Session 4 of Module 8: Evaluating an Immunological Correlate of Risk (Long Version, at http://faculty.washington.edu/peterg/SISMID2013.html)

Peter Gilbert

Summer Institute in Statistics and Modeling in Infectious Diseases

U of W - July 15–17, 2013

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Motivating Example: Evaluating Antibodies as CoRs in Vax004

- **Secondary objective**: Assess if various *in vitro* measurements of antibody levels in vaccinees correlate with HIV infection rate
- 8 antibody assays that measure binding/neutralization of the MN or GNE8 HIV strains
  - ELISAs to measure antibody binding: gp120, V2, V3, CD4 blocking
  - Functional assay: Neutralization of MN HIV-1

Assessing Antibodies as Correlates of Risk in Vax004

Sampling design

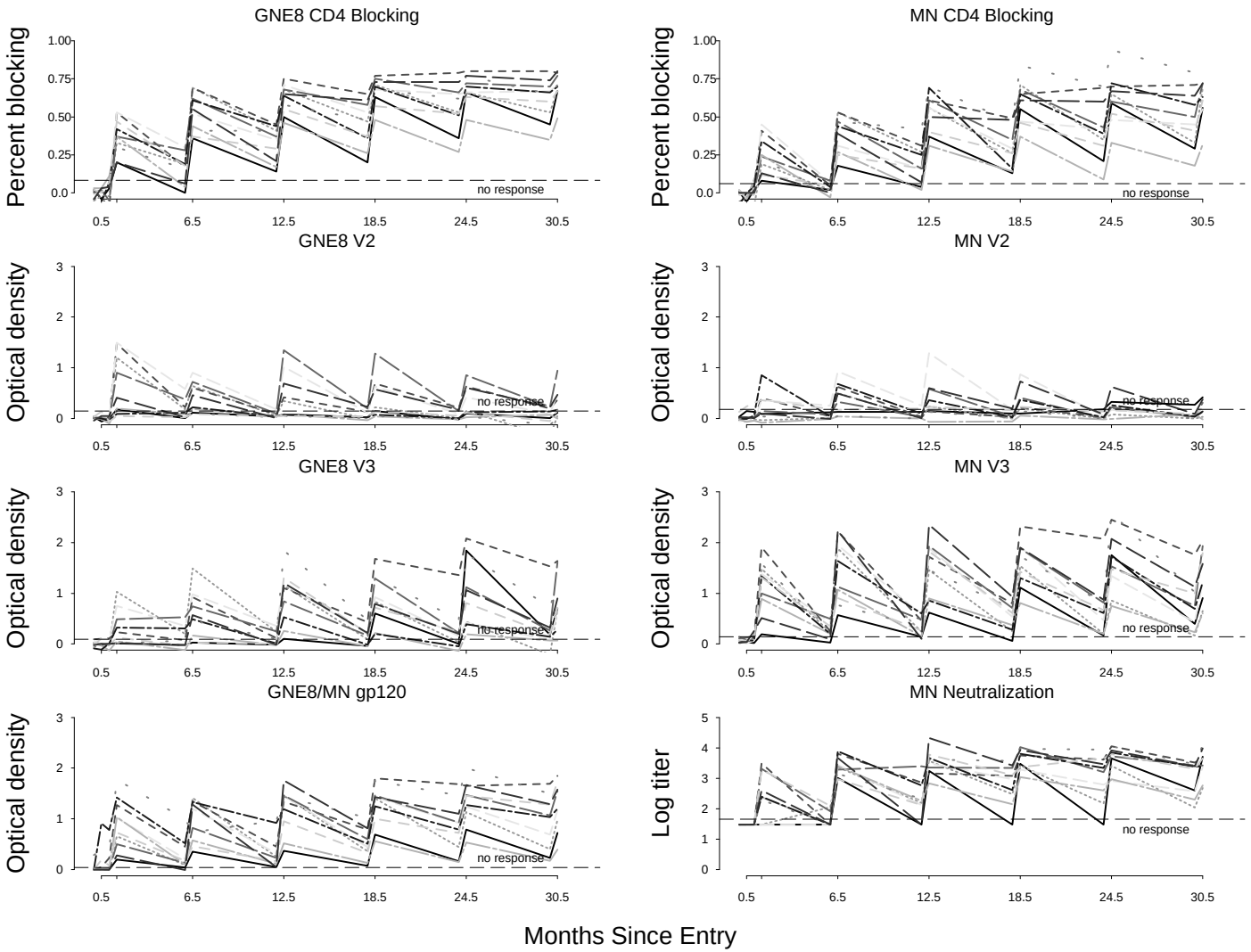
- Specimens collected:
  - Month 0, 1, 6, 12, 18, 24, 30, 36 (troughs)
  - Month 0.5, 1.5, 6.5, 12.5, 18.5, 24.5, 30.5 (peaks)
- Specimens assayed:
  - Random "subcohort" of 5% of all vaccinees ($n = 174$, all time points)
    - n=163/11 in subcohort uninfected/infected
  - All infected vaccinees ($n = 239$, last sample prior to infection)

