Estimation of natural HIV disease progression incorporating unknown infection dates

Ian James & Elizabeth McKinnon

Centre for Clinical Immunology and Biomedical Statistics
Murdoch University and Royal Perth Hospital
Perth, Western Australia
Carriage of HLA-Bw4 alleles is associated with inferior post-ART recovery despite apparently better natural progression.

Rauch et al. (2008) CID 46, 1921-5.
Natural CD4% progression – from first CD4
For a single individual:

![Graph showing CD4 levels over time]

Months from seroconversion
For a single individual:

\[ \alpha S(t) \]

Typically unknown

**Months from seroconversion**

Known

\textit{cf. seroconverters}
Model: \[ y_i = \alpha S(t_1 + \Delta_i) + \varepsilon_i \]

Individual specific (random effects)  \quad \text{Covariates}

eg. "reverse" von Bertalanffy

\[ y_i = \alpha^* (1 - \exp(-k[t_0 - t_i])) + \varepsilon_i \]
Growth:
Rate of growth proportional to remaining growth

von Bertalanffy

"reverse" von Bertalanffy

CD4

Rate of decline proportional to current "loss"

k=0.01
k=0.03
k=0.05

$0 \leq t \leq t_0$

$0 \leq CD4 \leq 0$

Months since seroconversion
$t_0 = 120$
Estimating equation:

\[ \eta_i = (y_1 - y_i) - (\alpha^* - y_i)(1 - \exp(-k\Delta_i)) \]

\[ E(\eta_i) = 0 \quad \text{if } k \text{ is } \sim \text{ non-random.} \]

Solve \[ X^T \eta = 0 \]

⇒ consistent estimates

Inferential properties from general estimating equation theory – \textbf{requires only 2 measures/person}

James, Biometrics 1991
Estimating equation:

\[ \eta_i = (y_1 - y_i) - (\alpha^* - y_i)(1 - \exp(-k\Delta_i)) \]

\[ E(\eta_i) = 0 \quad \text{if } k \text{ is } \sim \text{ non-random.} \]

Solve \[ T \eta = 0 \]

\[ \rightarrow \text{consistent estimates} \]

Inferential properties from general estimating equation theory – requires only 2 measures/person

James, *Biometrics* 1991
eg. $x = 1$ if carry Bw4, 0 if not

$$\alpha^* = \exp(\alpha_1 + \alpha_2 x) \quad \alpha_1 (k_1 ?)$$

$$k = \exp(k_1 + k_2 x) \quad \text{individual specific}$$

Solve

$$\sum \eta_i = 0 \quad \sum \Delta_i \eta_i = 0$$

$$\sum x_i \eta_i = 0 \quad \sum x_i \Delta_i \eta_i = 0$$

Optimality?
Non-linear mixed model:

\[ y_i = \alpha^* (1 - \exp(-k[t_0 - t_1 - \Delta_i])) + \epsilon_i \]

Assume log-linear effects as before

\[ S^+ / R / SAS \ldots \]

Taffe & May, Stats. Med. 2008
Incorporating known seroconverter dates:

**Unknown** seroconversion date:

\[ y_i = \alpha^* (1 - \exp(-k[t_0 - t_1 - \Delta_i])) + \epsilon_i \]

**Known** seroconversion date:

\[ y_i = \alpha^* (1 - \exp(-k[t_0 - t_1 - \Delta_i])) + \epsilon_i \]

Combine using indicators - possibly different parameters in each case
Example:

115 Caucasian males from WAHIV with at least 4 pre-ART CD4% measures over at least 12 months, no seroconversion date and went onto treatment.

- 70 carry Bw4, 45 no Bw4
- average 13.4 measurements per case
Example:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>SE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \log(\alpha^* ) )</td>
<td>3.50</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>( \log( t_0 - t_1 ) )</td>
<td>4.08</td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td>( \log( k ) )</td>
<td>-3.42</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>( k.Bw4 )</td>
<td>-0.50</td>
<td>0.15</td>
<td>0.0012</td>
</tr>
<tr>
<td>( t_0 - t_1.Bw4 )</td>
<td>0.50</td>
<td>0.11</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

(Asymptote \( p > 0.5 \))

- Slope \(-0.66\) for No Bw4
- Slope \(-1.08\) for Bw4

\[ p = 0.002 \]
IF we believed SC and non-SC have same natural progression:

\[ p = 0.002 \]

Median \( t_f \): SC 5.1 mths  Non-SC 24.6 mths
Estimating equation vs mixed model:

- First and last measurements only for estimating equation
- All measurements for NLME

<table>
<thead>
<tr>
<th></th>
<th>Est. Equ.</th>
<th>NLME</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log(\alpha^*)$</td>
<td>3.40 (0.05)</td>
<td>3.50 (0.04)</td>
</tr>
<tr>
<td>$\log(k)$</td>
<td>-2.37 (1.05)</td>
<td>-3.42 (0.14)</td>
</tr>
<tr>
<td>k.Bw4</td>
<td>-0.75 (1.37)</td>
<td>-0.50 (0.15)</td>
</tr>
</tbody>
</table>

Thank you