Proportional Hazards Models
With Continuous Marks
Yanqing Sun, Peter B. Gilbert, and Ian W. McKeague

Presented by Jason Shao

University of Washington Biostatistics
Stat/Biost 572

April 23, 2013
Randomized vaccine efficacy trials against HIV

Five major trials (beyond Phase II-A) since 1998

Major difficulty: Differential vaccine efficacy (VE)

Genetic diversity of HIV between and within infected individuals

Vaccines less effective against differing infecting forms
Differential vaccine efficacy: a huge problem

- The problem: Evaluating the possibility of broad protection

Characterizing genetic diversity

Sequence HIV from infected:

Mismatches at sequence loci:

Substitution matrix (weights):

- Methodology for **continuous measures** is the focus of this study
Goal: Evaluate VE in presence of competing risks, defined by a mark (covariate observed only in cases of failure).

Mark-specific VE in time-to-event data:

\[ VE(t, v) = 1 - \frac{\lambda(t, v|X = 1)}{\lambda(t, v|X = 0)} \]

\[ \lambda(t, v|X = x) = \lim_{\Delta t \to 0} \frac{\Pr(T \in [t, t + \Delta t), V = v|X = x, T > t)}{\Delta t} \]

- \( T \): failure time
- \( V \): mark variable
- \( X \): Vaccine (1) or placebo(0). Could be a vector, including other covariates of interest.
Discrete competing risks

- Prentice et al. (1978) proposed model for finite risk categories ($\nu = 1,...,m$) under the proportional hazards model:

$$\lambda(t, \nu | X) = \lambda_0(t, \nu) \exp(\beta(\nu)^T X)$$

- Baseline hazard $\lambda_0$ factors out in partial likelihood:

$$\mathcal{L}_\nu(\beta_\nu) = \prod_{i=1}^{d_\nu} \frac{e^{\beta_\nu^T X_{\nu(i)}(t)}}{\sum_{j=1}^{n} Y_{\nu(j)}(t_{\nu(i)}) e^{\beta_\nu^T X_{\nu(j)}(t)}}$$

- $t_{\nu(1)} < ... < t_{\nu(d_\nu)}$: uncensored failure times with cause $\nu$
- $X_{\nu(i)}(t)$: covariate vector for $\nu(i)$th individual (possibly $t$-dependent)
- $Y_{\nu(i)}(t)$: indicator of $\nu(i)$th individual being at risk for $\nu$ at $t$. 
Continuous competing risks

- The authors extend this model to continuous (bounded) marks (\textit{w.l.o.g.}, assume \( v \in (0, 1) \))
- Address a bivariate mark-specific hazard function:

\[
\lambda(t, v | X = x) = \lim_{\Delta t, \Delta v \to 0} \frac{\Pr(T \in [t, t + \Delta t), V \in [v, v + \Delta v) | X = x, T > t)}{\Delta t \Delta v}
\]
Continuous competing risks

- **Localized log partial likelihood contains a kernel function** $K$:

$$\ell_v(\beta_v) = \sum_{i=1}^n \int_0^1 \int_0^\tau K_h(u - v)$$

$$\times \left[ \beta^T(v)X_i(t) - \log \left( \sum_{j=1}^n Y_j(t) e^{\beta^T(v)X_j(t)} \right) \right] \times N_i(dt, du)$$

- $K_h(\cdot)$: Kernel function with bandwidth $h$
- $t \in (0, \tau)$: Follow-up period
- $N_i(t, v) = I(X_i \leq t, \delta_i = 1, V_i \leq v)$ is a marked point counting process, with $\delta_i$ being a failure indicator
- Kernel function ”borrows” from observations with mark near $v$
Point estimation and asymptotic results

- Estimation of $\beta(v)$ (log hazard ratio).
  - $\hat{\beta}(v)$ solves score equations derived from localized partial likelihood
  - Find solution using Newton-Raphson algorithms
- Authors show asymptotic consistency and normality of $\hat{\beta}(v)$ as well as the cumulative vaccine efficacy, $\hat{CV}(v) = \int_0^v 1 - \exp(\hat{\beta}(u)) du$.
- Also find a tractable estimator $\hat{\rho}^2(v)$ of the asymptotic variance of $\hat{CV}(v)$
Hypothesis testing

- Test for overall vaccine efficacy:
  \[ H_{10} : \text{VE}(v) = 0 \text{ for all } v \in [a, b] \text{ vs.} \]
  - \( H_{1a} \) (general): \( \text{VE}(v) \neq 0 \text{ for some } v \)
  - \( H_{1m} \) (monotone): \( \text{VE}(v) \geq 0 (> 0 \text{ for some } v) \)

- Test for differences in vaccine effect:
  \[ H_{20} : \text{VE}(v) \text{ does not depend on } v \in [a, b] \text{ vs.} \]
  - \( H_{2a} \) (general): \( \text{VE}(v) \text{ depends on } v \in [a, b] \)
  - \( H_{2m} \) (monotone): \( \text{VE}(v) \text{ decreases as } v \text{ increases over } [a, b] \)
**Problem:** Vaccine development hindered by genetic diversity of HIV and differential VE.
- Previous methods only work with discrete marks of failure

**Solution:** Develop MLE and hypothesis testing approaches for failure time data with *continuous* marks of failure.

**What’s next?**
- Simulation studies to assess power and size of hypothesis tests
- Analyze results from VaxGen 004, a Phase III trial \( n = 5403 \).
- Compare to discrete competing risk analyses: does power really increase?