Two Types of Correlates Questions

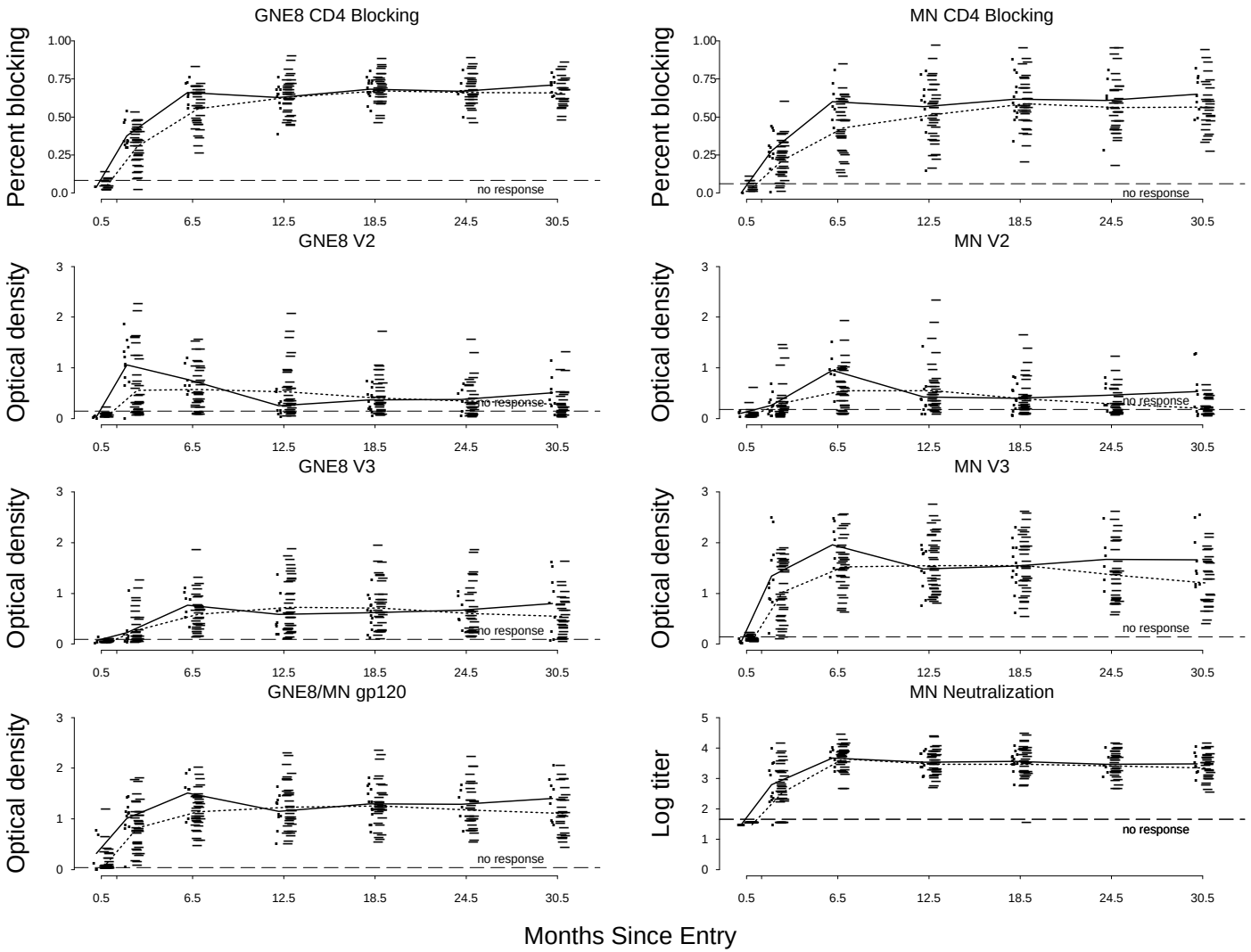
- For Vax004 and generally for efficacy trials, two types of correlates questions are of interest:
  1. If and how does the peak immune response level (e.g., 2 weeks after the last immunization) correlate with the subsequent rate of infection over a defined follow-up period?
  2. If and how does the immune response level near the time of exposure correlate with the rate of infection over a short follow-up period (e.g., until the next infection diagnostic test)?
- The first question is most useful for developing a surrogate endpoint (need something measured once near baseline)
- The second question is most useful for gaining insight into the mechanistic cause of protection (in theory immune response level at time of exposure is what matters)
- Both questions are interesting to ask, especially for vaccines with time-waning immune responses
- The following Vax004 results evaluate 'time-dependent' correlates

Example: Vax004 (Gilbert et al., 2005, JID)

- Randomly selected subject-specific antibody profiles



Peak Antibody Levels of Vaccinees (Solid/dotted = Uninfected/infected)



Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

The Cox Model with a Case-Cohort Sampling Design

- Cox proportional hazards model

$$\lambda(t|Z) = \lambda_0(t) exp\left\{\beta_0^T Z(t)\right\}$$

  - $\lambda(t|Z)$ = conditional failure hazard given covariate history until time t
  - β_0 = unknown vector-valued parameter
  - $\lambda_0(t) = \lambda(t|0)$ = unspecified baseline hazard function
    - Z are "expensive" covariates only measured on failures and subjects in the subcohort

Notation and Set-Up (Matches Kulich and Lin, 2004, JASA)

- T = failure time (e.g., time to HIV infection diagnosis)
- C = censoring time
- $X = min(T, C), \Delta = I(T \leq C)$
- $N(t) = I(X \leq t, \Delta = 1)$
- $Y(t) = I(X \geq t)$
- Cases are subjects with $\Delta = 1$
- Controls are subjects with $\Delta = 0$

Notation and Set-Up (Matches Kulich and Lin, 2004, JASA)

- Consider a cohort of n subjects, who are stratified by a variable V with K categories
- $\epsilon =$ indicator of whether a subject is selected into the subcohort
  - $\alpha_k = Pr(\epsilon = 1 | V = k)$, where $\alpha_k > 0$
- $(X_{ki}, \Delta_{ki}, Z_{ki}(t), 0 \leq t \leq \tau, V_{ki}, \epsilon_{ki} \equiv 1)$ observed for all subcohort subjects
- At least $(X_{ki}, \Delta_{ki} \equiv 1, Z_{ki}(X_{ki}))$ observed for all cases

Estimation of β_0

- With full data, β_0 would be estimated by the MPLE, defined as the root of the score function

$$U_F(\beta) = \sum_{i=1}^{n} \int_0^{\tau} \left\{ Z_i(t) - \bar{Z}_F(t, \beta) \right\} dN_i(t), \tag{1}$$

where

$$\bar{Z}_F(t, \beta) = S_F^{(1)}(t, \beta) / S_F^{(0)}(t, \beta);$$

$$\begin{aligned} S_F^{(1)}(t, \beta) &= n^{-1} \sum_{i=1}^{n} Z_i(t) exp \left\{ \beta^T Z_i(t) \right\} Y_i(t) \\ S_F^{(0)}(t, \beta) &= n^{-1} \sum_{i=1}^{n} exp \left\{ \beta^T Z_i(t) \right\} Y_i(t) \end{aligned}$$

Estimation of β_0

- Due to missing data (1) cannot be calculated under the case-cohort design
- Many modified estimators have been proposed, all of which replace $\bar{Z}_F(t, \beta)$ with an approximation $\bar{Z}_C(t, \beta)$, so are roots of

$$U_C(\beta) = \sum_{k=1}^{K} \sum_{i=1}^{n_k} \int_0^{\tau} \left\{ Z_{ki}(t) - \bar{Z}_C(t, \beta) \right\} dN_{ki}(t)$$

  - The double indices k, i reflect the stratification

Estimation of β_0

- The case-cohort at-risk average is defined as

$$\bar{Z}_C(t,\beta) \equiv S_C^{(1)}(t,\beta)/S_C^{(0)}(t,\beta),$$

where

$$\begin{aligned} S_C^{(1)}(t,\beta) &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\rho_{ki}(t)Z_{ki}(t)exp\left\{\beta^T Z_{ki}(t)\right\}Y_{ki}(t) \\ S_C^{(0)}(t,\beta) &= n^{-1}\sum_{k=1}^{K}\sum_{i=1}^{n_k}\rho_{ki}(t)exp\left\{\beta^T Z_{ki}(t)\right\}Y_{ki}(t) \end{aligned}$$

Estimation of β_0

- The potentially time-varying weight $\rho_{ki}(t)$ is set to zero for subjects with incomplete data, eliminating them from the estimation
- Cases and subjects in the subcohort have $\rho_{ki}(t) > 0$
  - Usually $\rho_{ki}(t)$ is set as the **inverse estimated sampling probability** (Using the same idea as the weighted GEE methods of Robins, Rotnitzky, and Zhao, 1994, 1995)
- Different case-cohort estimators are formed by different choices of weights $\rho_{ki}(t)$
- Two classes of estimators (N and D), described next

