HIV VACCINE TRIALS NETWORK

# Analysis of Vaccine Effects on Post-Infection Endpoints

## Biostat 578A Lecture 3

# Data Collected in Phase IIb/III Vaccine Trial

- Longitudinal data from all randomized subjects:
  - HIV tests
  - Vaccine-induced immune responses to a panel of HIV isolates
  - Covariates (e.g., risk behavior)
- Longitudinal data from infected subjects:
  1. Viral loads/CD4 cell counts
  2. ART initiation
  3. Clinical events
  4. Immune responses to the infecting HIV strain and to a panel of HIV isolates
  5. Genetic sequences (and phenotypes) of (multiple clones of) infecting HIVs

# Approach to Post-infection Endpoints for Assessing $VE_P$/$VE_I$

- Assess vaccine effects both in the randomized cohort and in the infected subcohort
  - Clinical endpoints (HIV-related conditions, WHO stage 2/3, etc.)
  - Surrogate endpoints for AIDS and secondary transmission (viral load, CD4, etc.)
- Analytic challenges of surrogate endpoints:
  - Surrogate effects may not predict clinical effects
  - ART use obscures direct assessment of mid/late vaccine effects on viral load/CD4 count
  - Post-randomization selection bias

Gilbert PB, DeGruttola V, Hudgens M, Self S, Hammer S, Corey L. (2003). *JID* **21**:2933-2947.

HIV VACCINE TRIALS NETWORK

# *Time-to-Event Late Endpoints*

- ART initiation
- HIV-related clinical events
- Composite endpoints
  - ART initiation or viral failure $>$ 100,000 copies/ml
  - ART initiation or CD4 failure $<$ 200 copies/ml

  - E.g., surrogate efficacy parameter

    $VE_C =$ Percent reduction (vaccine vs placebo) in composite endpoint rate by 2 years

# *Conduct Both Intent-to-Treat (ITT) and Infected Subcohort Analyses*

- **Pros randomized cohort analysis:**
  - Intent-to-treat (ITT) [unbiased]
  - Interpretation
  - Risks/benefits determined for all vaccinees
- **Cons:**
  - Cannot separate vaccine effects on infection and post infection
  - Follow-up period for counting endpoints restricted to infection monitoring period
    - Difficult to follow randomized cohort long-term

# *Why Must Restrict Follow-up to the Infection Monitoring Period*

- Suppose all subjects followed for 3 years for infection, and all infected subjects are followed 5 more years for a disease endpoint
- Ideal design would follow all subjects 8 years from entry
  - Too expensive to follow 3000-10,000 subjects for 8 years
- Kaplan-Meier analysis that censors those uninfected at 3 years would underestimate the survival probability
- Alternatively, assuming that all subjects uninfected at 3 years would not experience the disease endpoint would overestimate the survival probability
- Kaplan-Meier analysis following everyone for 3 years only is unbiased

# Conduct Both ITT and Infected Subcohort Analyses

- **Pros infected subcohort analysis:**
  - Interest to infer vaccine effects on HIV progression in HIV infected persons
  - Feasible to follow the infected subcohort long-term [$\sim$10% of randomized cohort]
- **Cons infected subcohort analysis:**
  - Post-randomization selection bias: Infected vaccine and infected placebo groups may differ in other ways besides vaccine/placebo assignment
  - Conduct sensitivity analyses (causal inference methods)

# Analysis of Longitudinal Viral Loads and CD4 Cell Counts

- Notation: $T$ scheduled fixed visit times post-infection Dx
  - $Y_t = \log_{10}$ viral load at visit $t$
  - $X_t =$ all covariates measured up to time $t$
  - $R_t =$ indicator of whether the subject attends visit $t$ and is ART-naive at visit $t$
- Estimands of interest:

$$E[Y_t|Vaccine] - E[Y_t|Placebo]$$

$$E[Y_t|Vaccine, X_t^{other}] - E[Y_t|Placebo, X_t^{other}]$$

**for a hypothetical situation where no one starts ART during follow-up**

# Analysis of Longitudinal Viral Loads and CD4 Cell Counts

- Consider analyses that count all $Y_t$'s missing that are measured after ART initiation
- Types of missing data
  - Missing Completely at Random (MCAR): The probability of missing $Y$ does not depend on $Y$ (it may depend on $X$)
  - Missing at Random (MAR): The probability of missing $Y$ depends on observed $Y$'s, but not on unobserved $Y$'s
  - Not Missing at Random (NMAR): The probability of missing $Y$ depends on unobserved $Y$'s

