

Statistical Science Issues in HIV Vaccine Trials: Part I

Outline

1. Study population
2. Criteria for selecting a vaccine for efficacy testing
3. Measuring effects of vaccination
 - biological markers to measure vaccine effect on HIV infectiousness and disease progression
 - direct measures
4. Sample size considerations
5. Complications in measuring vaccine effects due to antiretroviral treatment of infected trial participants
6. Correlates of vaccine protection
7. Behavioral Issues
8. Extrapolation of trial results

1. Which Population? - Ethics and Statistics

- Ethical issues
 - informed consent
 - dealing with potential HIV positive status
 - discrimination against participants in employment, housing, insurance, and travel
 - social stigma
 - effect of participation on risk behavior

- Statistical issues
 - need population with high infection incidence
 - need population with high retention rate

2. Criteria for Selecting a Vaccine for Efficacy Testing

- Which Vaccine? - Basic Science

General Criteria

- safe in diverse populations
- effective in preventing, eradicating, or suppressing multiple HIV strains
 - * **match** the vaccine to local HIV strains
- protective against systemic (needle) and mucosal (heterosexual sex) exposure
- practical: easy to administer, affordable, heat stable

- Protects animals against infection and/or disease in challenge studies with multiple strains of HIV, SIV, or SHIV
- Phase I/II surrogate markers
 - elicits strong, broad, prolonged neutralizing antibodies
 - elicits strong, broad, prolonged T cell responses

2. Current Problems

- Extensive genetic and antigenic variation of HIV- current results show limited cross-strain protection
- Stimulating humoral (antibodies) and cellular (CTLs and CD4 help) immune responses
- What are the correlates of immunity?
- No ideal animal model

3. Potential Effects of Vaccination

- Risk of HIV infection depends on **behavioral** and **biological** parameters:
 - Behavioral
 - * number unsafe contacts
 - * fraction unsafe contacts with infected individuals
 - Biological
 - * susceptibility probability
 - * infectivity probability

3. Potential Effects of Vaccination

- Measurement objectives
 - **program effectiveness:** the effect the vaccine would have on the spread of HIV if used in a vaccination program
 - * need a non-traditional design to assess program effectiveness
 - **biological potency:** the effect the vaccine would have on reducing the susceptibility and infectivity probability

3. Classical Preventive Trial Design

- Classical randomized, double-blind trial aims at assessing biological potency
- Three biological vaccine efficacy (VE) measures
 - VE_S : VE at reducing susceptibility to infection
 - VE_I : VE at reducing infectiousness of infected persons
 - VE_P : VE at slowing HIV disease progression

3. Measuring VE-susceptibility (VE_S)

- Estimate VE_S by comparing the rates or hazards of infection in the vaccine and placebo group
 - $VE_S = 1 - RR$
- **Key assumption for validity:** equal exposure in vaccine and placebo groups
- Randomization and blinding justify this assumption
- Useful to collect data on HIV exposure

3. Measuring VE-susceptibility (VE_S)

- **Binary endpoint**

- $V\widehat{E}_{S Bin} = 1 - \frac{AR_v}{AR_u} = 1 - \frac{n_{Iv}/N_v}{n_{Iu}/N_u}$

- $\widehat{\text{Var}}\left(\log\left\{\frac{n_{Iv}}{N_v}/\frac{n_{Iu}}{N_u}\right\}\right) = \frac{N_v - n_{Iv}}{n_{Iv}N_v} + \frac{N_u - n_{Iu}}{n_{Iu}N_u}$

- Test the hypothesis of no vaccine efficacy with the statistic

$$Z = \frac{\log\left\{\frac{n_{Iv}}{N_v}/\frac{n_{Iu}}{N_u}\right\}}{\sqrt{\left(\widehat{\text{Var}}\left(\log\left\{\frac{n_{Iv}}{N_v}/\frac{n_{Iu}}{N_u}\right\}\right)\right)}}$$

