On the Determination Coefficient and Global Adjustment Test of Goodness of Fit

HIV infection :: $CD4^+$ T cells dynamics

Cecilia Azevedo; Rita Rb-Silva; Ana Horta; Margarida Correia-Neves

Centre of Mathematics, School of Sciences & Life and Health Sciences Research Institute, School of Medicine.

University of Minho

Table of contents

Introduction

$CD4^+$ T Cells

T cells are regulators of the immune system. $CD4^+$ T cells are also known as helper T cells because they act on other cells of the immune system to promote various aspects of the immune response. They have a crucial role in the bodys immune system.

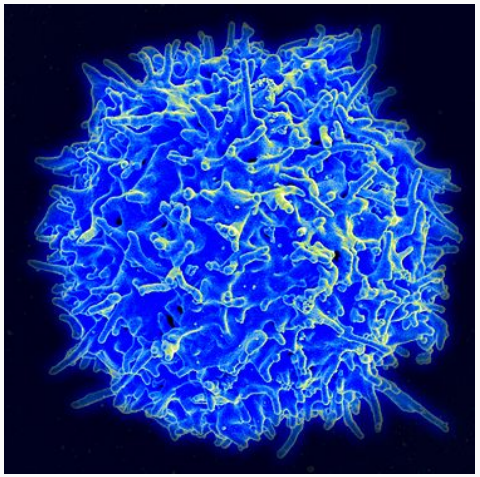


Figure 1: Human $CD4^+$ T Cell. *Image courtesy of NIAID*

HIV - Human immunodeficiency virus

When HIV virus enters in the bloodstream of a person it mainly attacks $CD4^+$ T cells, thought its CD4 protein (located in the border of the cell, Figure 2).

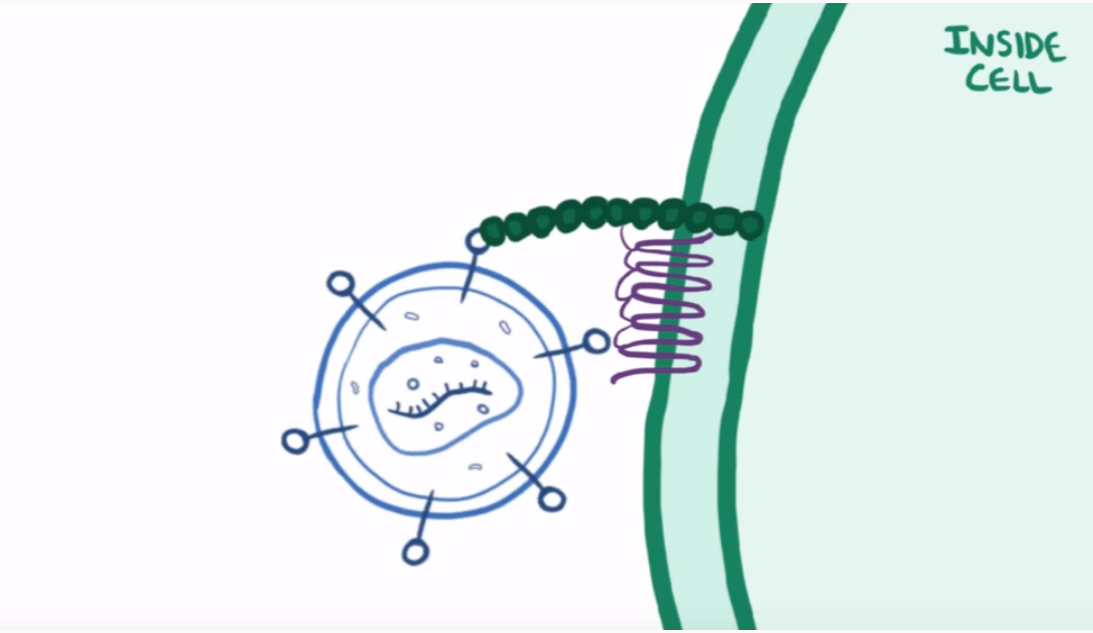


Figure 2: Membrane fusion

HIV - Human immunodeficiency virus

The HIV virus is a retrovirus as its genome is a single-stranded RNA and include in their replication cycle DNA intermediates (Figure 3).

