The Lupus Project

Jim Ramsay
Michal Abrahamowicz

Jiguo Cao, Carlotta Fok, Wen Zhang

1. Overview

- Systemic Lupus Erythematosus (SLE): the disease
- e-ry-them'-a-to-sus = "rash"
- The lupus data base
- What we hope to achieve
- A differential equation model for a flare event
- Fitting the differential equation to the data
- Some results

2. Systemic Lupus Erythematosus (SLE)

- Lupus is an auto–immune disease in the same family as rheumatoid arthritis. The body's immune system attacks itself, producing a wide spectrum of symptoms and affecting many organs.
- These attacks, called *flares*, occur suddenly and unpredictably, last for varying periods, and then disappear, sometimes for long periods.
- "Erythematosus" means reddening, referring to a characteristic skin rash by which it was first identified. Early accounts compared the rash to a wolf bite, hence the name.
- The disease is incurable.

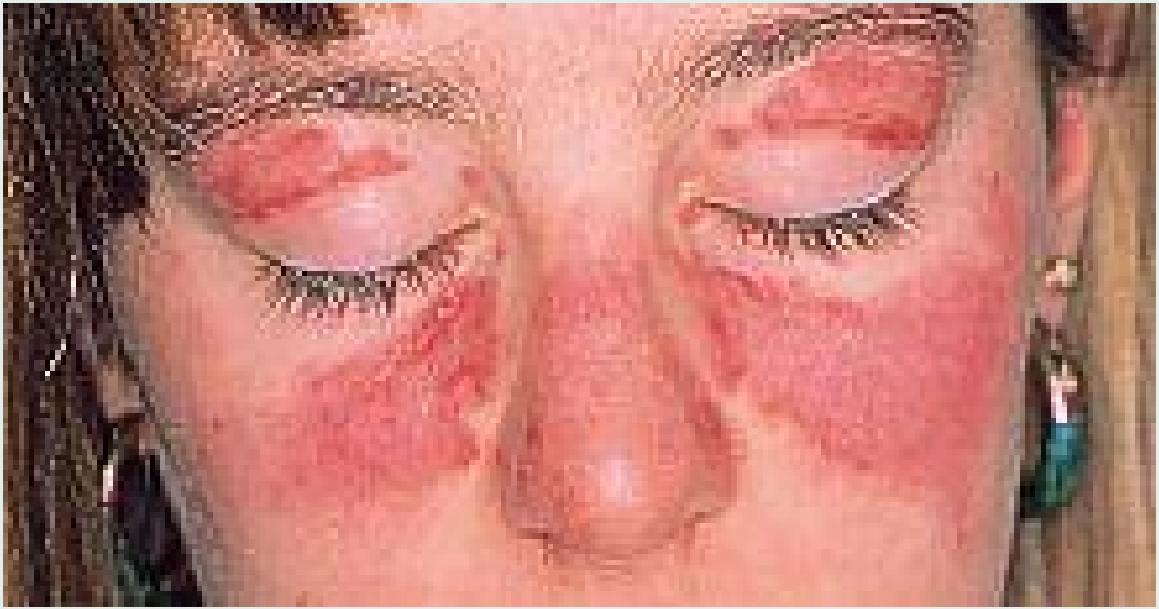
Who gets it?

- 9 times as many women as men
- blacks and some Asian groups more susceptible
- incidence ranges from 3 to 400 per 100,000
- lupus can appear at any age, and the earlier it appears, the more severe it tends to be
- lupus is on the increase, and in some places is now more common than rheumatoid arthritis
- genetic, environmental, and hormonal factors are all involved
- exposures to chemicals and ultra-violet light are suspects

What are the symptoms?

- symptoms range from mild to severe, and can cause permanent damage or be fatal
- a rash on the face and chest, pain and swelling in the joints and fatigue are common and early signs of a flare
- the kidneys are often affected, with swelling and loss of function, and end stage renal failure is a real risk
- the heart, arteries, lungs, eyes and central nervous system may also be involved
- the psychological effects of lupus are receiving more and more attention
- a typical flare goes from just noticeable to acute in the order of ten days or less

A typical facial rash



How is lupus treated?

- the variation in the nature and severity of symptoms combined with the unpredictability of flares makes treating this disease a huge challenge
- mild symptoms are treated with anti–inflammatory drugs (aspirin, etc.)
- more severe symptoms require the use of corticosteroids, usually *prednisone*
- response time to an increase in prednisone dose is usually rapid, and of the order of a few days
- corticosteroids are toxic if taken over long periods at high doses
- high dose levels must be tapered down gradually; sudden decreases can trigger a new flare
- patients are assessed at regular intervals