- adjusting for covariates:

- * fit a logistic regression model with HIV infection (Yes/No) as the response, and vaccination status and other covariates as predictor variables

- **Poisson-based estimate**

$$VE_S = 1 - \frac{\text{incidence rate in vaccinated}}{\text{incidence rate in unvaccinated}} = 1 - \frac{IR_v}{IR_u}$$

- estimate VE_S by $\widehat{VE}_{SPoiss} = 1 - \frac{\widehat{IR}_v}{\widehat{IR}_u} = 1 - \frac{n_{Iv}/PY_v}{n_{Iu}/PY_u}$

with PY_V = person-years at-risk in group V , $V \in \{v, u\}$

- $e^{\widehat{\beta}} = \frac{\widehat{IR}_v}{\widehat{IR}_u}$ can be obtained by fitting a generalized linear model (Poisson regression)

- adjusting for covariates

- fit a log-linear Poisson regression model including covariates,

estimate $\widehat{VE}_{SPoiss} = 1 - e^{\widehat{\beta}_{Adj}}$

- **Failure time endpoint**

- under a proportional hazards assumption,

$$VE_S = 1 - \frac{\text{hazard of infection in vaccinated}}{\text{hazard of infection in unvaccinated}} = 1 - \frac{\lambda_v}{\lambda_u}$$

- estimate VE_S by $\widehat{VE}_{Sph} = 1 - e^{\widehat{\beta}}$
 - * $e^{\widehat{\beta}} = \frac{\widehat{\lambda}_v}{\widehat{\lambda}_u}$ is an estimate of the hazard ratio obtained from fitting the proportional hazards model
- test the hypothesis of no vaccine effect (hazard ratio equals 1) with the log rank statistic
- adjusting for covariates:
 - * fit a proportional hazards model including covariates, estimate $\widehat{VE}_{Sph} = 1 - e^{\widehat{\beta}_{Adj}}$

3. Measuring VE-susceptibility (VE_S)

- If failure times are available, recommend use of the proportional hazards model for estimating VE_S
- Need to check the assumption of proportional hazards between vaccine and placebo groups
- vaccine efficacy may change over time, which violates the proportional hazards assumption
- What to do if the proportional hazards assumption is violated?

3. Measuring VE-susceptibility (VE_S)

- **Estimating time-varying vaccine efficacy**
- **One Approach:** Construct a nonparametric smoothed estimate $\widehat{VE}_S(t) = 1 - \widehat{RR}(t) = 1 - \exp\{\widehat{\lambda}_1(t)/\widehat{\lambda}_2(t)\}$, and calculate its standard error
- **Procedure:**
 1. Estimate $\lambda_1(t)$ and $\lambda_2(t)$ by nonparametric kernel smoothing:

$$\widehat{\lambda}_i(t) = \frac{1}{b_i} \int_{t_1}^{t_2} K\left(\frac{t-s}{b_i}\right) d\widehat{\Lambda}_i(s),$$

where $\widehat{\Lambda}_i(t) = \int_0^t (1/Y_i(s)) dN_i(s)$ is the Nelson-Aalen estimator of the cumulative hazard function $\Lambda_i(t) = \int_0^t \lambda_i(s) ds$, with

$$N_i(t) = \sum_{j=1}^{n_i} N_{ij}(t) = \sum_{j=1}^{n_i} I(X_{ij} \leq t, \Delta_{ij} = 1),$$