N-estimators

- The subcohort is considered a sample from all study subjects regardless of failure status
  - The whole covariate history $Z(t)$ is used for all subcohort subjects
  - For cases not in the subcohort, only $Z(T_i)$ (the covariate at the failure time) is used
- Prentice (1986, Biometrika): $\rho_i(t) = \epsilon_i/\alpha$ for $t < T_i$ and $\rho_i(T_i) = 1/\alpha$
- Self and Prentice (1988, Ann Stat): $\rho_i(t) = \epsilon_i/\alpha$ for all t

N-estimators

- General stratified N-estimator
  - $\rho_{ki}(t) = \epsilon_i / \widehat{\alpha}_k(t)$ for $t < T_{ki}$ and $\rho_{ki}(T_{ki}) = 1$
    - $\widehat{\alpha}_k(t)$ is a possibly time-varying estimator of α_k
    - α_k is known by design, but nonetheless estimating α_k provides greater efficiency for estimating β_0 (Robins, Rotnitzky, Zhao,1994)
    - A time-varying weight can be obtained by calculating the fraction of the sampled subjects among those at risk at a given time point (Barlow, 1994; Borgan et al., 2000, Estimator I)

D-estimators

- Weight cases by 1 throughout their entire at-risk period
- D-estimators treat cases and controls **completely separately**
  - α_k apply to controls only, so that α_k should be estimated using data only from controls
- Conditional on failure status, the D-estimator case-cohort design is similar to that of the case-control design whether or not the subcohort sampling is done retrospectively

D-estimators

- General D-estimator

$$\rho_{ki}(t) = \Delta_{ki} + (1 - \Delta_{ki})\epsilon_{ki}/\widehat{\alpha}_k(t)$$

  - Borgan et al. (2000, Estimator II) obtained by setting

$$\widehat{\alpha}_k(t) = \sum_i^n \epsilon_{ki}(1 - \Delta_{ki})Y_{ki}(t) / \sum_i^n (1 - \Delta_{ki})Y_{ki}(t),$$

    i.e., the proportion of the sampled controls among those who remain at risk at time t
  - the cch package in R (by Thomas Lumley and Norm Breslow) implements the case-cohort Cox model for N- and D-estimators (later we will practice with this function)

Main Distinctions between N- and D- Estimators

- D-estimators require data on the complete covariate histories of cases
- N-estimators only require data at the failure time for cases
  - For Vax004, the immune response in cases was only measured at the visit prior to infection, so N-estimators are valid while D-estimators are not valid

Main Distinctions between N- and D- Estimators

- For N-estimators, the sampling design is **specified in advance**, whereas for D-estimators, it can be **specified after the trial** (retrospectively)
    - D-estimators more flexible

Gaps of Both N- and D- Estimators

Estimator	Does Not Need Full Covariate Histories in Cases	Allows Outcome-Dependent Sampling
N	Yes	No
D	No	Yes

- For time-dependent correlates, none of the partial-likelihood based methods are flexible on both points
- All of the methods require full covariate histories in controls
- Full likelihood-based methods can help (later)

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Discrete Failure Time Weighted Likelihood Approach

- Work joint with Zhiguo Li and Bin Nan at the University of Michigan Biostatistics Department

- Li, Gilbert, and Nan (2008, Biometrics) developed a weighted likelihood two-phase case-cohort analysis method, for discrete failure times and allowing for time-dependent and missing immunological biomarker values
- Useful if the biomarker is time-dependent but the covariate histories are missing or partially missing in controls

Grouped Survival Data

- T_i: the underlying time to event (HIV infection) for subject i.
- C_i: the underlying censoring time for subject i.
- T_i is either known to be in one of the m fixed time intervals:

$$(t_0, t_1], (t_1, t_2], \ldots, (t_{m-1}, t_m),$$

  where $0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = +\infty$, or right censored at a visit time t_j, $1 \leq j \leq m-1$.
- X_i: a p-dimensional covariate that can be time-dependent, denoted as

$$X_i = (X_{i1}, \ldots, X_{im}).$$

- We only observe the first R_i intervals for subject i, $1 \leq R_i \leq m-1$.
- $\Delta_i = (\Delta_{i1}, \ldots, \Delta_{iR_i}, \Delta_{i,R_i+1})'$, where $\Delta_{ij} = 1$ if the event for the ith subject falls into the jth interval and $\Delta_{ij} = 0$ otherwise, $1 \leq j \leq R_i$, and $\Delta_{i,R_i+1} = 1 - \sum_{j=1}^{R_i} \Delta_{ij}$.

The Cox Model

$$\lambda(t|X(t)) = \lambda(t)\exp(X(t)'\beta).$$

- $\Lambda(t)$: the baseline cumulative hazard function.
- $\alpha_k = \Lambda(t_k) - \Lambda(t_{k-1})$, $\gamma_k = \log \alpha_k$, $k = 1, 2, \ldots, m$, and $\alpha_m = \gamma_m = \infty$.
- **Without considering censoring**, the conditional probability of the event for the ith subject falling into the jth interval given X_i is

$$P(\Delta_{ij} = 1|X_i) = e^{-\sum_{k=1}^{j-1} e^{\gamma_k + X_{ik}'\beta}} \left(1 - e^{-e^{\gamma_j + X_{ij}'\beta}}\right), 1 \leq j \leq m.$$

The Cox Model, Continued

Considering censoring, we have

$$
\begin{aligned}
& P(\Delta_i = \delta_i, R_i = j | X_i) \\
= \; & P\left\{ T_i \in (t_{j-1}, t_j], C_i \geq t_j \middle| X_i \right\}^{1-\delta_{i,j+1}} P\left\{ T_i \geq t_j, C_i \in (t_{j-1}, t_j] \middle| X_i \right\}^{\delta_{i,j+1}} \\
= \; & \prod_{\ell=1}^{j+1} \left(e^{-\sum_{k=1}^{\ell-1} e^{\gamma_k + X_{ik}'\beta}} \right)^{\delta_{i\ell}} \left(1 - e^{-e^{\gamma_j + X_{ij}'\beta}} \right)^{\delta_{ij}} f(\delta_i, j | X_i) \\
\equiv \; & L(\theta | \Delta_i = \delta_i, R_i = j) f(\delta_i, j | X_i), \qquad 1 \leq j \leq m-1,
\end{aligned}
$$

where $f(\delta_i, j | X_i) = \{P(C_i \geq t_j | X_i)\}^{1-\delta_{i,j+1}} \{P(t_j < C_i \leq t_{j+1} | X_i)\}^{\delta_{i,j+1}}$ does not contain any information about $\theta \equiv (\beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_{m-1})$ and hence can be dropped when constructing the likelihood function for θ. **Note that** $L_i(\theta)$ reduces to the likelihood contribution of the ith subject in Prentice and Gloeckler (1978).

