Analysis of Time Varying Data Using Mathematical Models

Sarah Holte Ph.D.

Fred Hutchinson Cancer Research Center

Outline

- Brief overview of differential equations and examples of use of differential equations in medicine and biology

Outline

- Brief overview of differential equations and examples of use of differential equations in medicine and biology
- The use formal statistical methods when using differential equations to analyze observed data.

Outline

- Brief overview of differential equations and examples of use of differential equations in medicine and biology
- The use formal statistical methods when using differential equations to analyze observed data.
- Examples of mathematical models in of HIV research

  Estimation of dynamic parameters;
  Failure of prediction of long term outcomes;
  Formal comparison of two models

Data where the mean structure is of interest

- Often statistical analysis involves evaluating the variability in a set of data in order to

  Compare mean of two or more groups;

  Obtain confidence intervals (precision) for an estimated mean

Data where the mean structure is of interest

- Often statistical analysis involves evaluating the variability in a set of data in order to

  Compare mean of two or more groups;
  Obtain confidence intervals (precision) for an estimated mean

- When longitudinal data exhibit a complex nonlinear pattern, describing the time-varying mean of the data is of interest.

Data where the mean structure is of interest

- Often statistical analysis involves evaluating the variability in a set of data in order to

  Compare mean of two or more groups;
  Obtain confidence intervals (precision) for an estimated mean

- When longitudinal data exhibit a complex nonlinear pattern, describing the time-varying mean of the data is of interest.

- If an underlying mechanism which generates the data is known or hypothesized, a mathematical model can be developed which describes the time-varying mean structure.

Differential equations

An ordinary differential equation (ODE) is an equation whose solution is a function of time. The differential equation relates various derivatives (wrt time) of the function to each other through the equation. Solving the ODE means finding the function that satisfies it

Example

$$\frac{dX(t)}{dt} = -\delta X(t),\ X(0) = X_0$$

Differential equations

An ordinary differential equation (ODE) is an equation whose solution is a function of time. The differential equation relates various derivatives (wrt time) of the function to each other through the equation. Solving the ODE means finding the function that satisfies it

Example

$$\frac{dX(t)}{dt} = -\delta X(t),\ X(0) = X_0$$

has solution

$$X(t) = X_0 e^{-\delta t}$$

Differential equations

An ordinary differential equation (ODE) is an equation whose solution is a function of time. The differential equation relates various derivatives (wrt time) of the function to each other through the equation. Solving the ODE means finding the function that satisfies it

Example

$$\frac{dX(t)}{dt} = -\delta X(t),\ X(0) = X_0$$

has solution

$$X(t) = X_0 e^{-\delta t}$$

which can be checked

$$\frac{dX(t)}{dt} = X_0 e^{-\delta t} * (-\delta) = -\delta X, \quad X(0) = X_0 e^0 = X_0$$

Differential equations in medicine and biology

- The most frequent use of ODE's in medicine and biology is in pharmacokinetics/dynamics to describe drug concentrations over time.

Differential equations in medicine and biology

- The most frequent use of ODE's in medicine and biology is in pharmacokinetics/dynamics to describe drug concentrations over time.
- In recent years, ODE models have been used to study the dynamics of a HIV viral load within an infected patient.

Differential equations in medicine and biology

- The most frequent use of ODE's in medicine and biology is in pharmacokinetics/dynamics to describe drug concentrations over time.
- In recent years, ODE models have been used to study the dynamics of a HIV viral load within an infected patient.
- The metabolism of environmental toxins is sometimes described using ODEs: Physiokinetic models.

Differential equations in medicine and biology

- The most frequent use of ODE's in medicine and biology is in pharmacokinetics/dynamics to describe drug concentrations over time.
- In recent years, ODE models have been used to study the dynamics of a HIV viral load within an infected patient.
- The metabolism of environmental toxins is sometimes described using ODEs: Physiokinetic models.
- Tumor growth (the size of tumors over time) has been modeled using ODEs.

Relationship between the model and data

Throughout the rest of the presentation, the time-varying data is assumed to have mean described by one (or more) compartments of a system of ODE's.

$$Y_t \sim N(X(t,\theta), \sigma^2)$$

where $X(t,\theta)$ is the solution to a system of ODEs and θ are parameters in ODEs.

Usually, parameter estimation based on observed data is the analysis of interest, and is conducted by minimizing

$$min_\theta \sum_{i=1}^{n} (X(t,\theta) - Y_{t_i})^2$$

This method is often referred to as 'nonlinear least squares regression'.

Why use a model?

Why use a model?

- To describe a nonlinear time-varying mean structure.

Why use a model?

- To describe a nonlinear time-varying mean structure.
- To relate a parameter in an unobserved compartment to a compartment which is observed