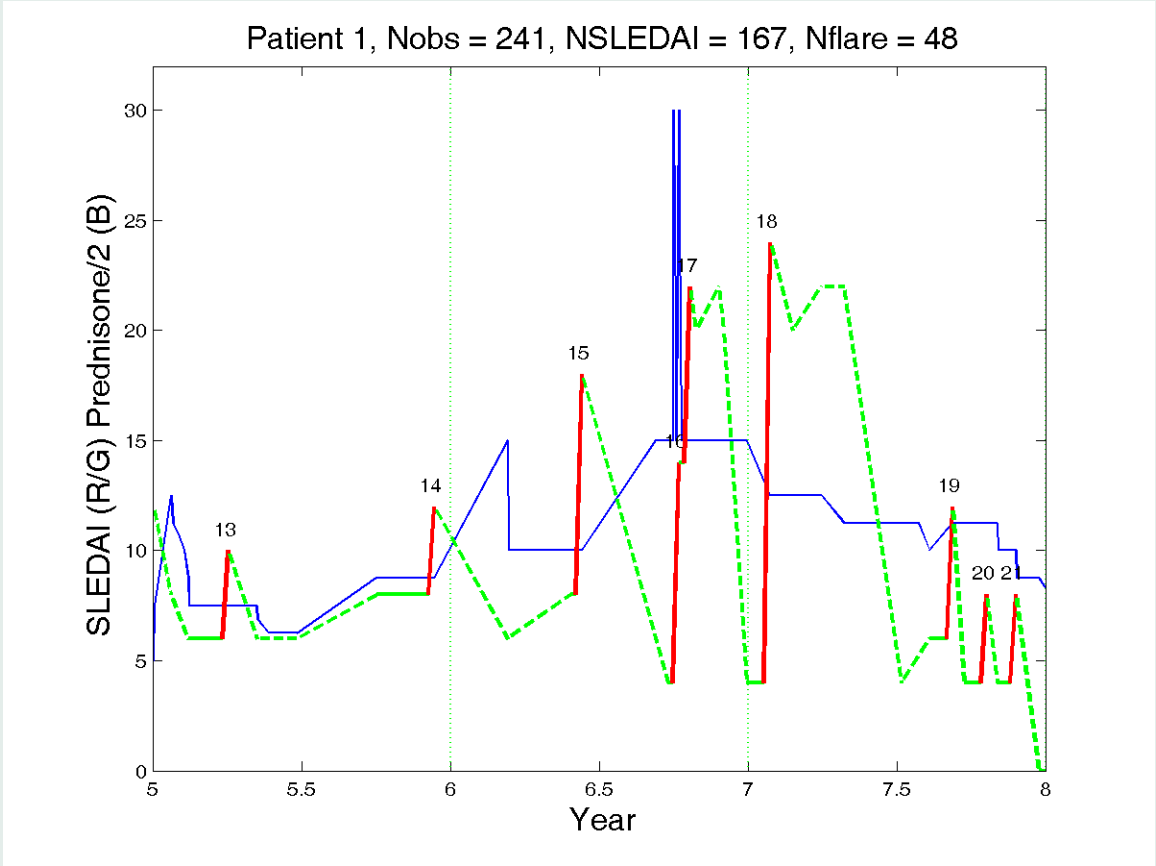
How is lupus severity assessed?

- although lupus symptoms are multidimensional, long term treatment requires some overall measure of disease severity
- a number of scales have been proposed
- the SLEDAI scale is widely used
- SLEDAI is a check list of 24 symptoms, each given a numerical weight
- SLEDAI item weights range from 1 for fever to 8 for seizures
- a flare has been defined as a SLEDAI score increase of 3 or more to a level of 8 or higher
- during flares SLEDAI scores of 25 to 30 are common

3. Our lupus database

- A joint McGill/University of Toronto team headed by Dr. Paul Fortin has complete histories for about 300 patients spanning, in many cases, around 20 years.
- This is one of the largest and highest quality set of patient records in the world.
- Our statistical investigations of these data are funded by the Canadian Institute for Health Research (CIHR).

A record segment



4. The statistical challenges

We require a model for:

- flare timings (a point process)
- flare intensities (a marked point process)
- flare durations (a marked interval process)
- flare dynamics: rate of onset and rate of recovery
- how flare characteristics depend on prednisone level and
- prednisone dynamics or rate of change
- individual differences in all of the above

It is fundamental to model the dynamics of the disease.

Data issues

- the SLEDAI scale score has limited reliability
- the dates at which these scores are assessed are themselves haphazard
- some data may be actually missing, eg: does SLEDAI = 0 always mean “no symptoms”?

We can, however, work closely with the physicians who work with these patients to identify flare characteristics, including flare onset times, flare durations, and to answer some questions.
For example, a SLEDAI score may not change, but the fact that prednisone was increased at that point suggests that the disease has nonetheless become acute.
We can also return to patient records to retrieve other information as required.

5. A simple model for flare dynamics

Let $u(t)$ be an indicator function for when lupus is in its active state and a flare is taking place.

- $u(t)$ takes only values 0 and 1.
- The times t_i at which $u(t)$ becomes positive can be estimated directly from the data, and therefore assumed known.
- The duration of an active state will be δ, and may vary from flare to flare.