$$Y_i(t) = \sum_{j=1}^{n_i} Y_{ij}(t) = \sum_{j=1}^{n_i} I(X_{ij} \geq t)$$

2. It can be shown that $n_i^{2/5}(\widehat{\lambda}_i(t) - \lambda_i(t)) \approx N(0, \sigma^2(t))$, $i = 1, 2$

3. A 95% confidence interval for $VE_S(t)$ is given by

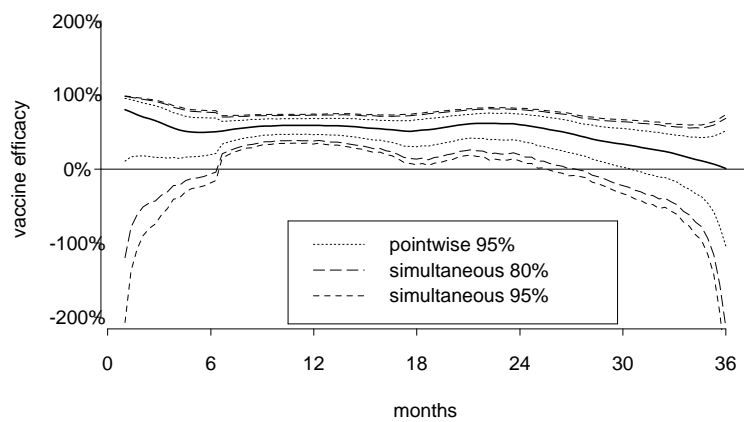
$$1 - \exp \left\{ (\widehat{\lambda}_1(t)/\widehat{\lambda}_2(t)) \pm 1.96 * \widehat{\text{Var}}(\widehat{\lambda}_1(t)/\widehat{\lambda}_2(t))^{1/2} \right\}$$

where $\widehat{\text{Var}}(\widehat{\lambda}_1(t)/\widehat{\lambda}_2(t)) = \frac{1}{\widehat{\lambda}_2(t)^4} \left[\widehat{\lambda}_2(t)^2 \frac{\widehat{\sigma}_1(t)^2}{n_1^{4/5}} + \widehat{\lambda}_1(t)^2 \frac{\widehat{\sigma}_2(t)^2}{n_2^{4/5}} \right]$

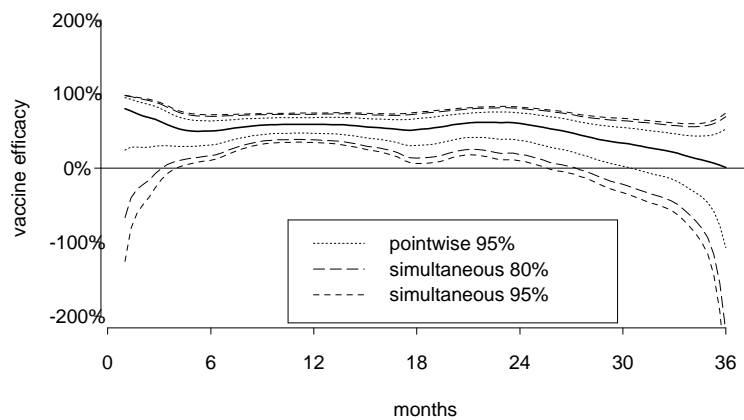
(using the delta method)

4. **Simultaneous** confidence intervals can also be computed

(a) Vaccine efficacy, simulation bands



(b) Vaccine efficacy, analytic bands



3. Measuring VE-susceptibility (VE_S)

- **Estimating time-varying vaccine efficacy**
- **Another Approach:** Parametrically model vaccine efficacy effects
 - the nonparametric smoothing method may suggest parametric forms for $VE_S(t)$
 - Fit a proportional hazards model

$$\lambda(t|V) = \lambda_0(t)e^{\beta(t)V}$$

with a specific parametric function for $\beta(t)$,
e.g., $e^{\beta(t)}$ linear, quadratic, or a step function

- for a preventive vaccine trial with follow-up period $[0, \tau]$, suppose vaccine injections are given during $[0, \tau_1]$. Then one might expect vaccine efficacy increasing from 0 to τ_1 , and then decreasing from τ_1 to τ . This suggests a parametric form like:

$$e^{\beta(t)} = [1 + \beta_1 t]I\{0 \leq t < \tau_1\} + [1 + \beta_1 \tau_1 + \beta_2(t)]I\{\tau_1 \leq t \leq \tau_1\},$$