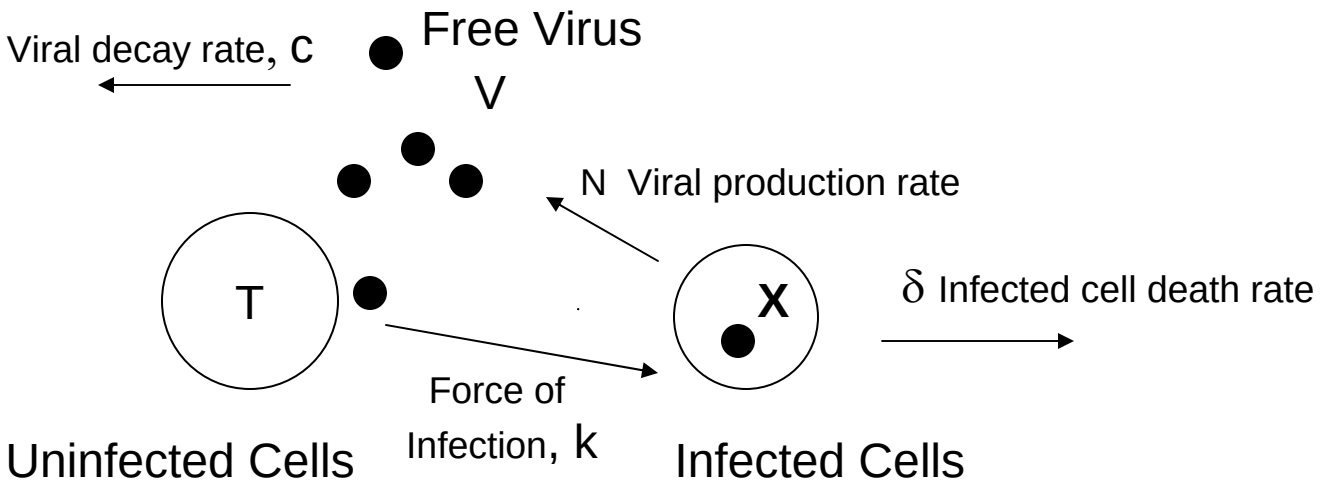
Why use a model?

- To describe a nonlinear time-varying mean structure.
- To relate a parameter in an unobserved compartment to a compartment which is observed
- Example: In HIV infected patients, Viral load is generally observed but infected cells are not generally observed, however the rate at which infected cells decay or die is of scientific and clinical interest.

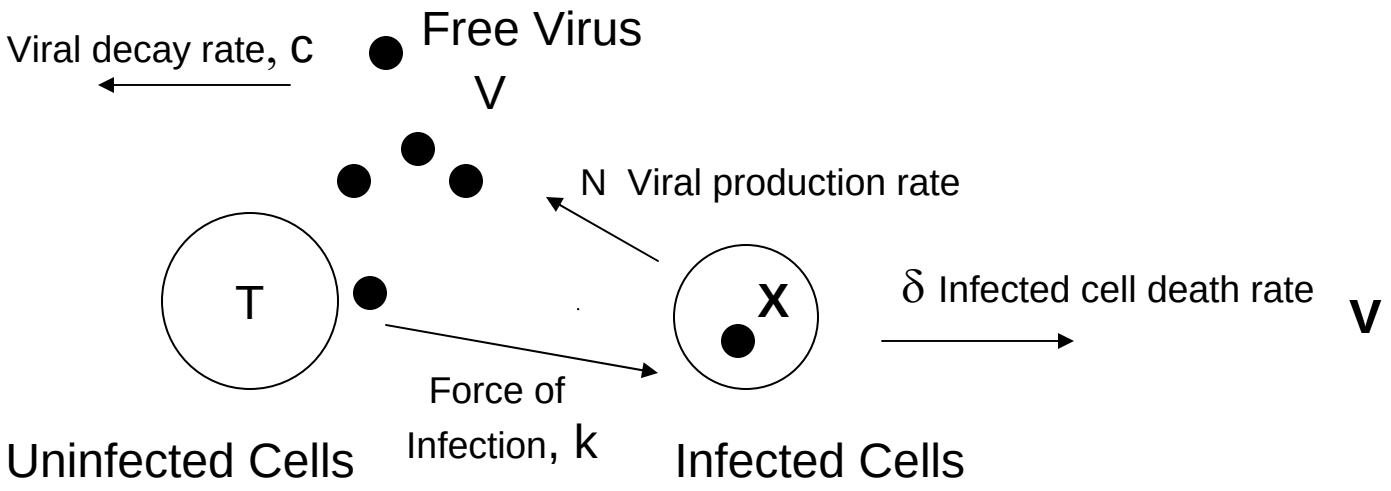
Why use a model?

- To describe a nonlinear time-varying mean structure.
- To relate a parameter in an unobserved compartment to a compartment which is observed
- Example: In HIV infected patients, Viral load is generally observed but infected cells are not generally observed, however the rate at which infected cells decay or die is of scientific and clinical interest.
- The model provides the relationship between observed viral load data and infected cell death rate, so that this rate can be estimated from data on observed viral load.

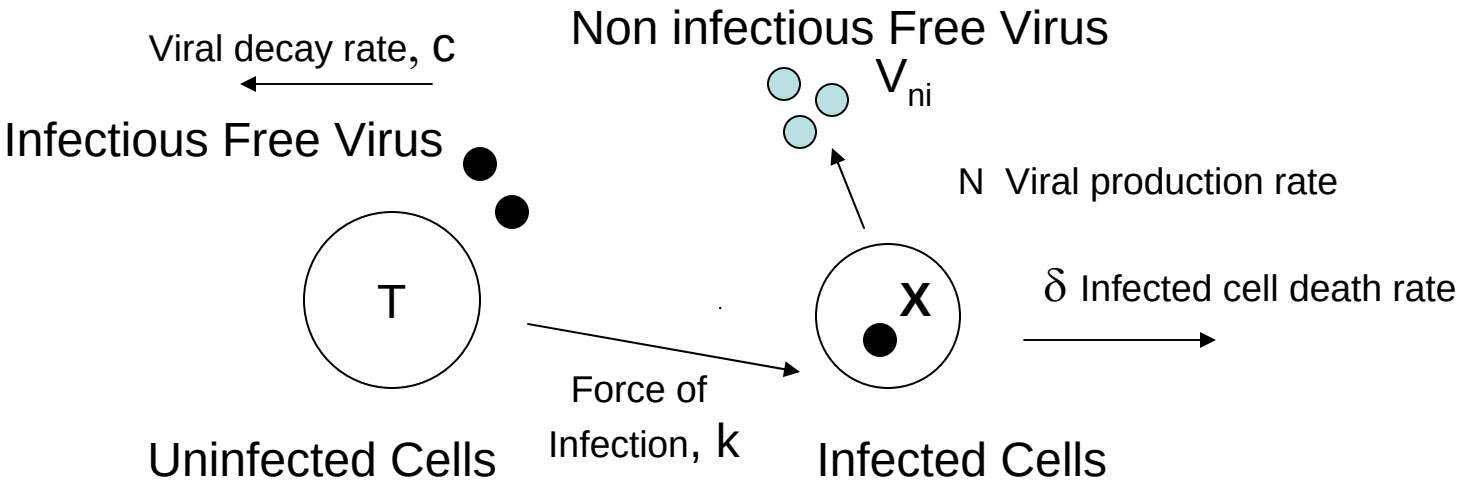
Pre-treatment Uninfected cell, Infected cell, and Virus,replication