We model symptom level $s(t)$ as a first order differential equation:

$$Ds(t) = -\beta(t)s(t) + \alpha(t)u(t)$$

What roles do $\beta(t)$ and $\alpha(t)$ play?

Consider the constant coefficient case:

$$Ds(t) = -\beta s(t) + \alpha u(t)$$

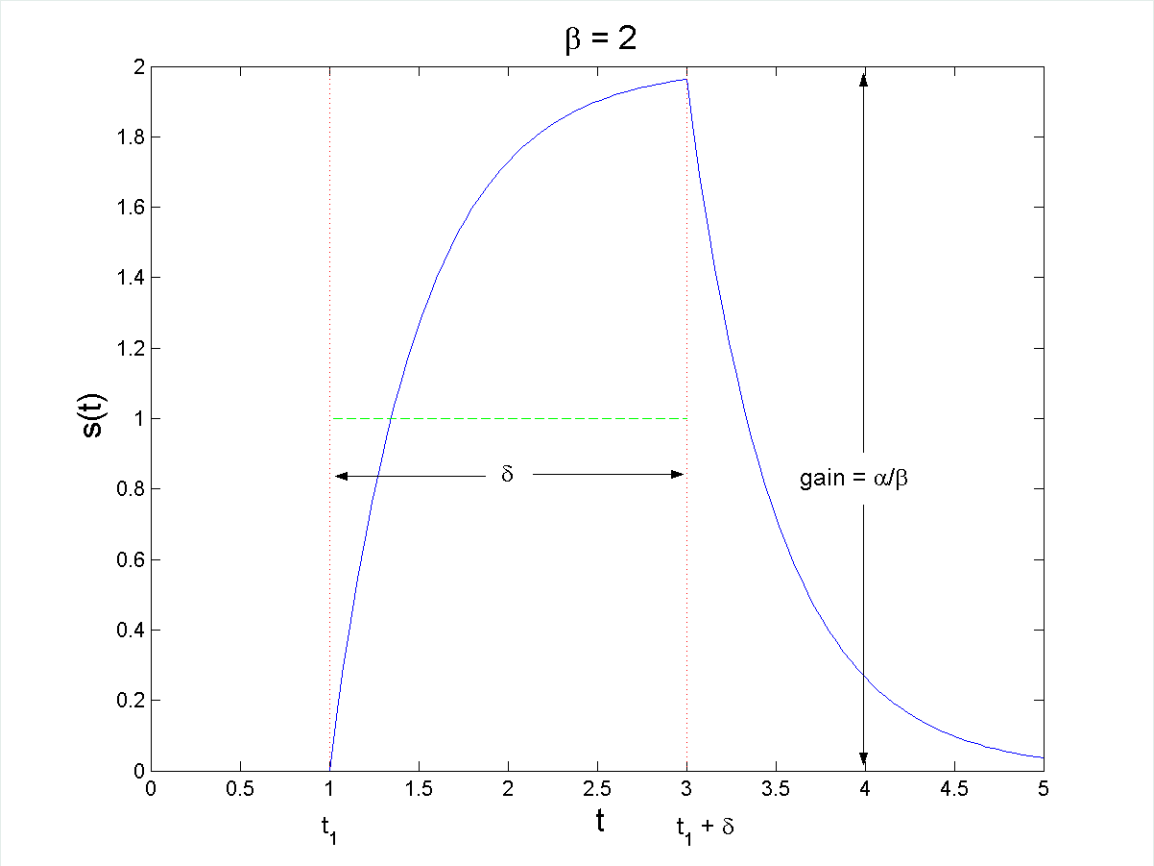
Let $s(0) = 0$, and $u(t)$ switch "on" at time t_1.
The equation has the general solution

$$s(t) = s(0)e^{-\beta t} + \alpha \int_0^t e^{-\beta(t-v)} u(v) dv$$

and for our $s(0)$ and $u(t)$, this becomes

$$s(t) = \begin{array}{ll} 0, & 0 \le t < t_1 \\ \frac{\alpha}{\beta}[1 - e^{-\beta(t-t_1)}], & t_1 \le t < t_1 + \delta \\ \frac{\alpha}{\beta}[1 - e^{-\beta\delta}]e^{-\beta(t-t_1-\delta)}, & t_1 + \delta \le t \end{array}$$