The Weighted Likelihood Method

Weighted likelihood:

$$L_{w,n}(\theta) = \prod_{i=1}^{n} \left\{ L_i(\theta) \right\}^{w_i},$$

where

$$w_i = (1 - \Delta_{i,R_i+1}) + \frac{I(i \in \mathcal{SC})}{P(i \in \mathcal{SC}|V_i)} \Delta_{i,R_i+1}, \quad 1 \leq i \leq n.$$

and V_i's are auxiliary variables that are observable for all subjects.

Logarithm of the weighted likelihood:

$$\ell_{w,n}(\theta) = \sum_{i=1}^{n} w_i \ell_i(\theta)$$

$$= \sum_{i=1}^{n} w_i \left\{ - \sum_{j=1}^{R_i+1} \left(\Delta_{ij} \sum_{k=1}^{j-1} e^{\gamma_k + X_{ik}' \beta} \right) + \Delta_{iR_i} \log \left(1 - e^{-e^{\gamma_{R_i} + X_{iR_i}' \beta}} \right) \right\}.$$

More Missing Data

- X_{i1} is observed for all subjects in the case-cohort study.
- X_{iR_i} is observed if the event is observed for subject i.
- A randomly chosen X_{ij} is observed if subject i is censored, $1 < j \leq R_i$.
- This data frame may be used to assess X as a CoR at the fixed time-point 1 and as a time-dependent CoR.

Multiple Imputation

- For each j, $1 < j \leq m-1$, within each stratum, fit a linear model using all completely observed pairs (X_{i1}, X_{ij}):

$$X_{ij} = c_0 + c_1 X_{i1} + \epsilon,$$

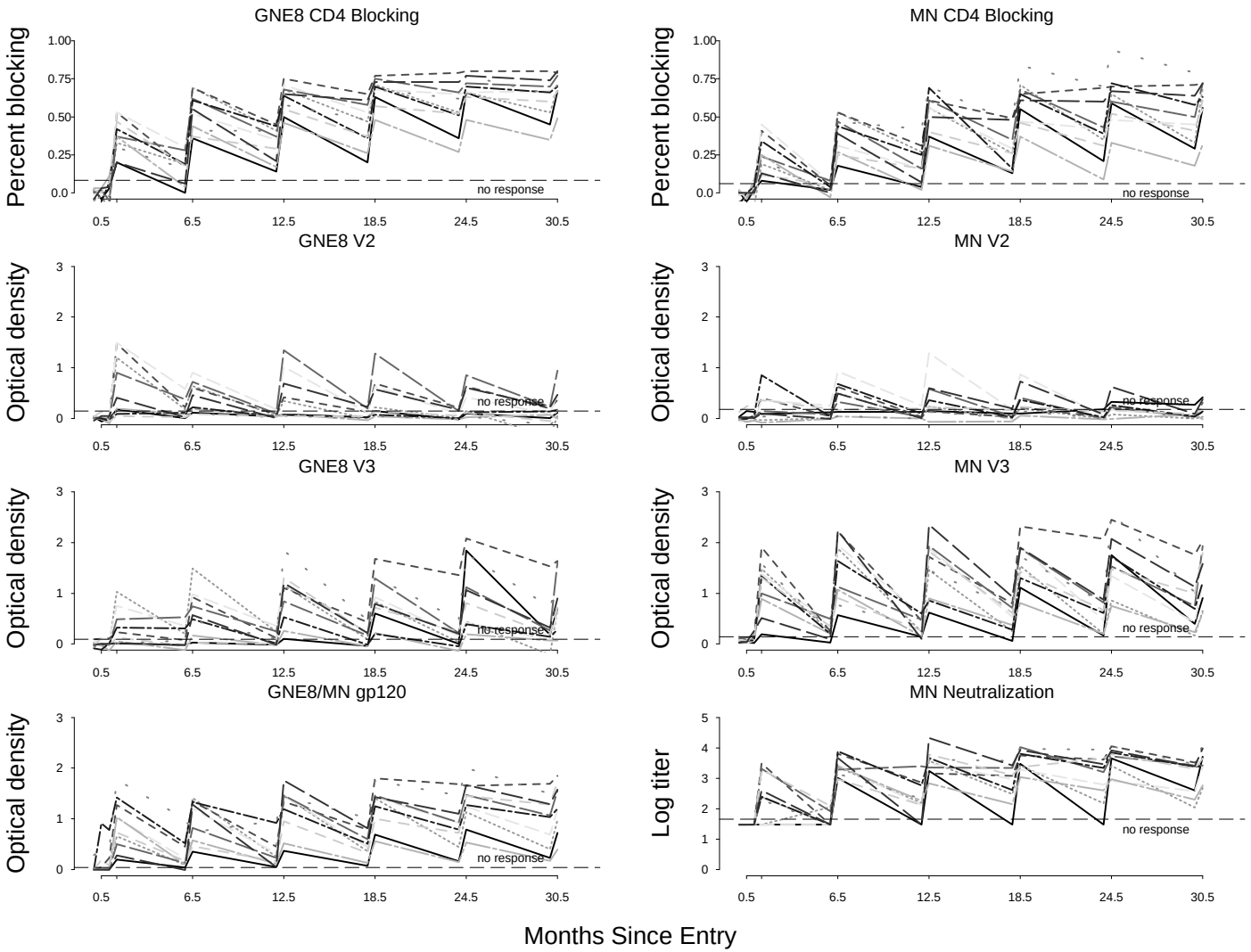
  where $\epsilon \sim N(0, \sigma^2)$.
- After obtaining estimates $\hat{c} = (\hat{c}_0, \hat{c}_1)'$ and $\hat{\sigma}^2$, we then take a random draw of σ^{*2} from $\hat{\sigma}^2 \chi_{n+1}$, where n is the number of subjects included in the linear regression, and c^* and ε^* are random draws from $N(\hat{c}, \sigma^{*2}(A'A)^{-1})$ and $N(0, \sigma^{*2})$, respectively, where A is the design matrix of the linear regression.
- Finally, we fill in the missing value X_{ij} by $\hat{X}_{ij} = c_1^* + c_2^* X_{i1} + \varepsilon^*$. We construct 10 complete data sets following this procedure.