# Analysis of Longitudinal Viral Loads and CD4 Cell Counts

- Two ways that $Y$'s may be missing: ART initiation and study drop-out
- Study drop-out may be MCAR
  - In Vax004, viral load over time did not predict the rate of study drop-out: $RR = 1.23$ per $1\log_{10}$ higher viral load, 95% CI 0.88-1.72, $p = 0.23$
- Starting ART is not MCAR
  - In Vax004, viral load predicted the rate of starting ART:
    - Viral load at 1 month: $RR = 1.57$ per $1\log_{10}$ higher viral load, 95% CI 1.22-2.01, $p = 0.0003$
    - Viral load over time: $RR = 1.88$ per $1\log_{10}$ higher viral load, 95% CI 1.51-2.34, $p < 0.0001$

# *Methods for Addressing the MAR Data*

- Weighted Generalized Estimating Equations (GEE) methods
- Likelihood based linear mixed effects (LME) models
- Other Approaches

HIV VACCINE TRIALS NETWORK

# Notation for Complete Data

- $T$ visit times common to all $n$ subjects
- $Y_i = (Y_{i1}, \cdots, Y_{iT})'$: Complete data response vector
- $X_i = (X_{i1}, \cdots, X_{iT})'$: Covariate matrix
  - $X_{it}$ a $p \times 1$ covariate vector at visit time $t$

HIV VACCINE TRIALS NETWORK

# Notation for Observed Data

- Subject $i$ has data at visits $1, \cdots, T_i, 1 \leq T_i \leq T$
- $Y_i^o = (Y_{i1}^o, \cdots, Y_{iT_i}^o)'$: Observed response vector
- $X_i^o = (X_{i1}^o, \cdots, X_{iT_i}^o)'$: Observed covariate matrix
  - $X_{it}^o$ a $p \times 1$ covariate vector at time $t$

HIV VACCINE TRIALS NETWORK

# Weighted GEE Approach

- Estimand of interest: $E(Y_{it}|X_{it}) = \mu_{it}$
- Generalized Linear Model (GLM): $g(\mu_{it}) = X_{it}'\beta$
  - $g$ a known link function
  - For $Y$ continuous, e.g., $\log_{10}$ viral load, $g(x) = x$, so that the GLM is

$$E(Y_{it}|X_{it}) = X_{it}'\beta$$

# Weighted GEE Approach

- Example models, with $V_i = 1$ if vaccine and $V_i = 0$ if placebo; $t_i$ a fixed visit time among $1, \cdots, T_i$
  - $X_{it} = (1, t_i, V_i) : E(Y_{it}|X_{it}) = \beta_0 + \beta_1 t_i + \beta_2 V_i$

    $$\Rightarrow \beta_2 = E[Y_{it}|V_i = 1, t_i] - E[Y_{it}|V_i = 0, t_i]$$

    [Constant vaccine effect over time]
  - $X_{it} = (1, t_i, V_i, t_i V_i) : E(Y_{it}|X_{it}) = \beta_0 + \beta_1 t_i + \beta_2 V_i + \beta_3 t_i V_i$

    $$\Rightarrow \beta_2 + \beta_3 t_i = E[Y_{it}|V_i = 1, t_i] - E[Y_{it}|V_i = 0, t_i]$$

    [Vaccine effect changing linearly over time]

# Weighted GEE Methods for Addressing MAR Data

- Standard GEE of Liang and Zeger (Biometrika, 1986):

$$\sum_{i=1}^{n} D_i^{o\prime}(X_i^o, \beta)(V_i^o)^{-1}\left[Y_i^o - \mu_i^o(\beta)\right] = 0$$

where $D_i^{o\prime}(X_i^o, \beta) = \partial \mu_i^o / \partial \beta$ is a $T_i \times p$ matrix and $V_i^o$ is a working covariance matrix for $Y_i^o$

  - $V_i^o = diag(var(Y_{it}^{1/2}))C_i^o(\rho)diag(var(Y_{it}^{1/2}))$ [$T_i \times T_i$ matrix]
  - $C_i^o(\rho)$ is a working correlation matrix depending on an unknown vector parameter $\rho$, which is estimated

# Weighted GEE Methods for Addressing MAR Data

- Examples
  - Independent working correlation: $C_i^o =$ identity matrix
  - Exchangeable working correlation: $C_i^o =$ 1's on the diagonal and common correlation $\rho$ for all off-diagonal elements
  - Auto-regressive-1: $C_i^o =$ 1's on the diagonal and $\rho^k$ for $k$ steps off the diagonal (logical choice for repeated measures data)