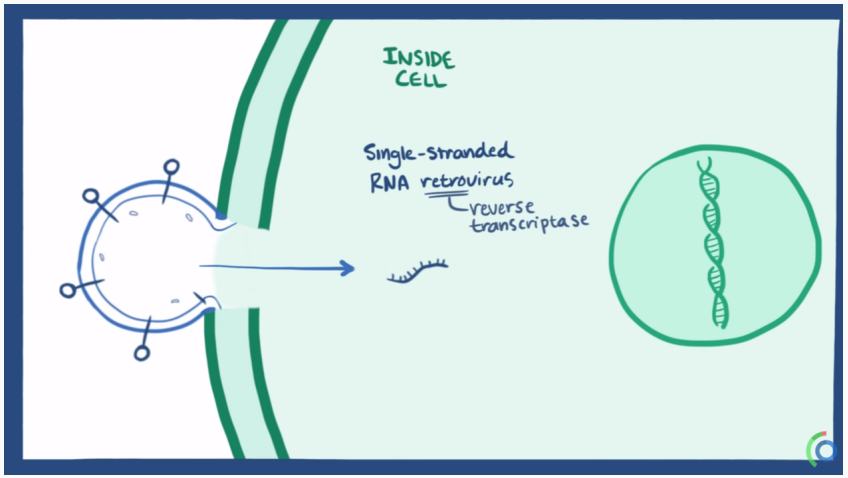


Figure 3: Viral RNA in the host cell cytoplasm

The virus uses its own reverse transcriptase enzyme to produce DNA from its RNA genome (Figure 4).

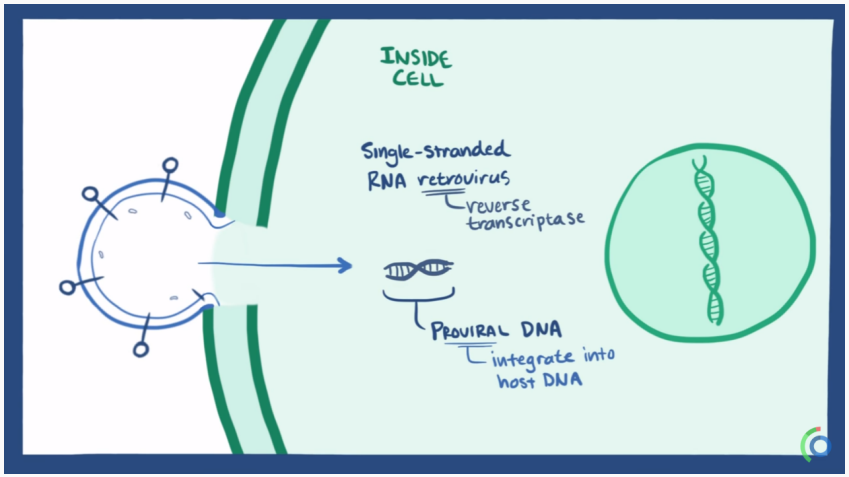


Figure 4: Viral DNA

That DNA is incorporated into the host cell genome who treats the viral DNA as part of its own genome, translating and transcribing the viral genes and producing the proteins required to assemble new copies of the virus (Figure 5).

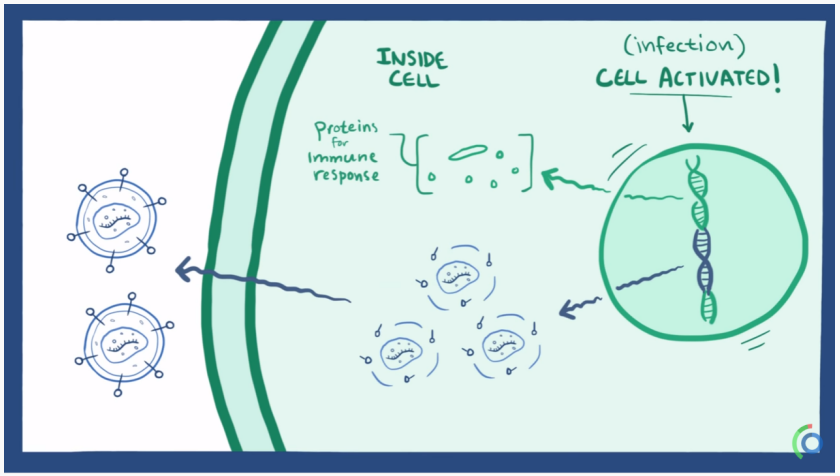


Figure 5: New copies of HIV

Number of CD4$^+$ T cells

The HIV replication results in massive depletion of CD4$^+$ T cells in the human blood. After this acute phase, several immune responses are generated by the host and viremia declines until a lower state level. However these responses fail to eliminate the virus leading to a chronic infection in most individuals during an asymptomatic period which can go on several years. During the chronic phase of the infection, blood CD4$^+$ T cell count declines slowly (see Figure 6).