Pre-treatment Uninfected cell, Infected cell, and Virus,replication



Post-treatment Uninfected cell, Infected cell, and Virus,replication



Model equations for viral dynamics before treatment

Perelson et.al. Science 1996

$$\begin{aligned} \frac{dX}{dt} &= kTV - \delta X \\ \frac{dV}{dt} &= N\delta X - cV \end{aligned}$$

has "states" which vary with time,

- $X(t)$ = population of infected cells at time t
- $V(t)$ = the population of viral RNA at time t

and "parameters" which are (generally) constant

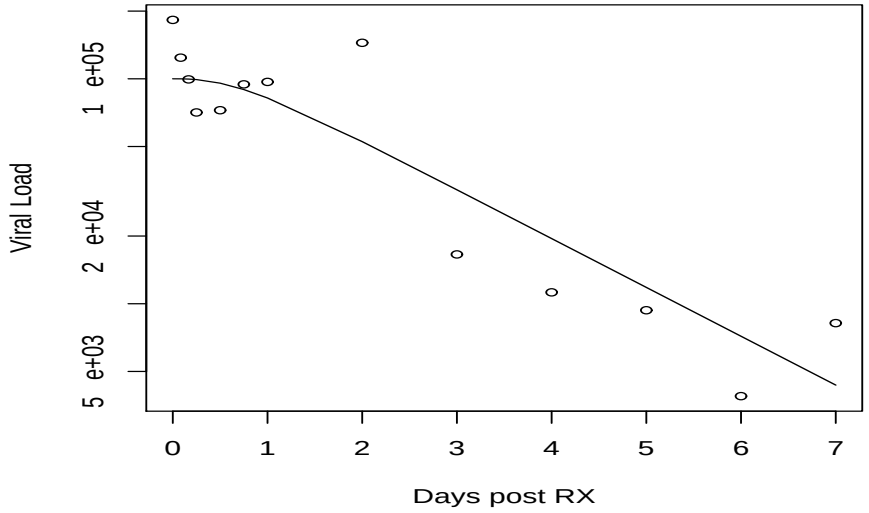
- k = infection rate, δ = infected cells death rate.
- N = the viral production, C viral clearance.

Model equations for viral decay after treatment

$$\begin{aligned}
\frac{dX}{dt} &= kTV_I - \delta X \\
\frac{dV_I}{dt} &= -cV_I \\
\frac{dV_{NI}}{dt} &= N\delta X - cV_{NI}
\end{aligned}$$

- V_I is the population of infectious virus (produced before treatment)
- V_{NI} is the population of non-infectious virus (produced after treatment)
- $V = V_I + V_{NI}$ is the total **observed** viral load
- Observed data, $(t_i, log(v_i)) \sim N(log(V(t, C, delta)), \sigma^2)$

Simulated viral load up to 7 days post treatment



Use nonlinear least squares (or other statistical methods) and model solution for $V(t, C, \delta)$ to estimate C and δ from time series data.

Conclusions based estimates of viral decay and infected cell clearance rates

Conclusions based estimates of viral decay and infected cell clearance rates

- Since HIV infection often lasts many years before the onset of AIDS, it was believed that the virus was relatively dormant during this time.

Conclusions based estimates of viral decay and infected cell clearance rates

- Since HIV infection often lasts many years before the onset of AIDS, it was believed that the virus was relatively dormant during this time.
- Estimates of δ and C showed that viral clearance and infected cell turnover were occurring at a much more rapid rate than previously believed.

Conclusions based estimates of viral decay and infected cell clearance rates

- Since HIV infection often lasts many years before the onset of AIDS, it was believed that the virus was relatively dormant during this time.
- Estimates of δ and C showed that viral clearance and infected cell turnover were occurring at a much more rapid rate than previously believed.
- Rapid viral clearance implies that virus is being produced rapidly and continuously and partially explains why HIV can mutate so rapidly.