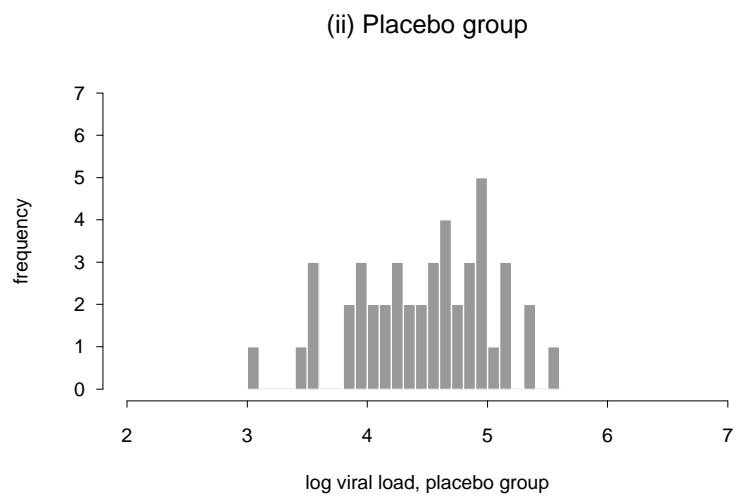
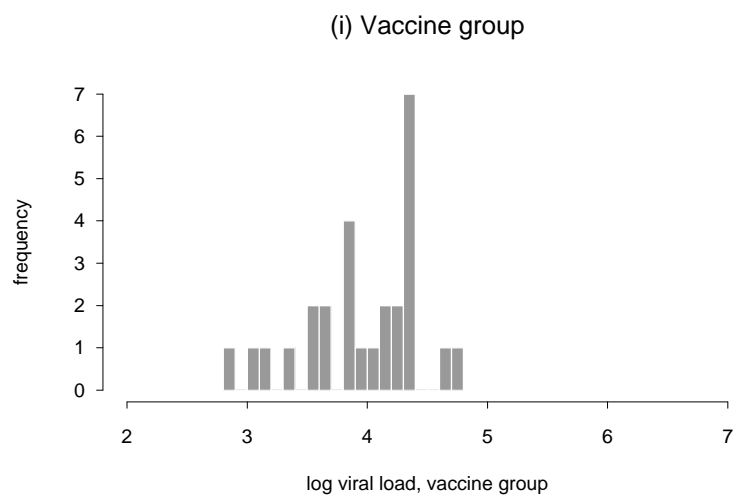
or, more generally, like

$$e^{\beta(t)} = [1 + g_1(\beta_1, t)]I\{0 \leq t < \tau_1\} + [1 + \beta_1 \tau_1 + g_2(\beta_2, t)]I\{\tau_1 \leq t \leq \tau_1\}$$

- advantages of parametric approach: increase efficiency of vaccine efficacy estimators, increase power of statistical tests, and give a clearly interpretable picture of time-varying vaccine effects

3. Measuring VE-infectiousness (VE_I)

- Cannot estimate VE_I from a classical trial design
- Can get **indirect** surrogate information on VE_I by measuring a biological marker of infectiousness
 - e.g., plasma or genital secretion viral load
 - * suppressed viral load may indicate reduced infectiousness



3. Statistical Power for Comparing Initial Viral Load

- Suppose a viral load measurement is taken from each infected subject shortly after infection is discovered
- Suppose 50% vaccine efficacy (VE_S)
80% power to detect the mean difference in log viral load

total number infected	mean group difference in log viral load
30	$1 \log_{10}$
50	$3/4 \log_{10}$
100	$1/2 \log_{10}$

- **Beware:** Potential selection bias in the viral load comparison
 – be careful interpreting the results!