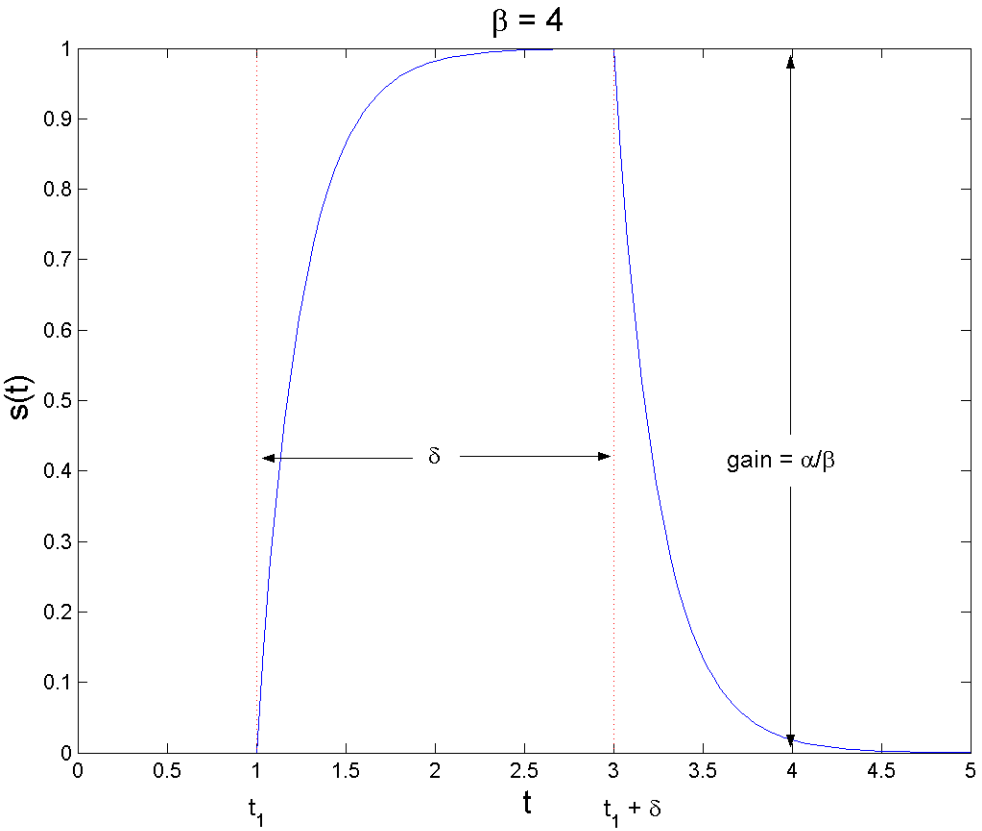
An example, $s(0) = 0, \alpha = 4$



What we see

- the maximum height, called the “gain,” is α/β
- β determines how rapidly the curve rises to this height, and how rapidly the curves decays to 0 when $u(t)$ switches off. It is the “dynamics” parameter.
- to see how this works, let $\beta = 4$

Another example, $s(0) = 0, \alpha = 4$



- Now, when we double β, we see a faster response.
- When $u(t)$ switches on, the curve will go from 0 to near the gain, α/β, in $4/\beta$ time units.
- Of course the gain can be held constant by increasing α.
- Think of a radio:
  - α as the volume control
  - β is the treble/bass control
- Engineers often parameterize the differential equation like this:

$$Ds(t) = -\frac{1}{\tau}s(t) + \frac{K}{\tau}u(t)$$

  where $\tau = 1/\beta$ is a time measure and K is the gain in SLEDAI units per unit increase in $u(t)$.

Isn’t this picture too simple?

- A lupus flare doesn’t look like $s(t)$ in these figures. The rise in symptoms is much more rapid than their decline.
- We can imagine that the disease also affects the body’s capacity to respond to the disease itself, as well as it’s capacity to recover.
- That is, $\beta(t)$ is also affected by the disease.
- When the patient is healthy between flares, $\beta(t)$ is high, leading to rapid response to the onset of the disease.
- When the patient is experiencing a flare, $\beta(t)$ is near zero, implying a slow recovery.