Predictions of long term outcomes and model comparison

Bi-phasic viral decay, more than one infected cell compartment produces virus

Perelson et.al. Nature 1997

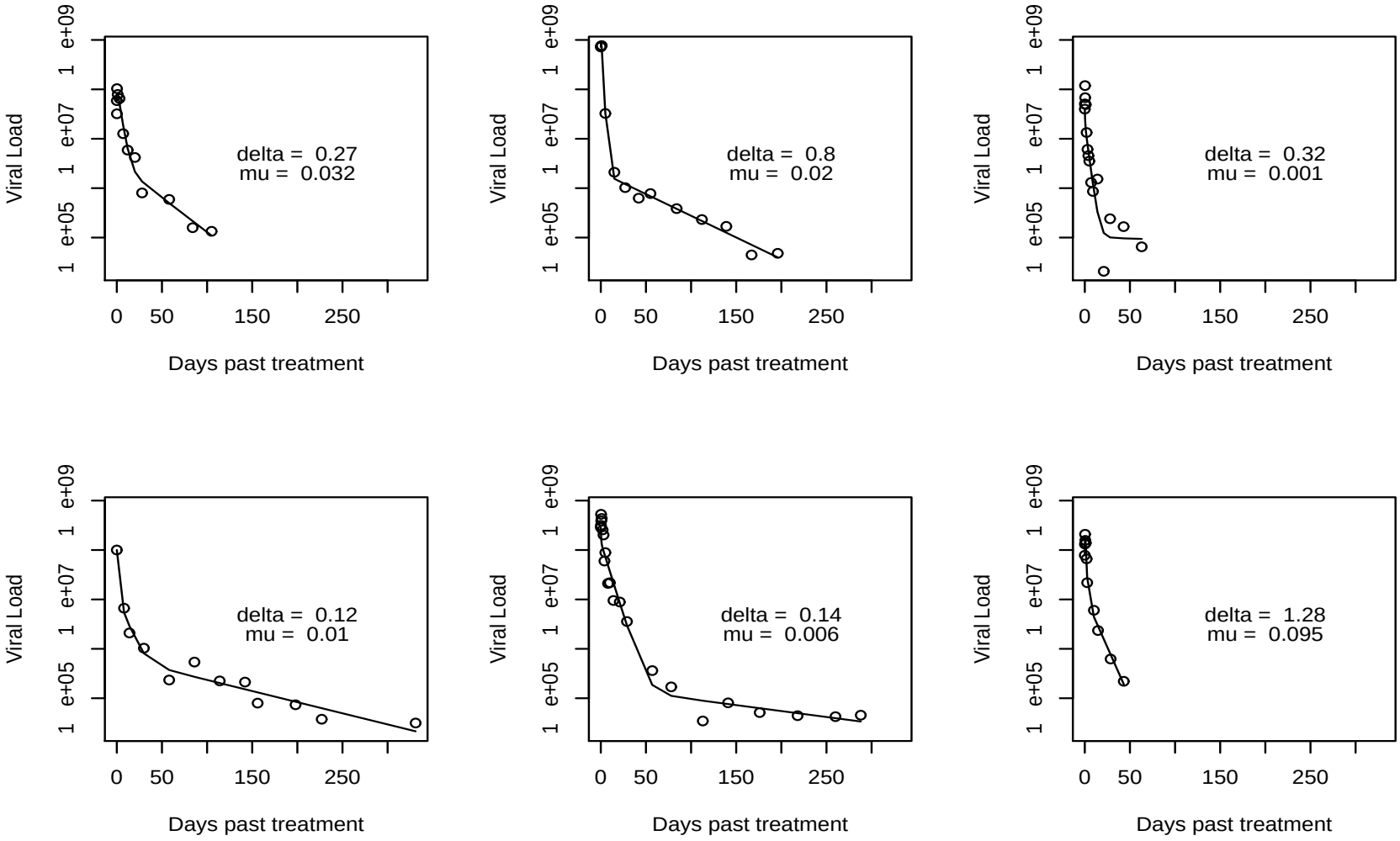
Constant Decay Model

$$\begin{aligned} \frac{dX}{dt} &= -\delta X \\ \frac{dY}{dt} &= -\mu Y \\ \frac{dV}{dt} &= p_x X + p_y Y - cV \end{aligned}$$

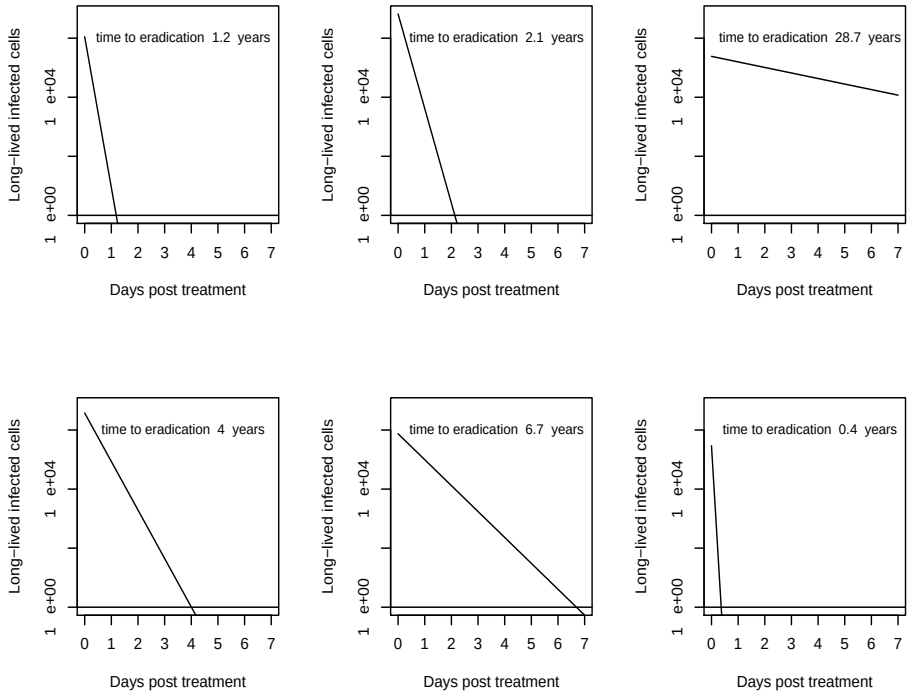
- X is the population of short lived infected cells
- Y is the population of longer lived infected cells
- V is cell free viral RNA

Viral decay after treatment in children

Constant decay model



Conclusions based on Constant Decay model



- Estimates of δ and μ using plasma viral load are obtained and used to estimate time on treatment of approximately 2 years to eradicate virus

Adjustment to the constant decay model

- As data from even longer periods of time post treatment became available it became apparent that virus was not eradicated even after many years of successful treatment.