3. Selection Bias in Comparing Viral Load Distributions

- a major challenge posed to conducting unbiased inference of the vaccine's effect on viral load is that the comparison groups are selected by the post-randomization event HIV infection
- as a consequence, the comparison of viral load between infected groups is prone to many of the selection biases which can occur in observational studies
 - one of the most relevant sources of potential selection bias:
partial efficacy of the vaccine at preventing infections

3. Selection Bias in Comparing Viral Load Distributions

- If in truth the vaccine has no impact on viral load, a standard two-sample test comparing viral load distributions between infected groups could draw the false conclusion that vaccination enhances or suppresses viral burden
- Or, if in truth the vaccine has a suppressive impact on viral load, the standard test may fail to detect it
- An erroneous inference that the vaccine enhances infection could be particularly destructive to continued expedient development of a safe and partially efficacious vaccine
 - e.g., at an interim look stop a trial prematurely

3. Measuring VE-infectiousness (VE_I)

- Can estimate VE_I directly by monitoring persons who are exposed to infected vaccine trial participants
 - e.g., through **Partner Designs**
 - e.g., through **Community Randomized Designs**

3. Measuring VE-infectiousness (VE_I)

- **Partner designs** as formulated by Longini, Datta, and Halloran (1996)
 - Suppose the design of a classic randomized, Phase III, double-blind, vaccine versus placebo preventive HIV vaccine efficacy trial
 - * To estimate VE_I , augment the design with sexual partners
 - **Augmented Design 1 (non-randomized)**: Add m_u and m_v steady sexual partners of unvaccinated and vaccinated primary participants, respectively, and monitor them for HIV infection
 - * for m_u partnerships, if either partner is infected during the trial, then his or her partner will be directly exposed to an **unvaccinated** partner during his or her primary infection
 - * for m_v partnerships, if either partner is infected during the trial, then his or her partner will be directly exposed to a **vaccinated** partner during his or her primary infection
 - **Augmented Design 2 (randomized)**: Randomize steady

sexual partners of primary participants to vaccine or placebo. Then, about one-quarter of partnerships will have both partners vaccinated, one-quarter will have both partners unvaccinated, and half will have one partner vaccinated and one partner unvaccinated.

- * let β be the probability that an infected, **unvaccinated** trial participant will infect his or her unvaccinated steady sexual partner during the follow-up period of the trial
- * define $VE_I = 1 - \phi$, where ϕ is the fractionate reduction in β for an infected **vaccinated** trial participant

- **Goal of augmented partner designs:**

- estimate VE_S and VE_I
- test
 1. $H_{0S} : VE_S = 0$
 2. $H_{0I} : VE_I = 0$
 3. $H_{0C} : VE_S = 0$ and $VE_I = 0$

Hypothesis tests:

$H_0 : VE_S = 0$ 1 df test with Wald statistic

$$Z = \widehat{VE}_S / \sqrt{\widehat{\text{Var}}(\widehat{VE}_S)}$$

$H_0 : VE_I = 0$ 1 df test with Wald statistic

$$Z = \widehat{VE}_I / \sqrt{\widehat{\text{Var}}(\widehat{VE}_I)}$$

$H_0 : VE_S = 0$ and $VE_I = 0$ 2 df likelihood ratio test with approximate Chi-square test statistic

$$\chi^2 \approx -2[\ln L(\widehat{\gamma}, \widehat{\beta}, 1, 1) - \ln L(\widehat{\gamma}, \widehat{\beta}, \widehat{\theta}, \widehat{\phi})]$$

3. Measuring VE-infectiousness (VE_I)

- Statistical power to detect VE_I from an augmented partners design depends on:
 - VE_S
 - number primary participants with steady partners
 - secondary attack rate
 - quality of HIV exposure data