Outline

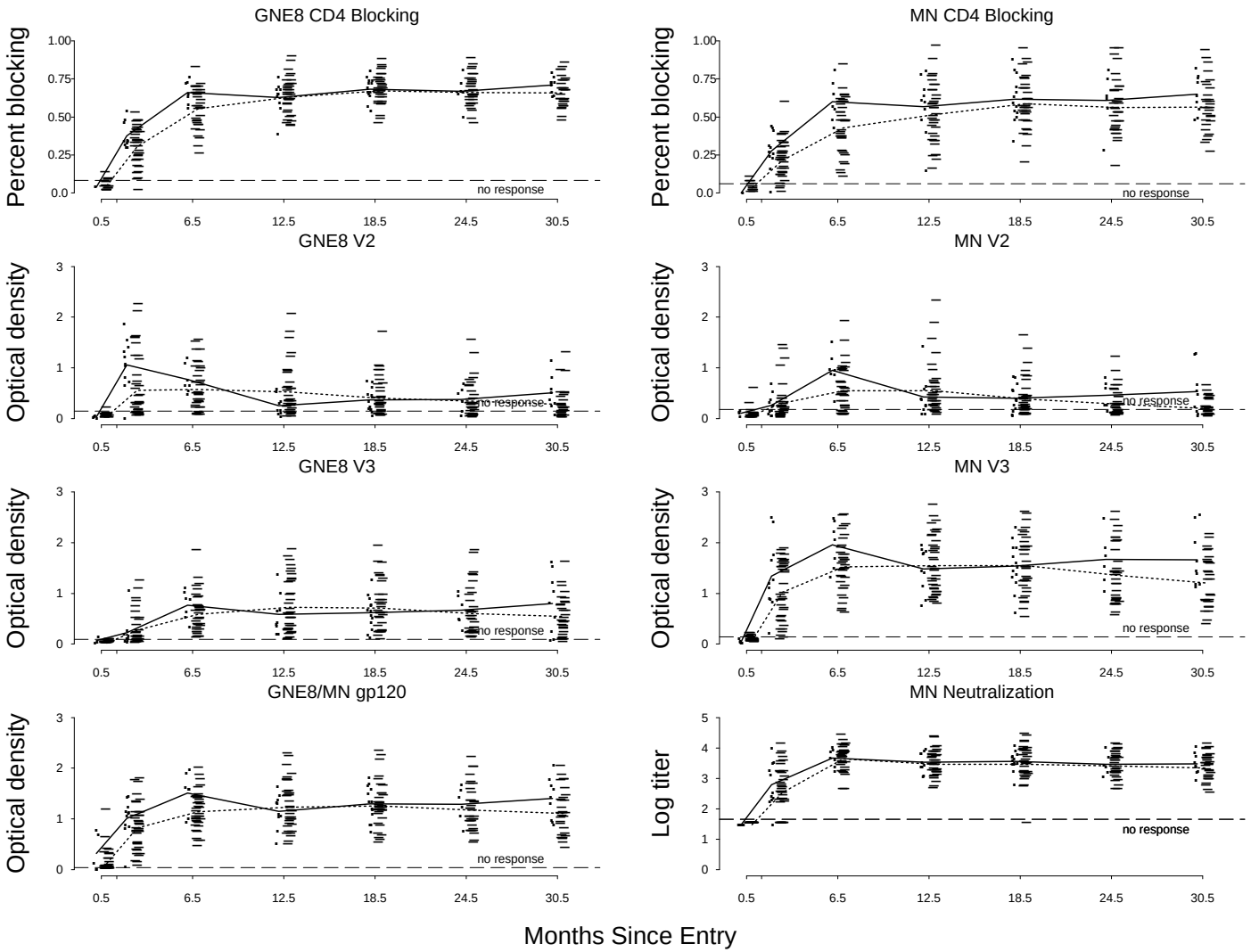
1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Example: Vax004 (Gilbert et al., 2005, JID)

- Randomly selected subject-specific antibody profiles



Peak Antibody Levels of Vaccinees (Solid/dotted = Uninfected/infected)



Tests for Different Antibody Levels, Uninfected vs Infected Vaccinees

- Wei-Johnson (1985, Biometrika) tests linearly combine Wilcoxon statistics across the 7 time-points
- Overall/aggregate tests of whether peak antibody levels differ between infected (n=239) and uninfected (n=163) vaccinees

Antibody Variable	Wei-Johnson p-value
MN CD4	0.074
GNE8 CD4	0.0045
MN V2	0.13
GNE8 V2	0.18
MN V3	0.21
GNE8 V3	0.031
MN/GNE8 gp120	0.39
MN Neutralization	0.60

Two Types of Correlates Questions

- For Vax004 and generally for efficacy trials, two types of correlates questions are of interest:
  1. If and how does the peak immune response level (e.g., 2 weeks after the last immunization) correlate with the subsequent rate of infection over a defined follow-up period?
  2. If and how does the immune response level near the time of exposure correlate with the rate of infection over a short follow-up period (e.g., until the next infection diagnostic test)?
- The first question is most useful for developing a surrogate endpoint (need something measured once near baseline)
- The second question is most useful for gaining insight into the mechanistic cause of protection (in theory immune response level at time of exposure is what matters)
- Both questions are interesting to ask, especially for vaccines with time-waning immune responses
- The following Vax004 results evaluate 'time-dependent' correlates

Results of Case-Cohort Cox Model Analysis

- Using continuous failure times (estimated time-to-HIV-acquisition), fit Prentice (1986) case-cohort Cox model with time-dependent antibody variables, using $\widehat{\alpha} = 174/3598 = 0.0484$

Antibody variable	*HR* of HIV infection by Ab Quartile Q1	Q2	Q3	Q4	P-value for difference	P-value for trend
MN CD4	1.0	0.45	0.39	0.33	0.008	0.009
GNE8 CD4 Binding	1.0	0.46	0.37	0.30	0.026	0.013
MN V2	1.0	1.56	0.95	0.88	0.044	0.17
GNE8 V2	1.0	0.72	0.66	0.49	0.052	0.009
MN V3	1.0	0.88	0.59	0.84	0.22	0.39
GNE8 V3	1.0	0.45	0.53	0.40	0.011	0.030
MN/GNE8 gp120	1.0	0.96	0.69	0.68	0.30	0.096
MN Neutralization	1.0	0.52	0.42	0.46	0.080	0.088

Interpretation of Vax004 Results

- MN CD4 blocking, GNE8 CD4 blocking, GNE8 V2, GNE8 V3, MN Neutralization responses inversely correlated with HIV infection rate
- Estimated VE negative for low responses, $\approx$ zero for medium responses, positive for high responses
- Two possible explanations
  - High antibody levels cause protection and low antibody levels cause increased susceptibility [**Causation Hypothesis**]
  - Antibody levels mark individuals by their intrinsic risk of infection [**Association Hypothesis**]
- Methods for evaluating a specific-SoP (the higher, second tier) are needed to discriminate between these possible explanations
  - Addressed in the other talks of the workshop

VaxGen Analysis with Discrete Failure Times

- 131 vaccine recipients became HIV infected by month 36.
- 277 uninfected vaccine recipients (254 censored at month 36 and 23 censored at an earlier visit time) were sampled using stratified sampling from 5 different strata.
- Peak GNE8 CD4 avidity levels (immune response) were measured at month 6.5 and the time interval in which the infection occurs for all infected subjects, and at month 6.5 and a randomly chosen time interval for selected uninfected subjects.

VaxGen Analysis with Discrete Failure Times

Table: Estimated log relative hazards (RHs) of HIV infection in the vaccine trial.