Adjustment to the constant decay model

- As data from even longer periods of time post treatment became available it became apparent that virus was not eradicated even after many years of successful treatment.
- Studies with longer follow-up tended to provide slower second phase decay estimates.

Adjustment to the constant decay model

- As data from even longer periods of time post treatment became available it became apparent that virus was not eradicated even after many years of successful treatment.
- Studies with longer follow-up tended to provide slower second phase decay estimates.
- Consider an alternate model and make a formal statistical comparison to see if the constant decay model can be rejected

Alternative to the constant decay model

Holte et.al. JAIDS 2006

Density Dependant Decay Model

$$\begin{aligned} \frac{dX}{dt} &= -\delta X \\ \frac{dY}{dt} &= -\mu Y \\ \frac{dV}{dt} &= p_x X + p_y Y - cV \end{aligned}$$

- X is the population of short lived infected cells
- Y is the population of longer lived infected cells
- V is cell free viral RNA

Alternative to the constant decay model

Holte et.al. JAIDS 2006

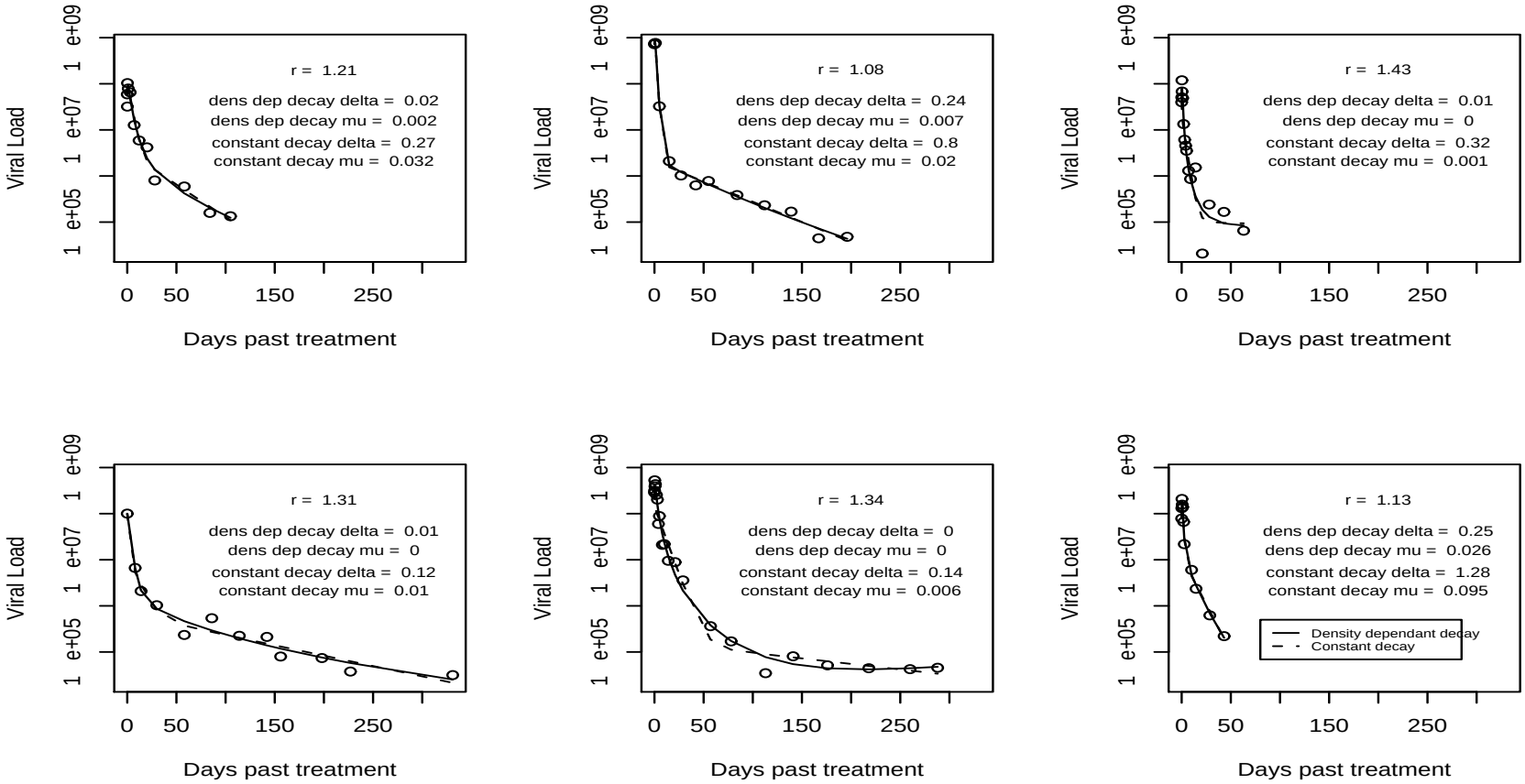
Density Dependant Decay Model

$$\begin{aligned} \frac{dX}{dt} &= -\delta X^{\mathbf{r}} \\ \frac{dY}{dt} &= -\mu Y^{\mathbf{r}} \\ \frac{dV}{dt} &= p_x X + p_y Y - cV \end{aligned}$$

- X is the population of short lived infected cells
- Y is the population of longer lived infected cells
- V is cell free viral RNA

Null hypothesis: Constant decay model is correct, $\mathbf{r = 1}$

Density Dependant Decay model results



Density Dependant Decay model results - Continued

- The parameter r is significantly greater than 1 for all but one child suggesting that that the constant decay model is not appropriate for the observed data.