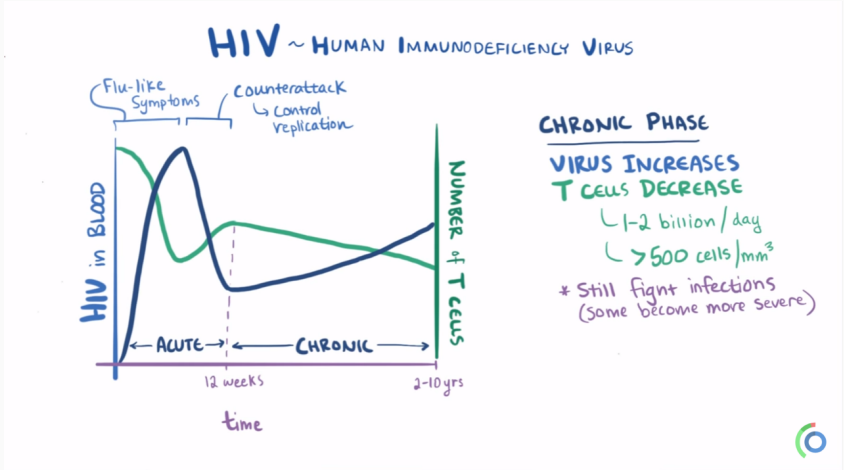


Figure 6: Schematic diagram of the course of HIV-1 infection

Motivation and data set

::The problem::

AntiRetroviral Treatment (ART) dramatically suppresses viral replication and reduces the plasma HIV-1 viral load to below the limits of detection of the most sensitive clinical assays (<50 RNA copies/mL).

ART causes a progressive reconstitution of the immune system.

That must imply an **increase in circulating CD4$^+$ T-lymphocytes**.

The mathematical problem includes

growth significance :: test if T-lymphocytes growth during ART is statistically significant;

growth magnitude :: quantify globally the *increase in circulating CD4$^+$ T-lymphocytes* during ART.

:: The data ::

We have $CD4^+$ T cells counts of 50 patients (10 female) with HIV-1 and in the same ART program. The counts were taken at baseline (BL) and at 1, 3, 6, 9, 12, 16, 20, 24, 28, 32, 36, 42, 48, 54, and 60 months during ART. In this problem we are considering data for male and female together because there is no statistical evidence that this two groups are different concerning to $CD4^+$ T cells counts at the different time points, age, or even body mass index (BMI at BL).

Table 1: Characteristics of the patients in the beginning of ART

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	
CD4 BL	9.0	116.0	262.5	247.1	320.5	1033.0	186.44
Age	22.00	32.75	41.00	40.55	46.25	67.00	10.24
BMI	16.23	20.75	23.05	23.45	25.82	33.61	3.487

Statistical methodology

There are an enormous variability in CD4 count data. This variability is within subjects and between subjects (as usual in this type of human data). In Figure 7, we have the data for the different subjects in separate panels.