	$(\text{Antibody})^{1/5}$	White	Sex	Med. RS	High RS
$\log(\text{RH})$	-1.56	-0.11	-1.41	1.27	1.14
95% CI	(-2.35, -0.76)	(-0.66, 0.45)	(-3.58, 0.76)	(0.74, 1.79)	(0.57, 1.72)
P value	0.001	0.708	0.202	0.000	0.000

White: 1 for white, 0 for nonwhite
Sex: 1 for female, 0 for male
Medium risk group (Med. RS): risk score is equal to 2 or 3
High risk group (High RS): risk score is greater than 3

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Some Sampling Questions to Consider Further

- Prospective or retrospective sampling?
- How much of the cohort to sample?
- Sampling design: Which subjects to sample?

Prospective or Retrospective Sampling?

Prospective sampling: Select a random sample for immunogenicity measurement **at baseline**

- Advantages of prospective sampling
  - Can estimate case incidence for groups with certain immune responses
  - Can study correlations of immune response with multiple study endpoints
  - Practicality: The lab will know what subjects to sample as early as possible, and there is one simple subcohort list

Prospective or Retrospective Sampling?

Retrospective sampling: At or after the final analysis, select a random sample of controls for immunogenicity measurement

- Advantages of retrospective sampling
  - Can match controls to cases to obtain balance on important covariates
    - E.g., balanced sampling on a prognostic factor gains efficiency (balanced sampling = equal number of subjects sampled within each level of the prognostic factor for cases and controls)
  - Can flexibly adapt the sampling design in response to the results of the trial
    - E.g., Suppose the results indicate an interaction effect, with VE $>>$ 0 in a subgroup and $VE \approx 0$ in other subgroups. Could over-sample controls in the 'interesting' subgroup.

Retrospective sampling may also sample controls at periodic intervals during the study follow-up period

Prospective or Retrospective Sampling?

For cases where there is one primary endpoint and it is not of major interest to estimate absolute case incidence, retrospective sampling may be typically preferred

How Many Controls to Sample?

- In prevention trials, for which the clinical event rate is low, it is very expensive and unnecessary to sample all of the controls
  - E.g., VaxGen trial: 368 HIV infected cases; 5035 controls
  - **Rule of thumb:** A $K:1$ Control:Case ratio achieves relative efficiency of $1 - \frac{1}{1+K}$ compared to complete sampling

K	Relative Efficiency
1	0.50
2	0.67
3	0.75
4	0.80
5	0.83
10	0.91

Which Controls to Sample?

Two-Phase Sampling

- **Phase I:** All N trial participants are classified into K strata on the basis of information known for everyone: N_k in stratum k; $N = \sum_{k=1}^{K} N_k$
- **Phase II:** For each k, $n_k \leq N_k$ subjects are sampled at random, without replacement from stratum k, and 'expensive' information (i.e., the immunological biomarker S) is measured for the resulting $n = \sum_{k=1}^{K} n_k$ subjects

Which Controls to Sample?

Principle: Well-powered CoR evaluation requires broad variability in the biomarker response and in the risk of the clinical endpoint

- Can improve efficiency by over-sampling the "most informative" subjects
  - Disease cases (usually sampled at 100%)
  - Rare or unusual immune responses; or rare covariate patterns believed to affect immune response (e.g., HLA subgroups)
- Baseline auxiliary data measured in everyone most valuable when they predict the missing data (i.e., the biomarker of interest)

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Measurement Error

Measurement error can reduce power to detect a CoR

Illustrative Example

- 'True' CoR $S^* \sim N(0,1)$
- 'Measured CoR' $S = S^* + \epsilon$, $\epsilon \sim N(0, \sigma^2)$
- Infection status Y generated from $\Phi(\alpha + \beta S^*)$

  with α set to give $P(Y = 1|S^* = 0) = 0.20$ and β set to give $P(Y = 1|S^* = 1) = 0.15$

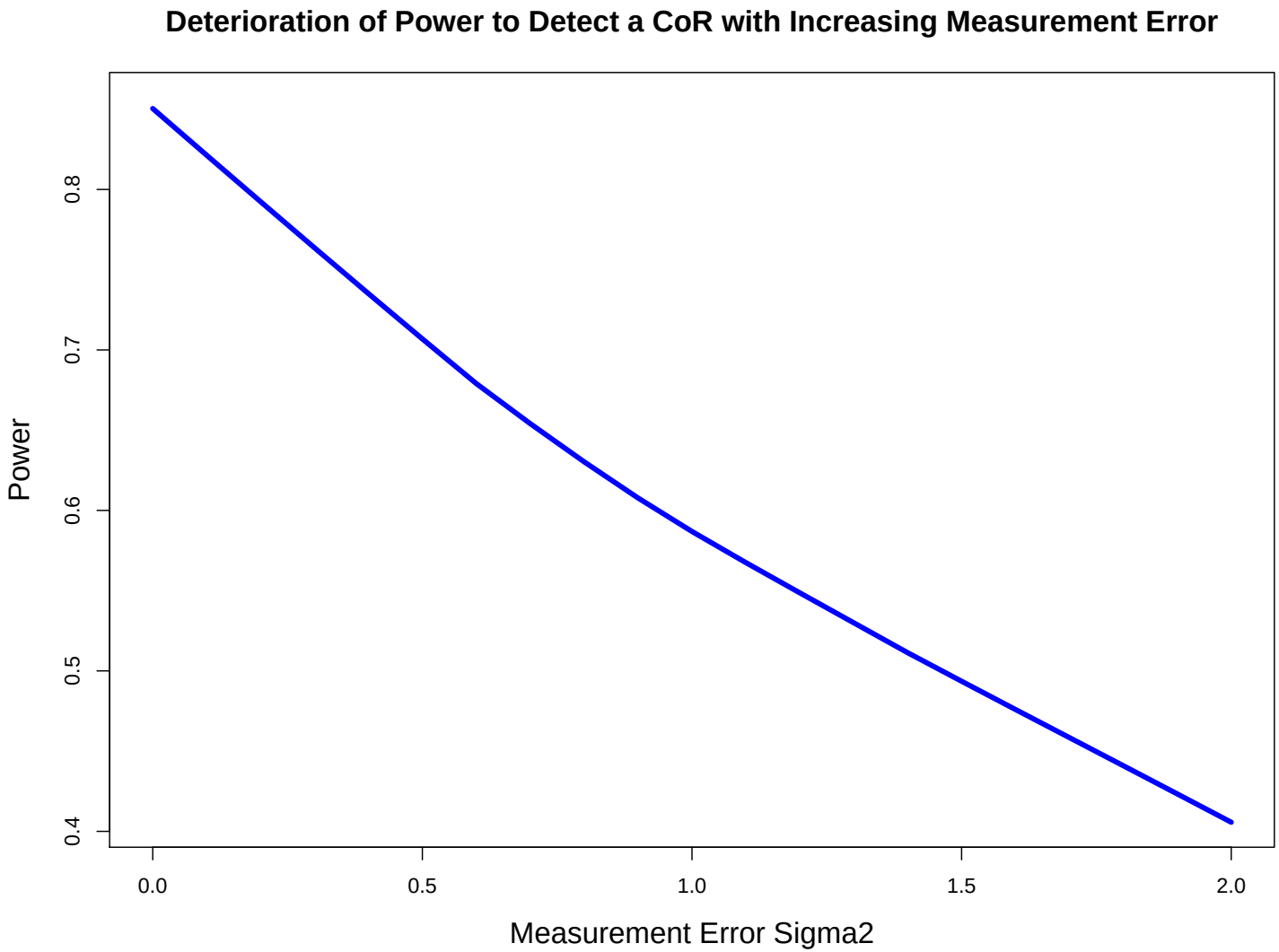
  σ^2 ranges from 0 to 2 (no-to-large measurement error)