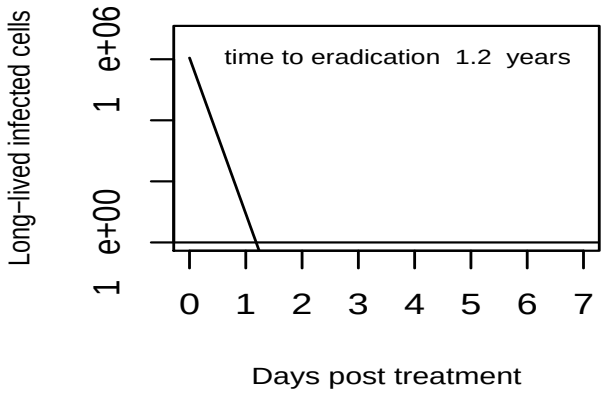
Density Dependant Decay model results - Continued

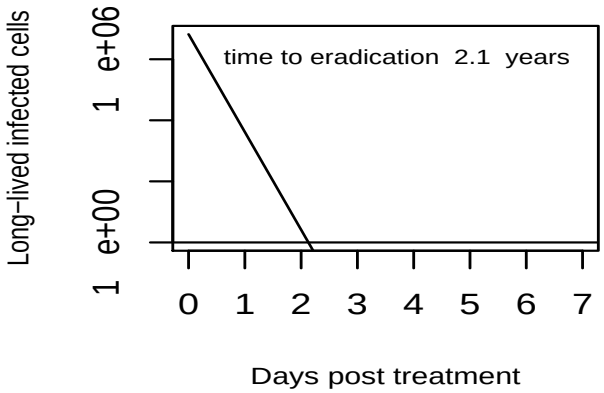
- The parameter r is significantly greater than 1 for all but one child suggesting that that the constant decay model is not appropriate for the observed data.
- Very different conclusions about the long term dynamics of viral load after treatment depending on which model is used for prediction and inference.

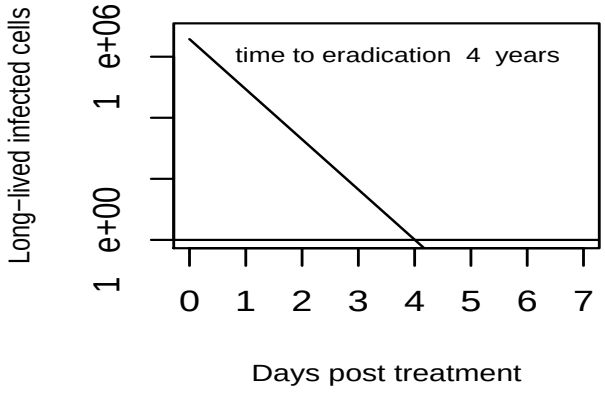
Density Dependant Decay model results - Continued

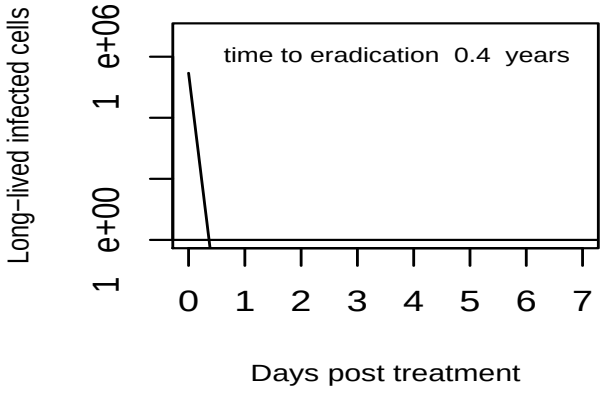
- The parameter r is significantly greater than 1 for all but one child suggesting that that the constant decay model is not appropriate for the observed data.
- Very different conclusions about the long term dynamics of viral load after treatment depending on which model is used for prediction and inference.
- This example demonstrates how very similar models can provide very different conclusions.