# Weighted GEE Methods for Addressing MAR Data

- Solution $\widehat{\beta}$ obtained by iteratively reweighted estimation of $\beta$ (McCullogh and Nelder, 1989, Generalized Linear Models)
- Variance of $\widehat{\beta}$ estimated by “sandwich variance estimator"

# *Weighted GEE Methods for Addressing MAR Data*

- Standard GEE provides unbiased estimation of $\beta$ under MCAR, but not under MAR
  - GEE not valid for vaccine trials
- Weighted GEE valid for MAR data (Robins, Rotnitzky, Zhao (RRZ), 1995, JASA)

# Weighted GEE Methods for Addressing MAR Data

- Weighted GEE (RRZ):

$$\sum_{i=1}^{n} D_i'(X_i,\beta)(V_i)^{-1}W_i\left[Y_i-\mu_i(\beta)\right]=0$$

where $D_i'(X_i,\beta)=\partial\mu_i/\partial\beta$ and $V_i=A_iC_iA_i$ is a working covariance matrix for $Y_i$

  - $W_i$ is a $T\times T$ diagonal matrix of time-specific weights:

$$W_i=diag(R_{i1}w_{i1},\cdots,R_{iT}w_{iT})$$

$R_{it}=I(Y_{it}$ observed$)=$ indicator of whether $i$th subject has a pre-ART value at visit $t$

$w_{it}>0$ for observed pre-ART visits; $=0$ o/w

# Weighted GEE Methods for Addressing MAR Data

- $w_{it} =$ reciprocal of the probability the $i$th subject is observed (pre-ART) at the $t$th visit
  - If observed with probability 1, assign weight 1
  - If observed with probability 1/2, assign weight 2
  - If observed with probability 1/10, assign weight 10
- Heuristically, reconstruct the complete dataset by weighting the observed data

# Weighted GEE Methods for Addressing MAR Data

- $w_{it}$ is estimated using a missing data model
- Let $\lambda_{it} = Pr(R_{it} = 1 | R_{i(t-1)} = 1, X_i, Y_i, \alpha)$
  - For the first time point, assume $R_{i1} = 1$ and set $\lambda_{i1} = 1$
  - Second and later time-points: The MAR assumption implies

$$\lambda_{it} = Pr(R_{it} = 1 | R_{i(t-1)} = 1, X_i, (Y_{i1}, \cdots, Y_{i(t-1)}), \alpha)$$

i.e., Missingness depends only on observed data and a parametric model with unknown parameter $\alpha$

# Weighted GEE Methods for Addressing MAR Data

- Example missingness model:

$$logit\{\lambda_{it}(\alpha)\} = Z_{it}'\alpha$$

where $Z_{it}$ may include anything (covariates and/or responses) observed prior to time $t$

- The MLE of $\alpha$ is computed, and $\lambda_{it}$ is estimated by

$$\widehat{\lambda}_{it} = Z_{it}'\widehat{\alpha}$$

# Weighted GEE Methods for Addressing MAR Data

- $w_{it}$ is then estimated as the reciprocal of the product of conditional probabilities:

$$\widehat{w}_{it} = \left[\widehat{\lambda}_{i1} \times \cdots \times \widehat{\lambda}_{it}\right]^{-1}$$

- $\widehat{\beta}$ and $Var(\widehat{\beta})$ computed similarly as in standard GEE

# Comments on Weighted GEE Approach for MAR Data

- If the marginal mean model is correctly specified, and the parametric model for missingness is correctly specified, then weighted GEE provides unbiased estimation of $\beta$, and the sandwich variance estimator is consistent, even if the working correlation matrix is misspecified
- If the missingness model is correctly specified, weighted GEE peforms better than standard GEE
- If the missingness model is misspecified, weighted GEE can perform poorly, and standard GEE sometimes does better

# *Comments on Weighted GEE Approach for MAR Data*

- Weighted GEE can have problems (biased estimation and huge variance estimates) if the $\widehat{w}_{it}$'s are large (i.e., estimated probability of being observed is near zero)