Measurement Error Reduces Power

Simple Simulation Study

- Consider a study with $n = 500$ participants
- Consider power of a logistic regression model to detect an association between S and Y

Measurement Error Reduces Power



Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Design of CoR Assessment

- The HIV Vaccine Trials Network (HVTN) has conducted design exercises for the evaluation of immunological CoRs in randomized double-blinded efficacy trials of 1, 2, or 3 vaccine regimens versus placebo*
  - Trials designed with primary objective to assess VE for infections occurring in first 18 months of follow-up
    - 90% power to reject $H_0 : VE \leq 0\%$ in favor of $H_1 : VE = 40\%$ with 1-sided $\alpha = 0.025$
  - Trials designed with secondary objective to assess immunological CoRs in the vaccine arm(s)

*Gilbert, Grove, et al. (2011, *Statistical Communications in Infectious Diseases*)

Design of CoR Assessment

- For the planned trials the immunizations are administered during a 12 month period, and 6.5 months is approximately the 'peak immunogenicity time-point'
  - E.g., 2 weeks after the second protein immunization

- Thus the goal is to evaluate the association between month 6.5 immunological measurements and the subsequent rate of HIV infection in the vaccine arm(s)

Design of CoR Assessment

Sketch of Analysis Plan

- Analyze vaccine recipients HIV negative at 6.5 months
- Apply two-phase methods, for assessing a CoR for infections occurring in the window 6.5–24 months, and for assessing a CoR for infections occurring in the window 6.5–36 months
- Use a balanced frequency-matched stratified random sample of uninfected vaccine recipients
  - E.g., stratify by gender. Suppose 110 men and 80 women in the vaccine arm are infected. Then, with 5:1 sampling, sample 5*110 uninfected men and 5*80 women in the vaccine arm.

Qualification of Immunological Measurements for CoR Assessment

- An immunological measurement must meet requirements to be worth studying as a potential CoR
  - Require a greater response rate or level in vaccine group than placebo group
  - Require a low enough 'noise-ratio' of 'protection-irrelevant' vs 'protection-relevant' variability of the immune response variable in the vaccine group
    - Protection-irrelevant variability includes technical measurement error and hour-to-hour or day-to-day intra-subject variability
  - A lower noise-ratio implies greater plausibility that there exists a strong association between the measured immune response and infection
  - Power erodes as this ratio decreases (as seen earlier and seen again later)

Analysis for Multiple Vaccine Arms

- For trials with multiple vaccine arms, two analyses are complementary:
  1. Assess immune correlates for each vaccine arm separately
  2. Assess immune correlates pooling over the vaccine arms

- The pooled analysis provides greater sample size and greater heterogeneity in immune responses and hence greater statistical power
  - However this approach relies on the assumption that the immune correlate is common to the vaccine arms, which may or may not hold
  - Hence the rationale for both vaccine-specific and pooled evaluations

Power Calculations for Assessing a CoR

- The most influential factor for the power calculations is the number of HIV infections in the vaccine arm(s)

- The HVTN is considering trial designs with 2, 3, or 4 study arms (all with one placebo arm), and CoR assessment for:
  1. HIV infections diagnosed in $6.5 - 24$ months, where the analysis occurs when the last enrollee has 24 months of follow-up
  2. HIV infections diagnosed in $6.5 - 36$ months, where the analysis occurs when the last enrollee has 36 months of follow-up

Expected Number of Vaccine Recipients for Which To Measure Immune Responses

Sample Size for CoR Study During Period 6.5 – 24 Months

Number of Vaccine Arms	Enrollment of Vaccinees	Expected Number Infected Between 6.5 and 24 Months	Number Uninfected Vaccinee Controls (5:1 Ratio)	Total Number of Measurements
1	2150	53	265	318
2	4300	106	530	636
3	6450	159	795	954

Sample Size for CoR Study During Period 6.5 – 36 Months

1	2150	87	435	522
2	4300	174	870	1044
3	6450	261	1305	1566

Power Calculations for Assessing a CoR

- Power is calculated for evaluating as a CoR a normally distributed quantitative immunological measurement taken at Month 6.5
  - Specifically, compute power to detect a relative hazard of infection (RR) per 2 sd higher value of the observed immune response
  - Test

$$H_0 : RR = 1 \quad \text{vs} \quad H_1 : RR < 1$$

  - Use 1-sided $\alpha = 0.025$

Components of Variability

- The variability of the observed immune response is decomposed into protection-relevant (PR) and protection-irrelevant (PIR) variability

$$Var(\text{observedIR}) = Var^{PR}(IR) + Var^{PIR}(IR)$$

- Power results are shown for an underlying model where the infection rate in the vaccine arm decreases by the fraction RR per 2 sd increase in the protection-relevant variability of the immune response
- Power is computed for 4 scenarios of protection-irrelevant sd (noise)
  - Define the 'noise-ratio' as