Density dependant vs constant decay model - time to eradication

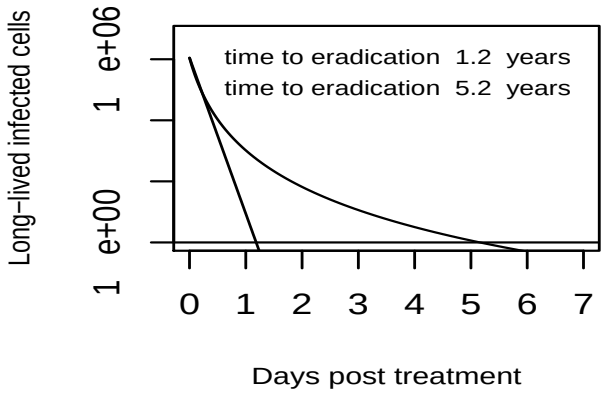


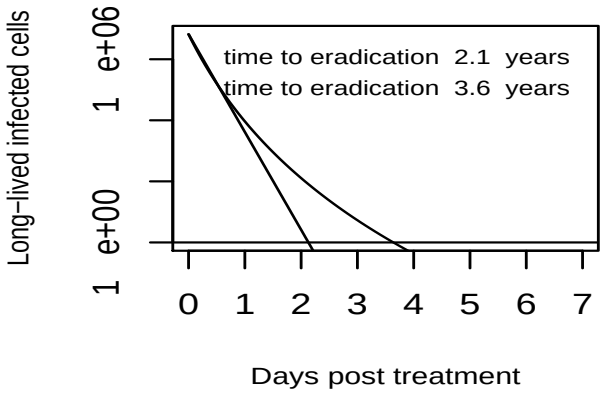


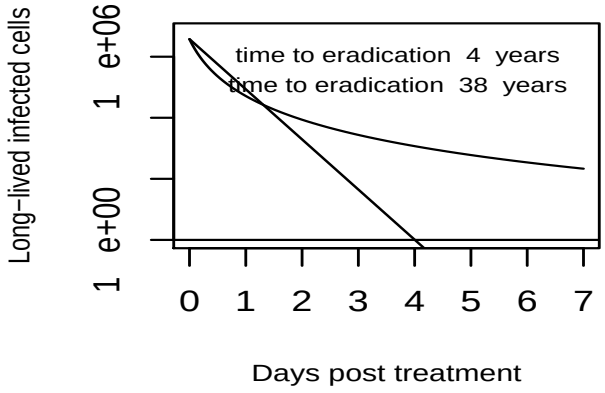


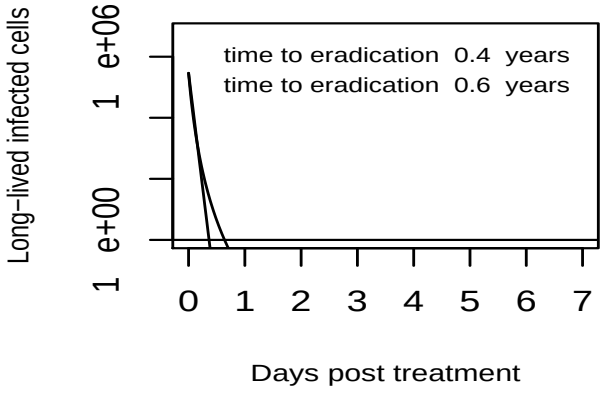


Density dependant vs constant decay model - time to eradication

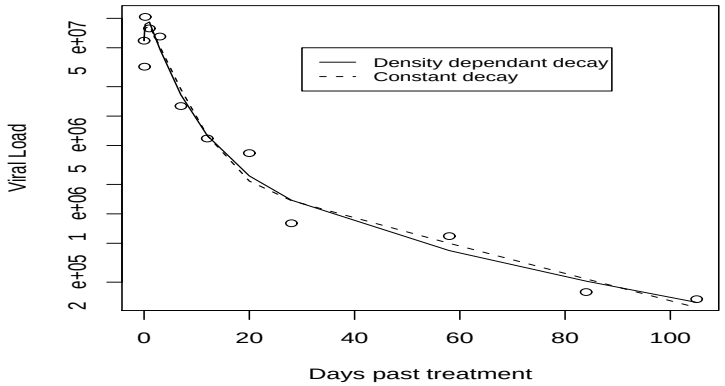


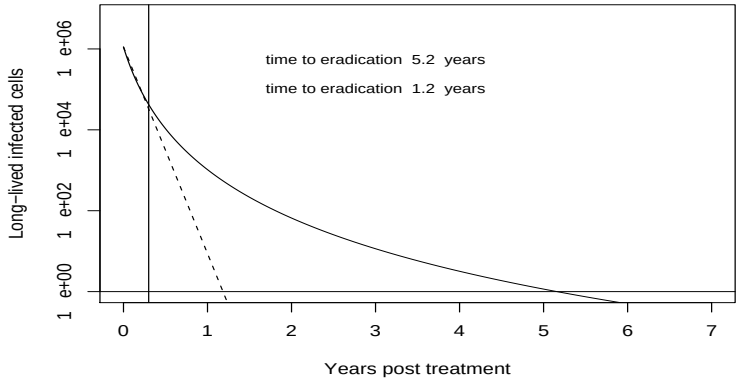






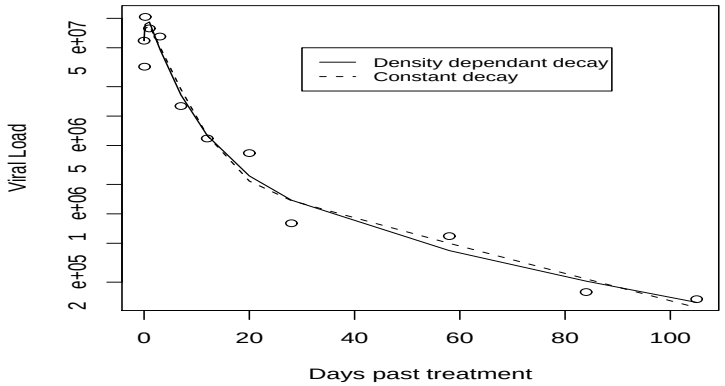
Conclusions based on models for viral decay after treatment

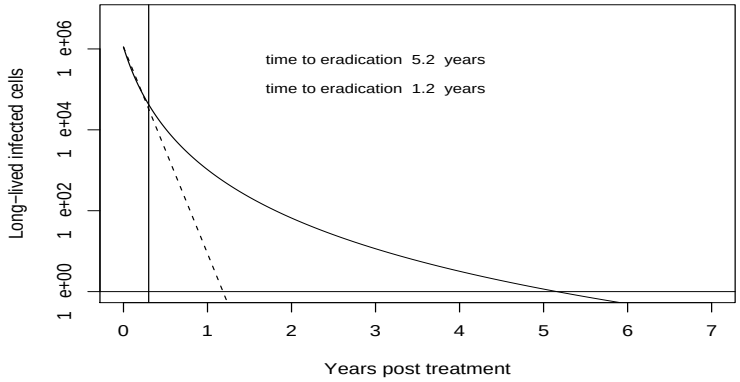




- Using models to make predictions is subject to the dangers of the potential for incorrect mathematical models....