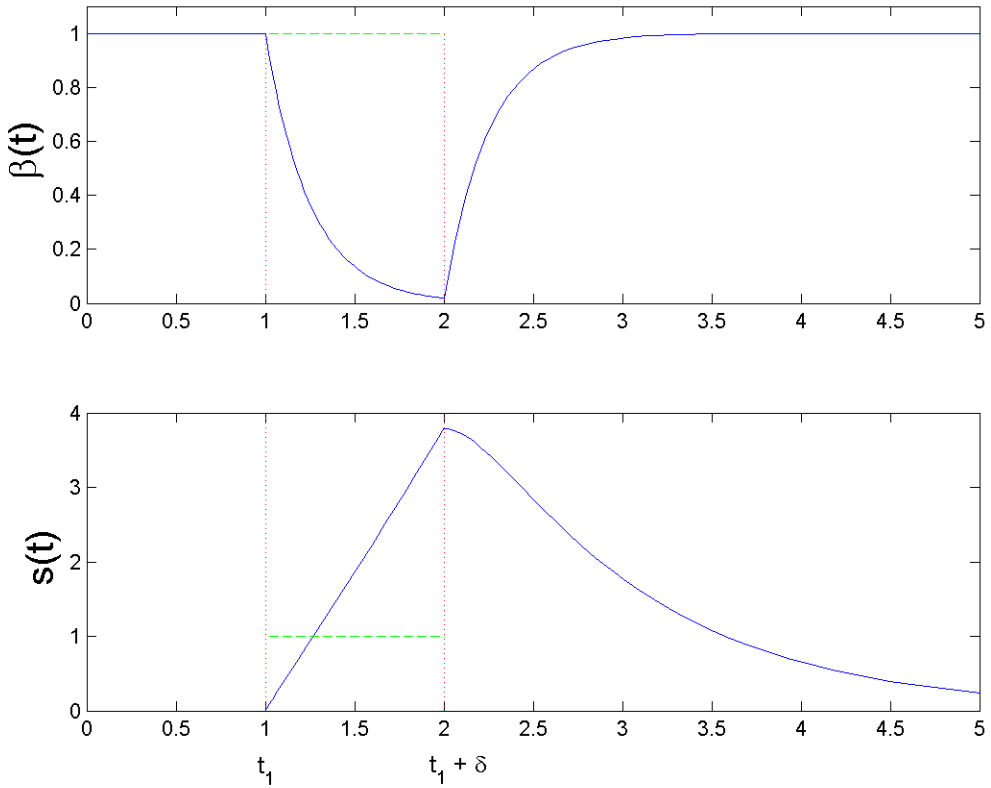
A differential equation for $\beta(t)$

- We tried this differential equation for β

$$D\beta(t) = -\gamma\beta(t) + \theta[1 - u(t)]$$

- When $u(t)$ switches on, $\beta(t)$ decays to zero, and $Ds(t)$ tends to equal $\alpha u(t)$; that is, $s(t)$ increases linearly while $u(t) = 1$.
- When $u(t)$ switches off, $\beta(t)$ returns to the level of its gain, θ/γ, and $s(t)$ tends to decay exponentially with rate equal to $\beta(t)$'s gain.

An example, $\alpha = \gamma = \theta = 4, \delta = 1$



- Now we have the general shape of a lupus flare.
- The increase in symptoms is essentially linear because $\beta(t)$ decays rapidly to 0 inside a flare.
- When $\beta(t) \approx 0$, the gain becomes $\alpha\delta$.
- the intensity of the flare is then proportional to δ.

6. Fitting the differential equation to data

- We use *profiled least squares estimation*.
- In *profiled* estimation, we have two sets of parameters.
- Let's call them ϕ and $\boldsymbol{\theta}$.
- We're primarily interested in estimating $\boldsymbol{\theta}$.
- We make the nuisance parameters ϕ functions of $\boldsymbol{\theta}$, that is, $\phi(\boldsymbol{\theta})$.
- This leaves us with only $\boldsymbol{\theta}$ to estimate.

- In this application, the nuisance parameters ϕ are the coefficients defining the function $s(t)$ fitting the SLEDAI scores.
- The parameters that interest us are those that define the functions $\beta(t)$ and $\alpha(t)$ defining the differential equation that we want to model the data.

Here are the steps:

1. convert the DIFE to a differential operator L
2. derive function $s(t|\lambda, L)$ that smooths the data using the differential operator L as a roughness penalty; that is, that minimizes

$$\sum_{i}^{N}[s_i - s(t_i)]^2 + \lambda \int [Ls(t)]^2 dt$$

3. minimize the *profiled* error sum of squares

$$SSE(\lambda, L) = \sum [s_i - s(t_i|\lambda, L)]^2$$

with respect to the parameters α, γ and θ that define operator L

The linear differential operator L

$$Ls(t) = \beta(t)s(t) + Ds(t) - \alpha u(t)$$

This is just a rearrangement of the differential equation

$$Ds(t) = -\beta(t)s(t) + \alpha(t)u(t)$$

If function $s(t)$ solves the differential equation, then $Ls(t) = 0$ for all t.

The smoothing function $s(t|\lambda, L)$

- Let $s(t)$ have the basis function expansion in terms if K basis functions $\phi_k(t)$:

$$s(t) = \sum_{k}^{K} c_k \phi_k(t) = \mathbf{c}'\boldsymbol{\phi}(t)$$

- Let N by K matrix $\mathbf{\Phi}$ contain the basis function values $\phi_k(t_i)$, and $\mathbf{y}$ contain the SLEDAI scores to be smoothed.
- Then $\mathbf{c}(\lambda, L) = [\mathbf{\Phi}'\mathbf{\Phi} + \lambda \mathbf{R}(L)]^{-1}[\mathbf{\Phi}'\mathbf{y} + \lambda \mathbf{r}(L)]$
- where $\mathbf{R}(L) = \int [L\boldsymbol{\phi}(t)][L\boldsymbol{\phi}(t)]'\, dt$
- and $\mathbf{r}(L) = \int [L\boldsymbol{\phi}(t)] u(t)\, dt$

The profiled error sum of squares $SSE(\lambda, L)$

- Substituting our expression for $\mathbf{c}(\lambda, L)$, we have

$$SSE(\lambda, L) = \|\mathbf{y} - \mathbf{\Phi}\mathbf{c}(\lambda, L)\|^2$$

- We have effectively eliminated the coefficients $\mathbf{c}$ defining the expansion of $s(t)$ from the problem by profiling.
- We no longer need the explicit penalty term because the penalty is implicit in the expression for $\mathbf{c}(\lambda, L)$.
- It only remains to optimize with respect to the parameters defining L.
- For a small number of parameters, numerical optimization with numerical gradient computation works fine.
- For larger numbers of parameters, an expression for the gradient will be helpful, and can be worked out with a little effort.