3. Measuring VE-infectiousness (VE_I)

- Example

- 4000 primary trial participants
- 1000 monogamous partners enrolled
- secondary attack rate = 50%
- $VE_S = 20\%$
 - Nonrandomized partners design: 57% power to detect $VE_I = 60\%$
 - Randomized partners design: 87% power to detect $VE_I = 60\%$
- A secondary attack rate of 50% may be unrealistically high
 - * if it is lower, then sensitivity to detect VE_I is lost
- Other references: Longini et al. 1998, Statistics in Medicine 17:1121-36, Rida 1996, Statistics in Medicine 15:2393-2404

3. Measuring VE-infectiousness (VE_I)

- Limitation of partner designs: their feasibility is in question
 - in order to estimate the potential reduction in infectiousness, couples must continue to engage in risky behaviors when one partner is known to be infected. Furthermore, the uninfected partner must avoid risk outside the partnership before and after his/her partner becomes infected. Such behaviors must occur despite counseling interventions which will be implemented as part of the trial. At present it is unknown how couples would behave in a trial designed to reduce infectiousness.

3. Measuring VE-infectiousness (VE_I)

- **Alternative: Cluster randomization trials**
 - definition: experiments in which intact social units (e.g., families, villages, clinics, brothels) rather than independent individuals are randomly allocated to intervention groups
- **General advantages of design:**
 - administrative convenience
 - increase compliance
 - avoid ethical considerations which might otherwise arise
 - out of necessity

- General disadvantages of design:

Presence of between-cluster variation implies:

- reduction in effective sample size
 - * extent depends on degree of within-cluster correlation and on average cluster size
- standard approaches for sample size estimation and statistical analysis do not apply
 - * application of standard sample size approaches leads to an underpowered study
 - * application of standard statistical methods generally tends to bias p-values downwards, i.e. could lead to spurious statistical significance

3. Measuring VE-infectiousness (VE_I)

- How is the dependence between responses of cluster members measured?
 - intracluster correlation coefficient

$$\rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2}$$

σ_B^2 = variability between clusters

σ_W^2 = variability within clusters

$\rho = 1$ when the responses of cluster members are identical (i.e. when $\sigma_W^2 = 0$)

ρ gets closer to zero the less correlation there is among subjects in the same cluster

- Variance inflation factor: when all clusters have exactly m subjects the variance of the estimated effect of intervention is inflated by

$$1 + (m - 1) * \rho$$

3. Measuring VE-infectiousness (VE_I)

- Sample size adjustment, completely randomized design
 - compute number of subjects required per treatment group using standard sample size formulas
 - multiply result by the variance inflation factor $1 + (\bar{m} - 1)\rho$, where

\bar{m} = average cluster size

ρ = prior estimate of intraclass correlation

3. Measuring VE-infectiousness (VE_I)

- There are three designs which are commonly used when intact clusters are assigned to interventions:
 - completely randomized (involving no pre-stratification or matching)
 - matched pair (in which one of two clusters in a stratum are randomly assigned to each intervention)
 - stratified (involving the assignment of two or more clusters to each combination of intervention and stratum)

3. Measuring VE-infectiousness (VE_I)

- Examples of potential HIV vaccine trials:
 - randomize n IDU clinics to vaccine or placebo (e.g., in Bangkok)
 - randomize n STD clinics to vaccine or placebo (e.g., in Southern Africa)
 - randomize n factories to vaccine or placebo (all employees within a factory receive the same preparation)
 - * potential stratification factors:
 - country, city, or village
 - predominant mode of risk and/or level of risk
 - predominant virus type

3. Measuring VE-disease progression (VE_P)

- Follow infected trial participants 5-10 years to assess if HIV disease progresses differently in vaccinated individuals compared to unvaccinated individuals
 - takes too long!
- Measure a biological marker of disease progression
 - e.g., viral load or CD4 cell count over time
 - * compare vaccinated and unvaccinated groups by a summary measure of marker change over time, such as the slope coefficient fit by a simple linear regression model