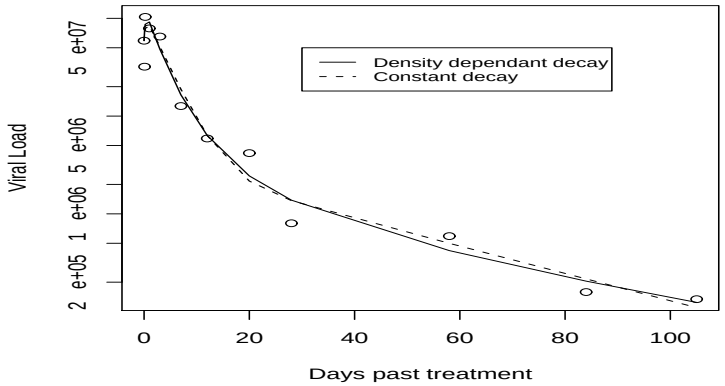
Conclusions based on models for viral decay after treatment

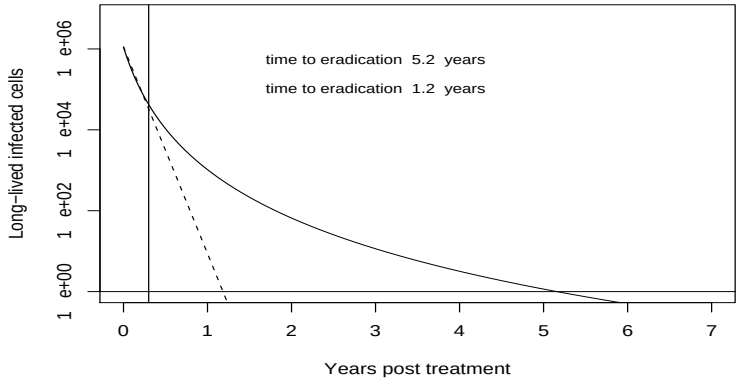




- Using models to make predictions is subject to the dangers of the potential for incorrect mathematical models....
- ... in addition to extrapolating beyond the range of observed data

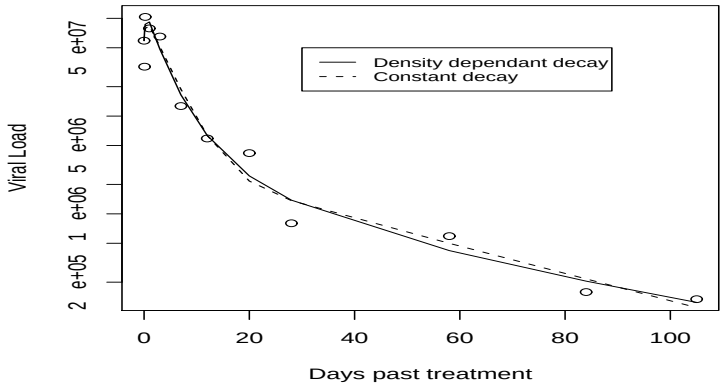
Conclusions based on models for viral decay after treatment

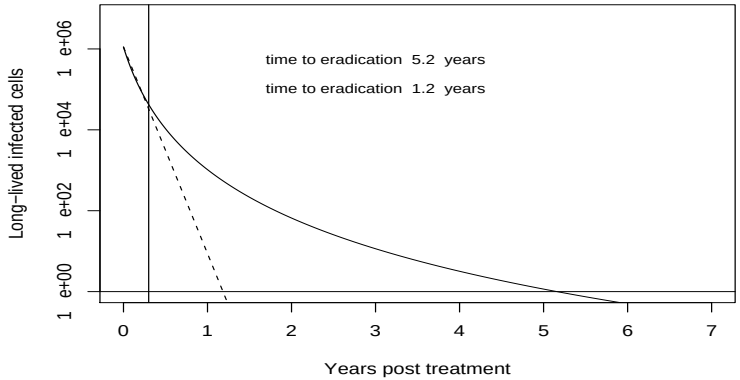




- Using models to make predictions is subject to the dangers of the potential for incorrect mathematical models....
- ... in addition to extrapolating beyond the range of observed data
- Analysis using density dependant decay model allow us to **reject the null hypothesis** that the constant decay model accurately describes the data.

Conclusions based on models for viral decay after treatment





- Using models to make predictions is subject to the dangers of the potential for incorrect mathematical models....
- ... in addition to extrapolating beyond the range of observed data
- Analysis using density dependant decay model allow us to **reject the null hypothesis** that the constant decay model accurately describes the data.
- Additional methods for model comparison are needed.

Conclusions

- Combing mathematical models to describe time-varying mean structure with formal statistical techniques for estimation and inference can provide insight to complex biological behaviors.
- However, models can be (and usually are) wrong over long periods of time so that extrapolating over long periods of time is dangerous.
- A good analyst will know the strengths and weaknesses of using mathematical models in the analysis of time-varying data.