# Comments on Weighted GEE Approach for MAR Data

- In vaccine trials, the estimated probability of having pre-ART viral loads might be near zero!
  - This is due to standardized treatment guidelines
- Current HVTN policy: Provide ARTs to all infected participants when they meet pre-specified criteria
  - E.g., CD4 $<$ 300 cells/mm$^3$, viral load $>$ 100,000 copies/ml, or HIV-related clinical symptoms
  - $Z_{it}$ may predict perfectly whether a participant will start ART!
  - $w_{it} = \infty$ if and only if
    Pr(subject $i$ drops out by $t$ or starts ART by $t$) = 1
    - If everybody receives ART when they should, then $w_{it} = \infty$ for some subjects and weighted GEE breaks down

# *Comments on Weighted GEE Approach for MAR Data*

- Irony with weighted GEE: Want to predict missingness to handle MAR missingness, but if predict missingness too well then the method fails
- Weighted GEE does not handle well the censoring of responses
  - Can use ad hoc approaches, such as assigning all left-censored viral loads a value equal to half the detection limit
  - Can study the marginal mean conditional on non-censored response (truncated mean)

# *Linear Mixed Effects (LME) Models Approach*

- LME models provide unbiased inferences under assumptions:
  1. Multivariate normality of viral loads
  2. All predictors of ART initiation are captured in observables
  - The LME models can accommodate censored viral values (Jim Hughes, 1999, Biometrics)

HIV VACCINE TRIALS NETWORK

# Standard LME Model

- $Y_i = X_i'\beta + Z_i'\gamma_i + e_i$

  $\beta$ a vector of fixed effects
  $\gamma$ a vector of random effects for subject $i$
  $e_i$ a vector of random errors

- Assume $\gamma_i$ and $e_i$ are independent with

$$\begin{aligned} \gamma_i &\sim N(0, \Sigma) \\ e_i &\sim N(0, \sigma^2 I) \end{aligned}$$

HIV VACCINE TRIALS NETWORK

# Standard LME Model

- Hughes (1999, Biometrics) developed an EM algorithm to obtain the MLEs of $\beta$, $\Sigma$, and $\sigma^2$, allowing for any number of subjects to have left-censored or right-censored viral loads
- Louis' (1982, JRSS B) method used to estimate $Var(\widehat{\beta})$
  - This method adjusts the variances of the esimated fixed effects for the information lost due to censoring

HIV VACCINE TRIALS NETWORK

# Standard LME Model

- Viral load assay used in the VaxGen trials: quantification range 400-750,000 copies/ml
  - Left-censoring: $Y_{it} \leq 400$ copies/ml
  - Right-censoring: $Y_{it} \geq 750,000$ copies/ml
    - In Vax004, 259 values $\leq 400$ and 43 values $\geq 750,000$ (18.2% of all values censored)
- In simulations Hughes (1999) showed that his method does much better (less bias in estimating $\beta$, $\Sigma$, and $\sigma^2$) than typically applied hoc methods that assign an arbitrary value for censored values (e.g., the detection limit or one-half the detection limit)
  - Efficiency gain derived from the parametric distributional assumption

# Comparison of Weighted GEE and LME Models

- LMEs handle MAR data without needing a missingness model, and performance improves the extent to which the variables included in the model predict missingness
- LMEs can handle left-censoring and right-censoring of viral loads
  - LMEs preferred if ART initiation is predicted very well, if a credible missingness model cannot be built, or if left/right censoring is heavy

# Comparison of Weighted GEE and LME Models

- Weighted GEE avoids the assumption of multivariate normal viral loads, and is more robust to specification of the correlation structure of measurements over time
  - Weighted GEE preferred if its assumptions are credible, ART is not predicted very well, and left/right censoring is light

# *Example: Analysis of VAX004 Data*

- Marginal mean model of $Y = log_{10}$ viral load with covariates:
  - vaccination status
  - white/nonwhite
  - baseline risk score (0-7)
  - time after infection diagnosis in years
  - education (4 levels)
  - region (6 regions)
  - calendar time of infection diagnosis (4 intervals)
  - age (5 levels)

HIV VACCINE TRIALS NETWORK

# *Methods to be Compared*

- Weighted GEE of RRZ
- Multiple imputation for GEE (Paik, JASA, 1997)
- Robust efficient score (Annie Qu, 2006, unpublished manuscript)