- Use of a biological marker to measure VE_P may be misleading
 - to validate a surrogate marker, it must be a correlate of the clinical outcome (easy to check) and fully capture the net effect of vaccination on the clinical outcome (difficult or impossible to check)
 - Long-term follow-up is necessary to validate a surrogate marker

4. Sample size considerations

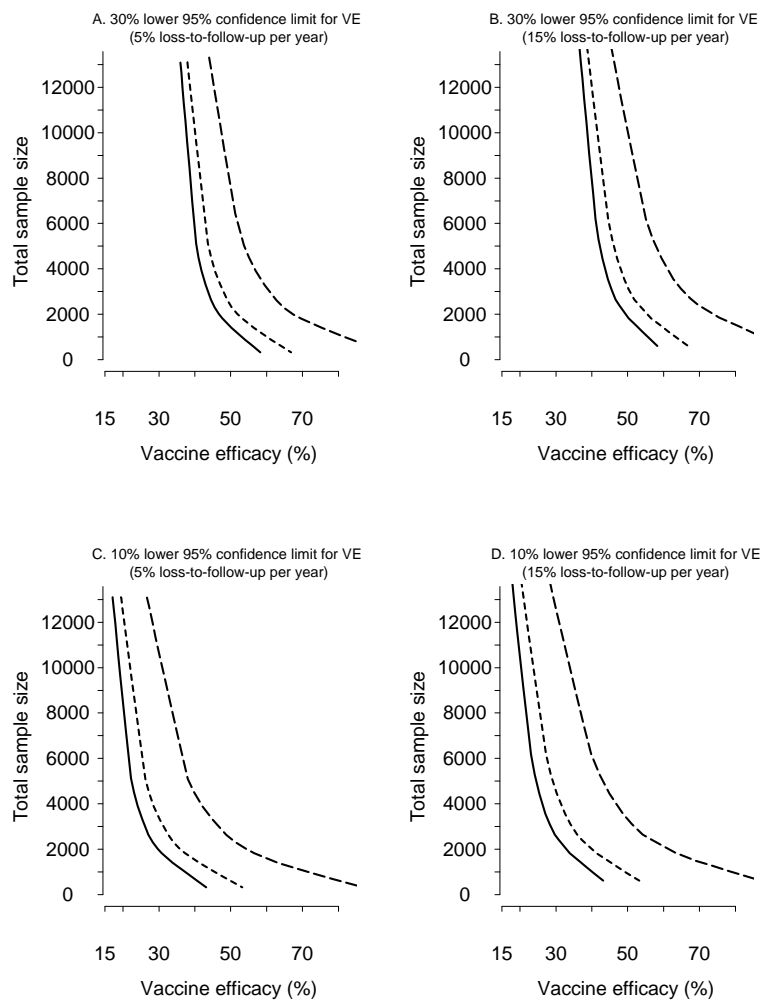
- Vaccine efficacy trials are typically powered to test the null hypothesis

$$H_0 : VE_S = \delta \quad \text{versus} \quad H_1 : VE_S > \delta$$

with $\delta = 0$ or $\delta > 0$ (e.g., $\delta = 0.10$ or 0.30)

- Factors determining sample size
 - rate of infection in population of interest
 - lowest acceptable vaccine efficacy (e.g., 30%)
 - desired type I error and power to detect the effect size of interest
 - rate of enrollment of eligible subjects
 - minimum length of follow-up required
 - anticipated rate of loss to follow-up

Assume 3 years of follow-up. The solid, dotted, and dashed lines are for 5%, 3%, and 1% annual incidence in the placebo arm.