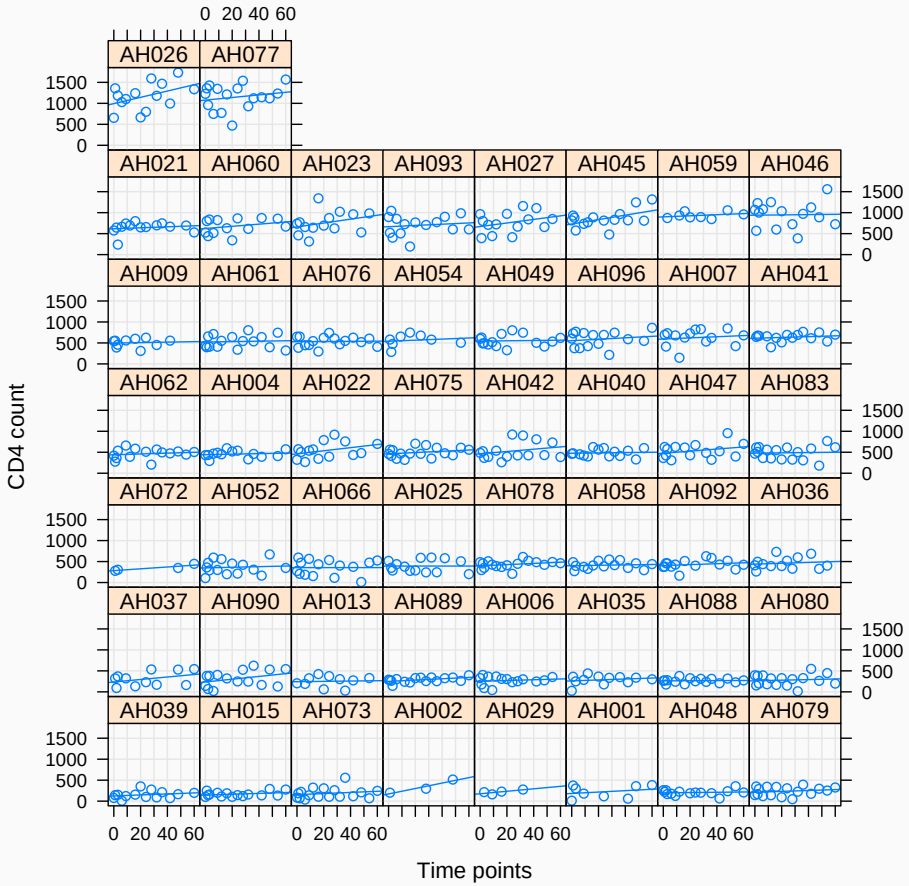


Figure 7: Scatter diagram

A Linear Model

Globally $CD4^+$ counts grow over time, see Figure 8.

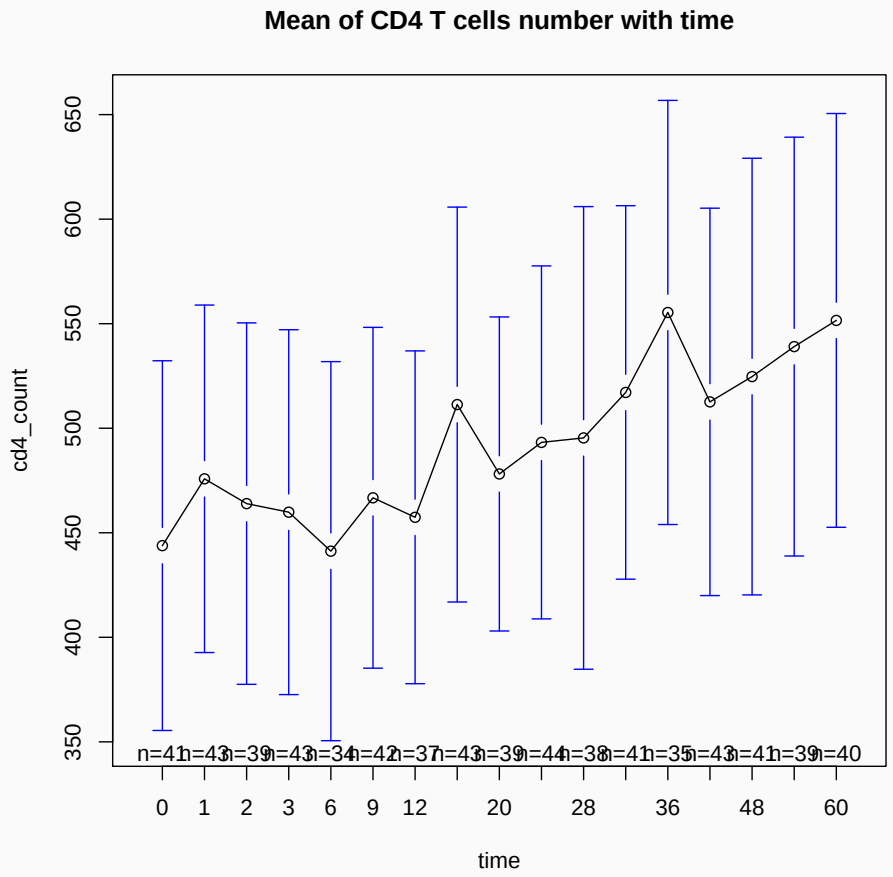
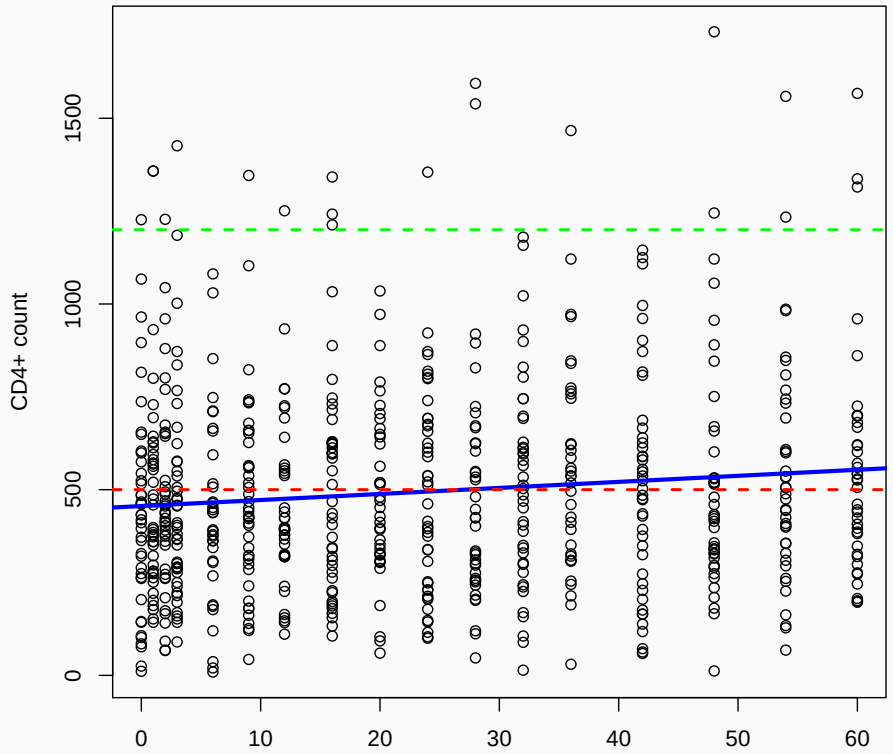