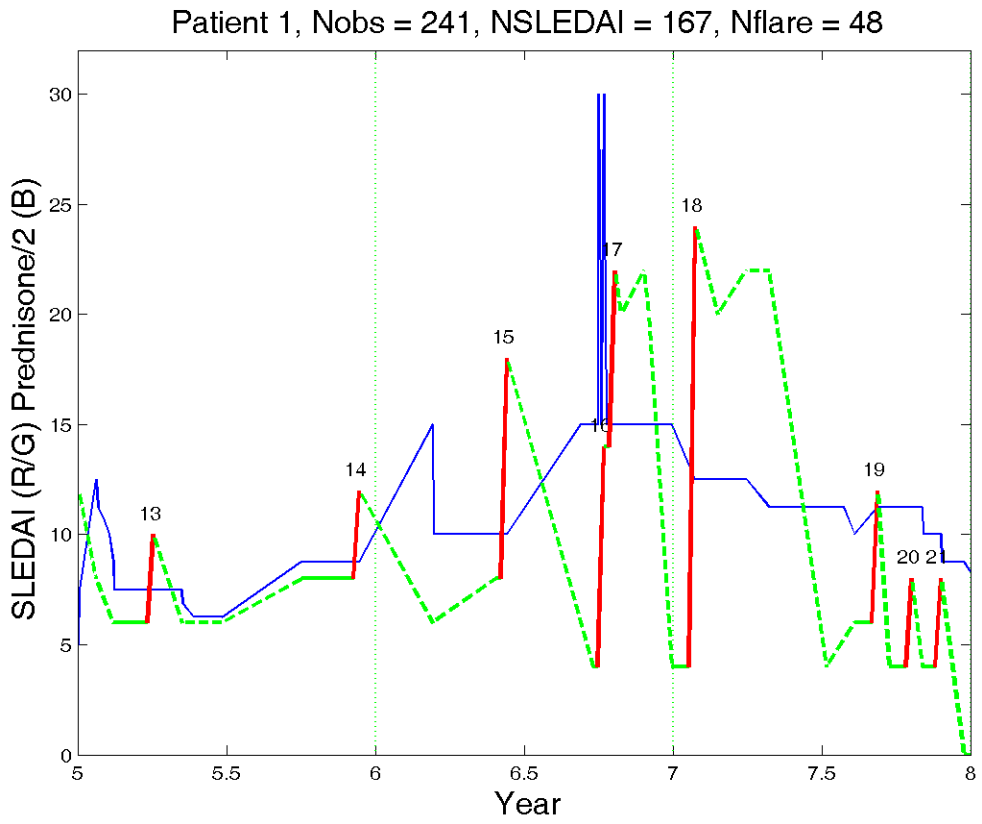
How do I choose smoothing parameter λ?

- The larger λ, the more the smoothing function $s(t_i|\lambda, L)$ will satisfy $Ls(t) = 0$, and will be a solution to the differential equation.
- However, we may want to use in–between values of λ that allow for effects in the data that cannot be accommodated by the DIFE.
- In simulated data experiments, choosing λ by generalized cross-validation has worked well, in the sense of finding a good estimate of the DIFE's parameters.

What about the basis functions $\phi_k(t)$?

- We use order 4 B-spline basis functions, with a knot at every data point, and
- three coincident knots at the times of onset and offset of flares.
- Coincident knots allow the first derivative to be discontinuous at flare boundaries, as required by the model.

7. Some results for the lupus record



- There 30 SLEDAI measurements and 8 flares in this segment.
- Flare onsets were set at 0.02 years prior to the time of the first flare score in a flare.
- The duration δ of a disease period was set to 0.02 years.
- The 30 SLEDAI scores were augmented by the 8 onset time values paired with SLEDAI scores equal to the previous scores.
- Three knots were placed at each of the 16 disease boundaries.
- One knot was placed at the mid-point of each of the 9 non-disease periods.
- The total number of knots was 59 and 61 order 4 B-spline basis functions defined the smoothing function $s(t)$.