4. Sample size considerations

- Special challenges in HIV vaccine trials:
 - Loss to follow-up rate may be higher in high-risk cohorts
 - immunization period makes up substantial part of observation period
 - counseling may have large impact on event rate
 - * counseling \Rightarrow reduced risk \Rightarrow reduced power
 - * participation \Rightarrow perceived immunity \Rightarrow higher risk behavior
(?)
- Vaccine preparedness studies
 - assess effects of counseling
 - estimate HIV incidence rates

5. Impact of Treating Infecteds with Antiretrovirals on Estimation of VE_S, VE_I, VE_P

- VE_S
 - ARVs will not interfere
- VE_I
 - augmented partners design
 - * ARVs will likely reduce the secondary attack rate, thus necessitating a larger sample size
 - viral load a marker for infectiousness
 - * only 1-3 measures of viral load can be taken on infected trial participants before initiation of ARV therapy. Thus, do not observe the long-term longitudinal viral load profile.
- VE_P
 - ARVs will make estimation difficult or impossible
- Antiretrovirals will make validation of biological markers of infectiousness or disease progression difficult

6. Assessing HIV Strain-specific Protection

- Measure characteristics of HIV isolates taken from infected trial participants
 - HIV genetic sequence
 - HIV serotype
 - other HIV phenotypes; e.g. syncytium inducing capacity, cellular tropism, tissue tropism, replicative capacity, co-receptor usage, etc.
- “Sieve Analysis”: assess how VE_S depends on characteristics of the exposing HIV [More on sieve analysis later]

7. Behavioral Issues

- Infection preventable by changes in unsafe sexual and drug injection practices
- Question: does participation lower or raise risk behavior?
- Question: does vaccination lower or raise risk behavior?
- Blinding versus unblinding
- to assess these questions, need an **unblinded trial** (or add a third observational arm)
- Evaluate program effectiveness

7. Behavioral Issues

- Assessing risk behavior
- Five reasons to assess behavior (Chesney et al., 1995):
 1. Enroll a cohort at sufficient risk for HIV infection
 2. Assess trends in subject risk behavior
 3. Estimate the extent and consequences of unblinding
 4. Permit tailoring of counseling to individual subject behaviors
 5. Allow investigation of vaccine efficacy in behavioral subgroups

8. Extrapolation of trial results

- Particular features of a vaccine trial
 - population (men, women, children, infants, IDU, commercial sex workers)
 - route of transmission (heterosexual sex, homosexual sex, needles, mother-to-child)
 - strain of HIV virus
- e.g., can results from a trial in the Bangkok IDU setting (subtypes B and E, intravenous and sexual transmission) be extrapolated to Southern Africa (subtypes A, C, and D, predominantly heterosexual transmission)?

8. Extrapolation of trial results

- Example: male-to-female heterosexual transmission
 - $p \equiv$ probability that uninfected female becomes infected from 1 sexual contact with HIV⁺ male
- We know:
 - p is **highly** variable in individuals (due to variability in susceptibility and/or infectiousness)
 - knowledge of p , how and how often people interact, and HIV prevalence determine # of infections seen in a vaccine trial
- Knowledge of p can, in theory, enable us to predict the consequences of a vaccination program in a population with different kinds/type of exposure than in the vaccine trial
- To estimate p , need to know number, type, and frequency of sexual contacts among trial subjects. Also need information on who interacts with who.
- Estimation of p complicated by several factors (heterogeneity in

susceptibility, infectiousness, mixing patterns, vaccine mechanism of efficacy)

- Under strong assumptions, p is estimated by \widehat{VE}_{SPoiss} and \widehat{VE}_{Sph}
 \implies in general, considerable uncertainty in estimate of p

Implications of Uncertainty about p

- Except for vaccines with very high efficacy (e.g., $> 90\%$), extrapolation of vaccine efficacy results to different populations can be very imprecise
 \implies in designing a vaccine trial, it may be wise to choose a study population similar to the population that would be given the vaccine if the trial result is positive

Summary

- Extra studies of effects of infectiousness and consequences of vaccine failure can be built onto classic susceptibility designs
- Extensive laboratory analysis of immune responses and genetic, antigenic, and phenotypic properties of infected participants' viruses can provide insights into the correlates of protective immunity and mechanisms of protection
 - fuels hypothesis driven iterative process of HIV vaccine development