- Analyze 319 subjects with pre-ART viral load data

HIV VACCINE TRIALS NETWORK

# Results

| Variable | WGEE | | | Mult. Imp. | | | Rob. Effi c. Score | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\widehat{\beta}_j$ | $s.e.$ | $Z$ | $\widehat{\beta}_j$ | $s.e.$ | $Z$ | $\widehat{\beta}_j$ | $s.e.$ | $Z$ |
| Intercept | 4.084 | 0.342 | 11.928$*$ | 4.277 | 0.326 | 13.086$*$ | 4.420 | 0.344 | 12.849$*$ |
| Vaccine | -0.013 | 0.097 | -0.138 | -0.030 | 0.092 | -0.334 | 0.016 | 0.080 | 0.199 |
| White | 0.058 | 0.164 | 0.354 | 0.208 | 0.131 | 1.586 | 0.198 | 0.114 | 1.735 |
| Risk score | 0.035 | 0.038 | 0.920 | 0.0196 | 0.033 | 0.587 | 0.015 | 0.025 | 0.602 |
| Time (years) | 0.025 | 0.058 | 0.441 | -0.048 | 0.057 | -0.844 | 0.062 | 0.036 | 1.723 |
| Educ.2 | 0.362 | 0.172 | 2.097$*$ | 0.048 | 0.162 | 0.298 | 0.177 | 0.136 | 1.302 |
| Educ.3 | 0.131 | 0.154 | 0.851 | -0.033 | 0.159 | -0.207 | 0.110 | 0.131 | 0.840 |
| Educ.4 | 0.076 | 0.200 | 0.380 | -0.133 | 0.185 | -0.721 | -0.021 | 0.159 | -0.134 |
| Region.1 | -0.164 | 0.228 | 0.113 | -0.252 | 0.251 | -1.006 | -0.399 | 0.289 | -1.376 |
| Region.2 | 0.034 | 0.236 | 0.147 | 0.036 | 0.266 | 0.138 | -0.307 | 0.302 | -1.016 |
| Region.3 | 0.197 | 0.228 | 0.863 | 0.064 | 0.251 | 0.254 | -0.212 | 0.287 | -0.739 |
| Region.4 | 0.098 | 0.229 | 0.431 | -0.091 | 0.252 | -0.362 | -0.315 | 0.289 | -1.090 |
| Region.5 | 0.144 | 0.245 | 0.588 | -0.030 | 0.268 | -0.114 | -0.263 | 0.293 | -0.890 |

$*P < 0.05$

# Results, Continued

| | WGEE | | | Mult. Imp. | | | Rob. Effi c. Score | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | $\widehat{\beta}_j$ | $s.e.$ | $Z$ | $\widehat{\beta}_j$ | $s.e.$ | $Z$ | $\widehat{\beta}_j$ | $s.e.$ | $Z$ |
| CaltimeD.2 | -0.383 | 0.218 | -1.756 | -0.188 | 0.173 | -1.084 | -0.228 | 0.160 | -1.426 |
| CaltimeD.3 | -0.591 | 0.199 | -2.970$*$ | -0.309 | 0.174 | -1.769 | -0.406 | 0.156 | -2.600$*$ |
| CaltimeD.4 | -0.164 | 0.185 | -0.886 | 0.050 | 0.158 | 0.320 | 0.001 | 0.144 | 0.007 |
| CaltimeD.5 | -0.445 | 0.208 | -2.139$*$ | -0.292 | 0.180 | -1.623 | -0.390 | 0.163 | -2.393$*$ |
| age.2 | 0.184 | 0.174 | 1.053 | -0.017 | 0.138 | -0.129 | 0.011 | 0.130 | 0.082 |
| age.3 | 0.171 | 0.184 | 0.930 | -0.068 | 0.140 | -0.488 | -0.001 | 0.125 | -0.006 |
| age.4 | 0.389 | 0.188 | 2.066$*$ | 0.168 | 0.160 | 1.050 | 0.130 | 0.148 | 0.881 |
| age.5 | -0.333 | 0.221 | -1.509 | -0.196 | 0.243 | -0.805 | -0.147 | 0.236 | -0.624 |

$*P < 0.05$

HIV VACCINE TRIALS NETWORK

# Notes on Results

- The 3 methods are all stable (standard errors reasonably small) and perform fairly similarly
  - Estimates differ for region and time
- The assumption that $1/w_{it} > 0$ is evidently reasonable
- Robust efficient score approach provides lower standard errors except for intercept and region

# Other Approaches to Analyzing the MAR Data

- If fewer than $<$ 50% of infected subjects start ART by the late time-point $t$, can minimize dependent censoring problem by estimating the median difference (vaccine vs placebo) in pre-ART viral load
- "Utility" ITT approach: Analyze ranks of outcomes for all randomized participants
  - E.g., ranks from best to worst:
    1. Not infected
    2. Infected and did not start ART within 2 years
       -Rank by level of viral loads
    3. Infected and started ART within 2 years