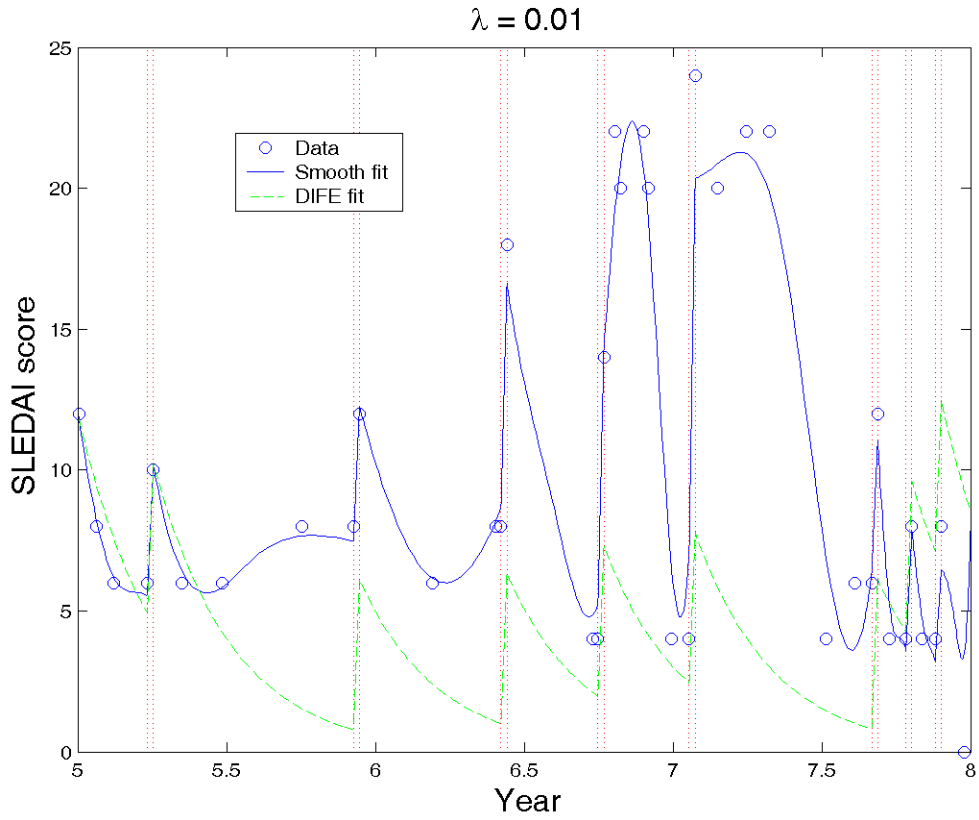
A model simplification

- Preliminary results indicated huge values for γ, implying that $\beta(t)$ moved extremely rapidly between virtually zero and the maximum value, defined by θ.
- We decided to simplify the differential equation for $\beta(t)$ to

$$D\beta(t) = \theta[1 - u(t)]$$

- This implies linear increase within a flare episode, and exponential decrease afterwards, with a rate constant θ.

Results for $\lambda = 0.01$



What we see

- With this small value of λ, the fitting function $s(t)$ is not required to be a good solution to the differential equation.
- Consequently it can fit the data well.
- The differential equation solution captures the general structure of a flare well, but cannot accommodate the variation in flare intensities that are in the data.
- We need to improve the differential equation.

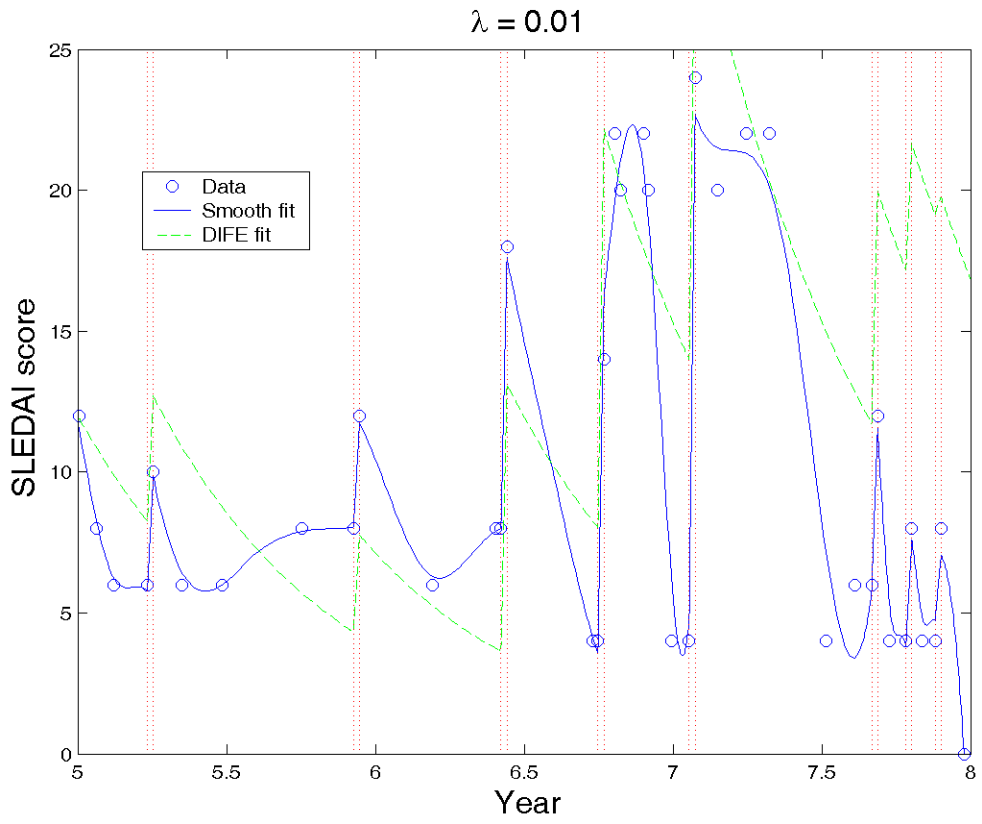
A variable gain model

- Flare intensities are determined by the gain in the differential equation.
- The gain in

