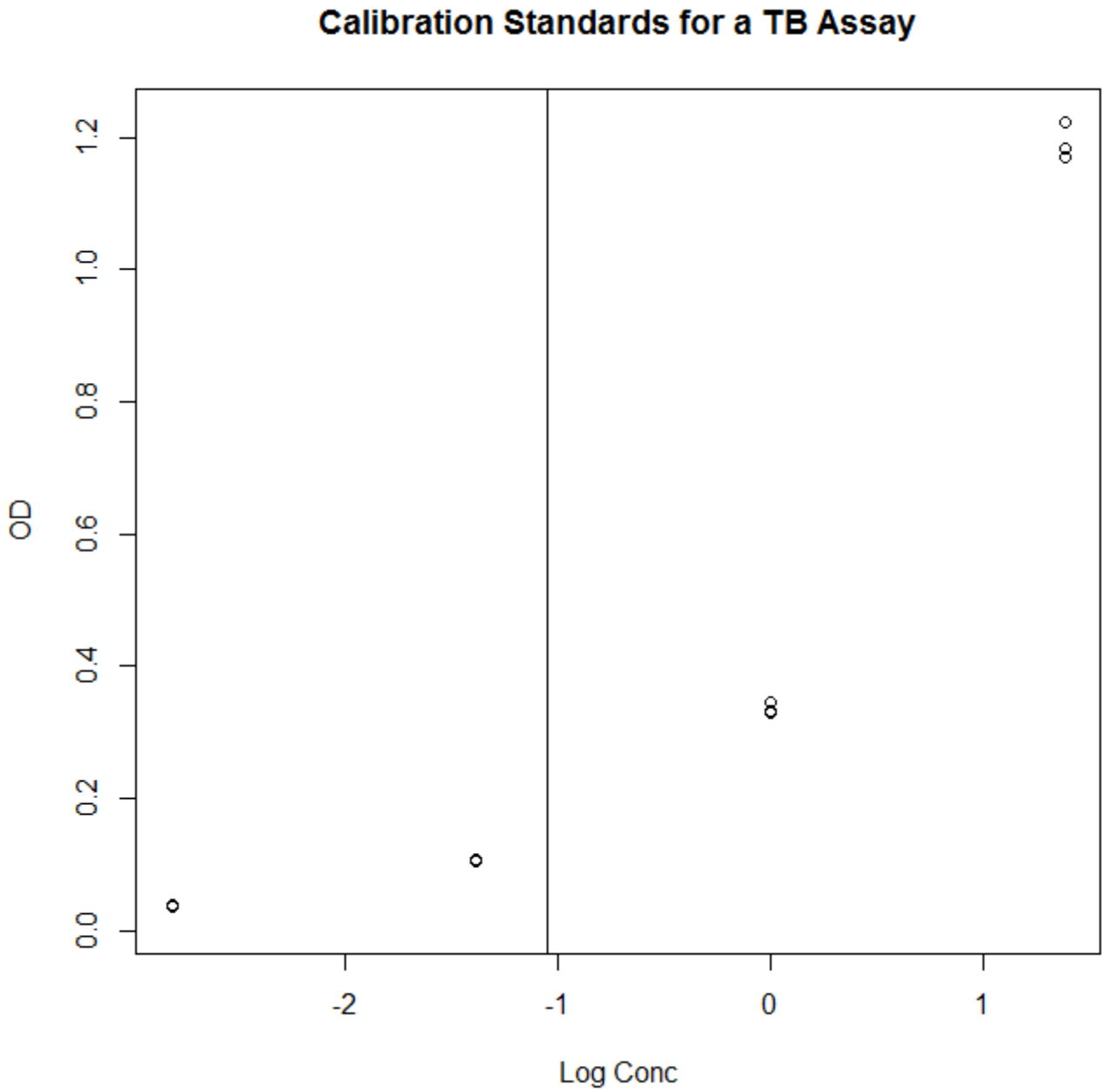
$$\frac{A - D}{y - D} - 1 = (x / C)^B$$

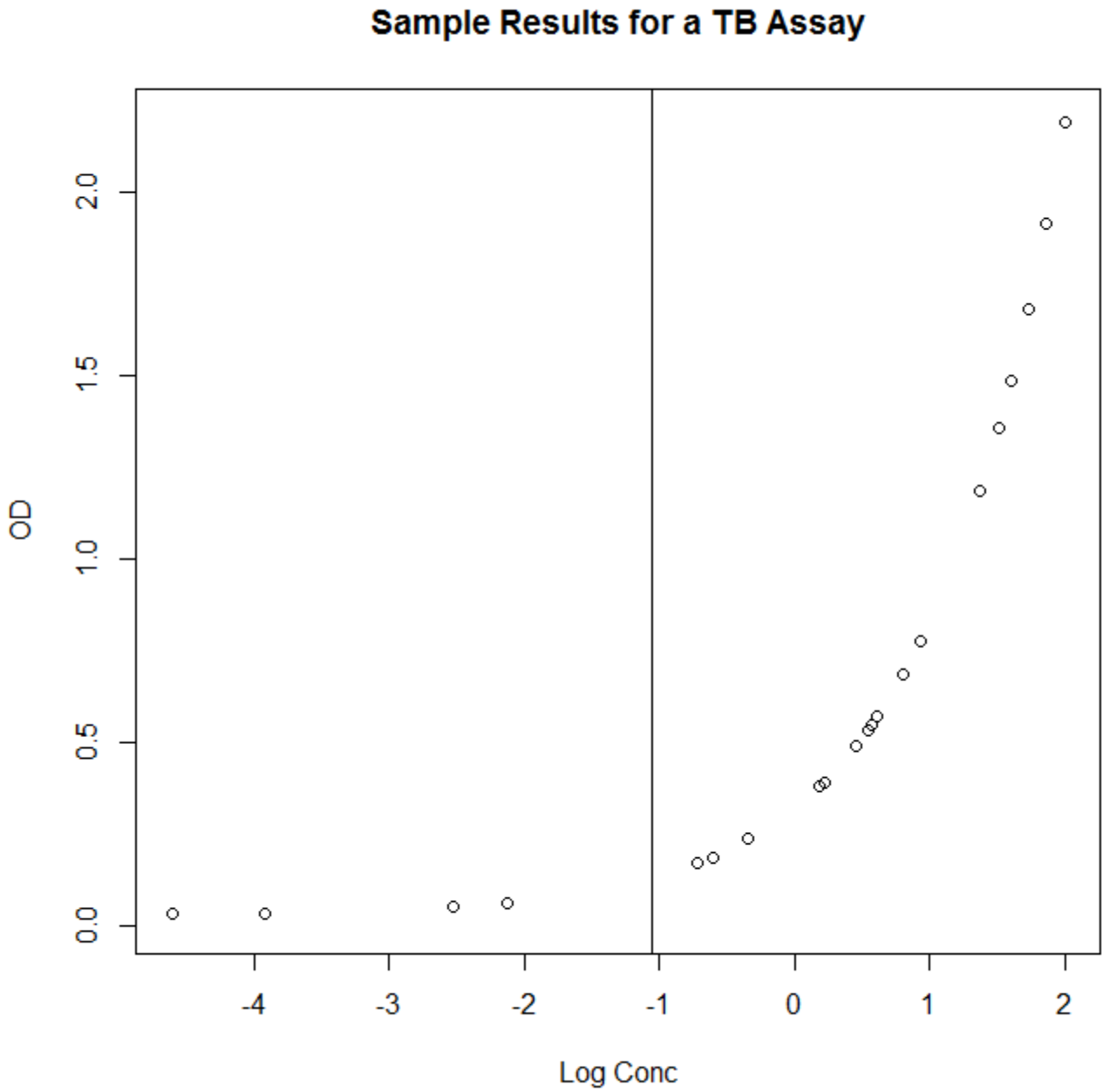
$$\ln\left(\frac{A - y}{y - D}\right) = B[\ln(x) - \ln(C)]$$

- ELISAs are often run on 96-well plates. Calibration data of known concentrations in some of the wells so that the parameters of the 4-parameter log-logistic curve can be estimated.
- OD values above A are interpreted as an estimated concentration of 0. OD values below D are out of bounds high.
- In some forms of ELISA, the OD is low for low concentrations and high for higher ones.
- The assay is generally good in a region around the center, not so good at the ends


Analytical Error
Curve Estimation Error
Total Error
CV
0.0
0.2
0.4
0.6
0.8
5
10
20
50
100
200
500
1000
2000
5000
Concentration

- If the concentration is too low, then the assay will not yield useful results.
- If the concentration is too high, then a dilution of the sample can fall into the accurate range.
- Often, a dilution series is run.


Calibration Standards for a TB Assay
OD
0.0
0.2
0.4
0.6
0.8
1.0
1.2
-2
-1
0
1
Log Conc


Sample Results for a TB Assay
OD
0.0
0.5
1.0
1.5
2.0
-4
-3
-2
-1
0
1
2
Log Conc