Figure 8: Global mean growth of number of T cells with time

Results

Coefficient of Determination

If we use all data (all counts available for the 50 patients), we obtain the equation $y = 455.887 + 1.628x$ for least-squares regression line and this linear model has a coefficient of determination very low, R^2=0.012, so more than 98% of the total variability on data is not explained by this model, see Figure 9. The model explains nothing!!!



Utility of the model

The model presented is an **useless** model?

Global Adjustment Test based on β (the slope parameter)

$H_0 : \beta = 0$ (useless model) vs $H_1 : \beta \neq 0$.

The test statistic is

$$T = \frac{\widehat{\beta}}{\widehat{\sigma}_{\widehat{\beta}}} = \frac{\widehat{\beta}}{\sqrt{\frac{MSE}{s_x^2}}}\sqrt{n-2} \sim t_{n-1}$$

and the p-value $= 2 \times \left(1 - \Phi(n-1, |T_{\text{obs}}|)\right)$, where Φ is the df of t_{n-2} and $|T_{\text{obs}}|$ is the absolute value of the observed test statistic T. In our case we have the estimative of the parameter, $\widehat{\beta} = 1.63$ and a low standard error associated $\widehat{\sigma}_{\widehat{\beta}} = 0.572$ which implies a statistic value T_{obs}=2.842.
In this case we have 680 degrees of freedom (50×17-168-2), so the p value is 0.0046, and we reject the useless of the model.

Discussion

Conclusions

The low p-value of the Global Test of Adjustment

is a consequence of the large number of observations.

For medicine is a realistic model. There are a global (constant) mean growth of 20 T cells per year of ART. As far as we know is consensual that there are a global growth of this kind of cells during ART but there was not a quantification of the growth.
Here we quantify the global growth but this model should not be used to make predictions.

This result allows answering a part of the question:

if there is growth, it is a slow one. It should not be forgotten that only the positive balance is being considered. It has to be taken into account growth that prevents decay.

to be continued . . .

THANK YOU VERY MUCH FOR YOUR ATTENTION !!!!