$$Ds(t) = \beta(t)s(t) + \alpha u(t)$$

  is controlled by parameter α.
- Clearly the gain is not constant.
- We now let $\alpha(t)$ be time-varying.
- We will use only 6 order 4 spline basis functions (equally spaced knots), so as to allow for only mild variation in gain.

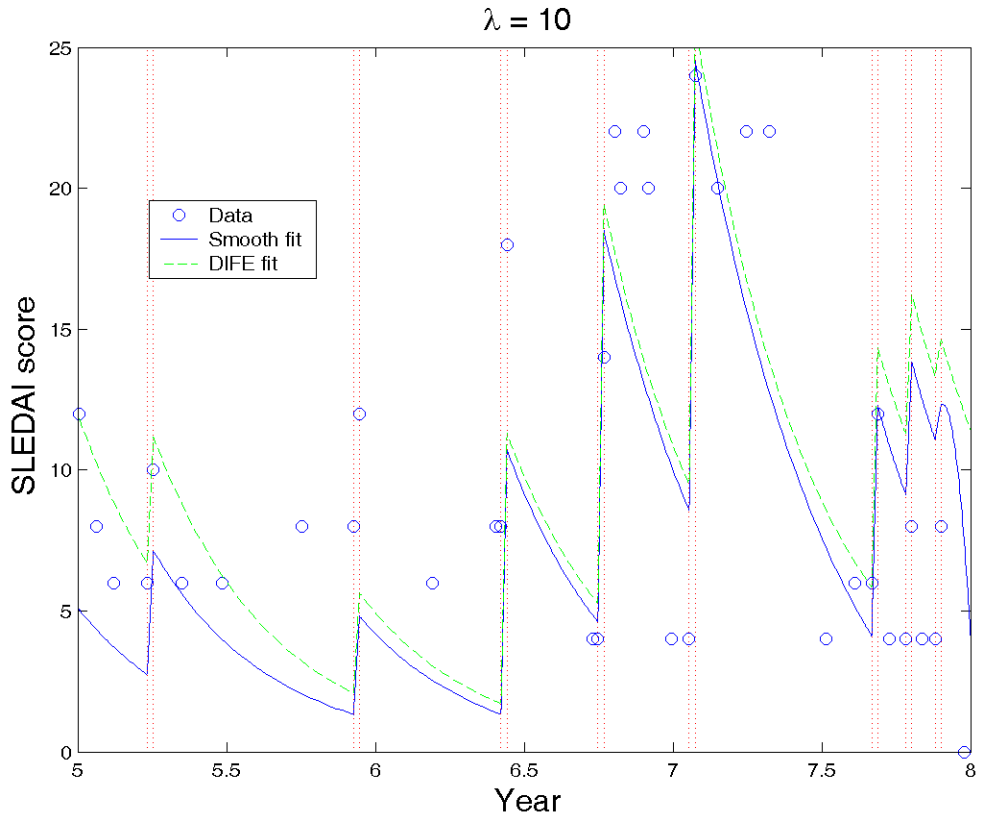
Variable gain results for $\lambda = 0.01$



What we see

- Now the differential equation also fits the data fairly well.
- With this level of λ we can still see substantial differences between the smoothing function and the DIFE, especially on the right.
- Also, $s(t)$ can capture the large oscillations in SLEDAI scores for flares 3, 4 and 5, but the DIFE cannot.
- How well will we do if we increase λ so that $s(t)$ comes much closer to satisfying the differential equation?.

Variable gain results for $\lambda = 10$



What we see now

- The smoothing function $s(t)$ is very close to being a solution to the DIFE.
- We miss the large SLEDAI score at the 3rd flare, and the low values at the beginning of the 5th flare.
- It looks like the rate constant θ is too small to capture the very rapid decay in symptoms for some flares.
- We may need to make $\theta(t)$ time–varying, too.

8. Conclusions

- It looks like we have a simple (2 parameters) model for flare structure that works.
- We have the flexibility to extend the model to capture medium-term changes in flare intensities and possibly decay rates.
- Our capacity to fit the data seems adequate given the imprecision of a SLEDAI score.
- The next step is to show the interaction with prednisone dosage.
- The interaction is likely to go both ways:
  - Prednisone affects symptoms
  - Symptoms affect the prednisone dose
- We need a system of two differential equations, one for SLEDAI score and one for prednisone.
- Each equation will have a term for the other variable.