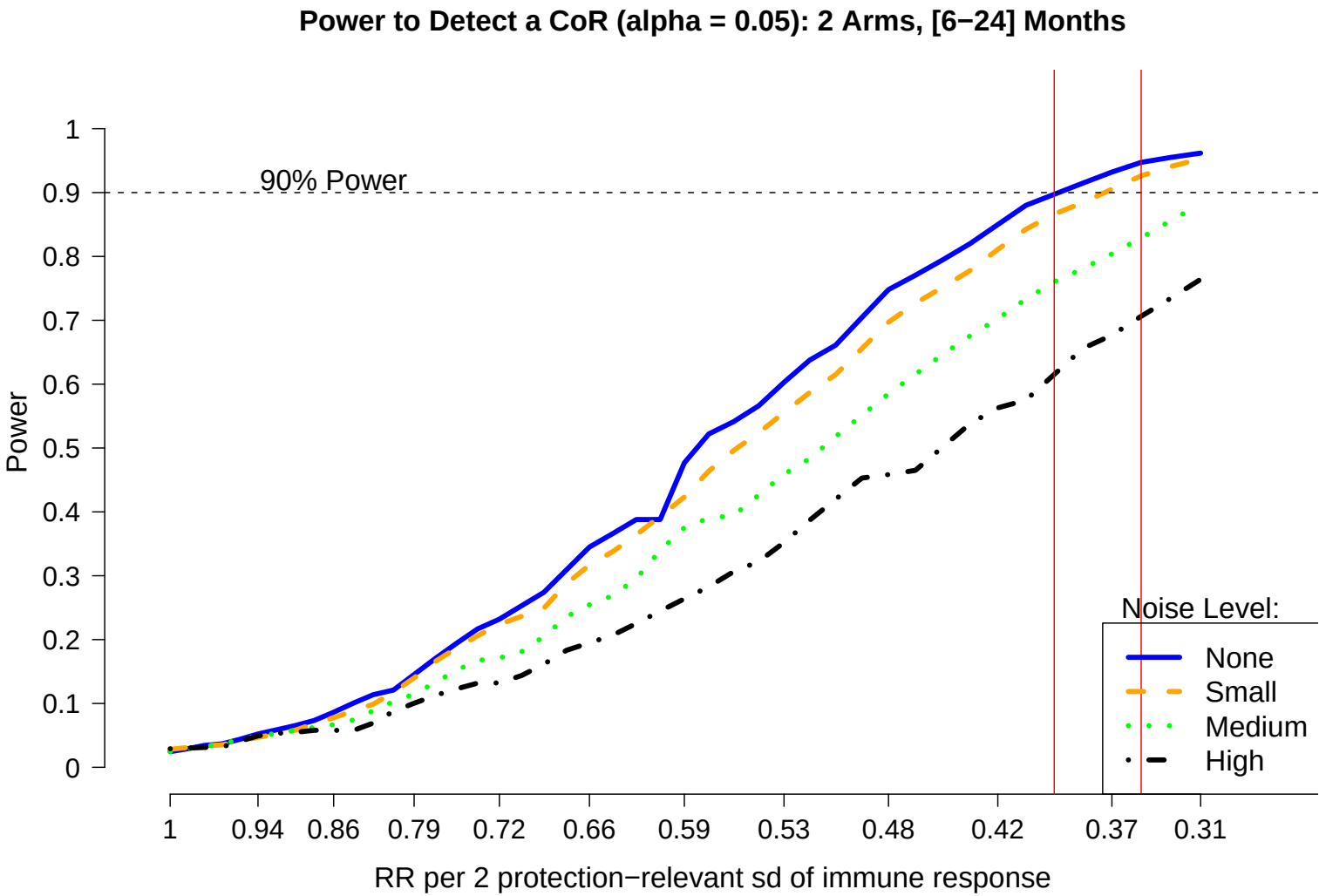
$$\text{noise} - \text{ratio} = Var^{PIR}(IR)/Var^{PR}(IR)$$

  - We evaluate measurements with noise-ratio = 0%, 33%, 67%, 100% (i.e., none, low, medium, high noise)
  - Equivalently, 100%, 90%, 69%, or 50% of the variability in the immune response is protection-relevant

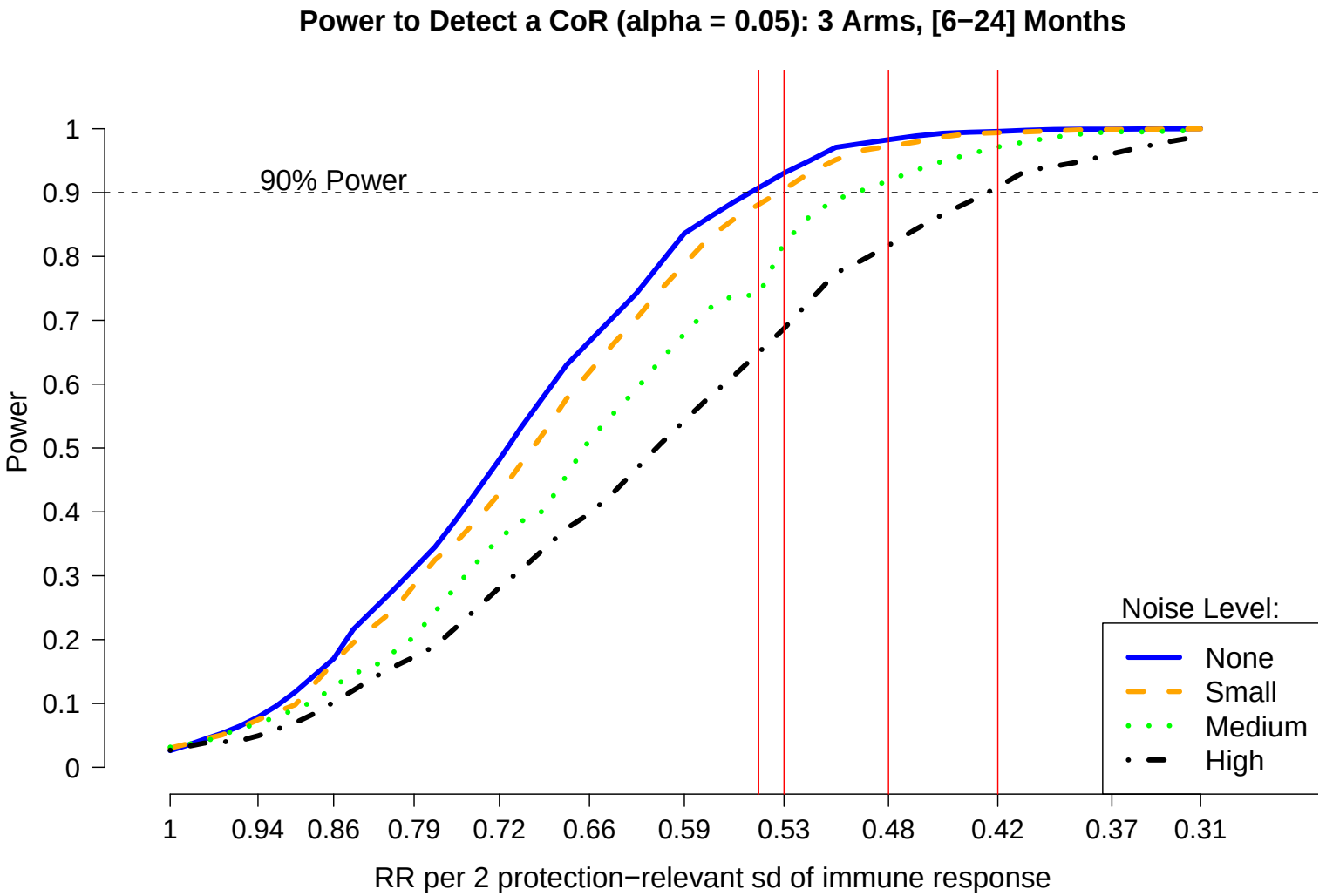
Data Analysis Method

- For testing $H_0 : RR = 1$ vs. $H_1 : RR < 1$, we apply a 1-sided Wald test from the stratified inverse-probability weighted Cox proportional hazards model (Borgan et al., 2000, Estimator II)
  - A D-estimator appropriate for retrospective selection of subjects for measuring the immune responses
- Given the observed set of infected vaccine recipients, take a random sample of uninfected vaccine recipients that gives a 5:1 uninfected:infected ratio
- Stratified sampling: Draw 5:1 random samples separately for men and women
- Implemented with cch (R code later)

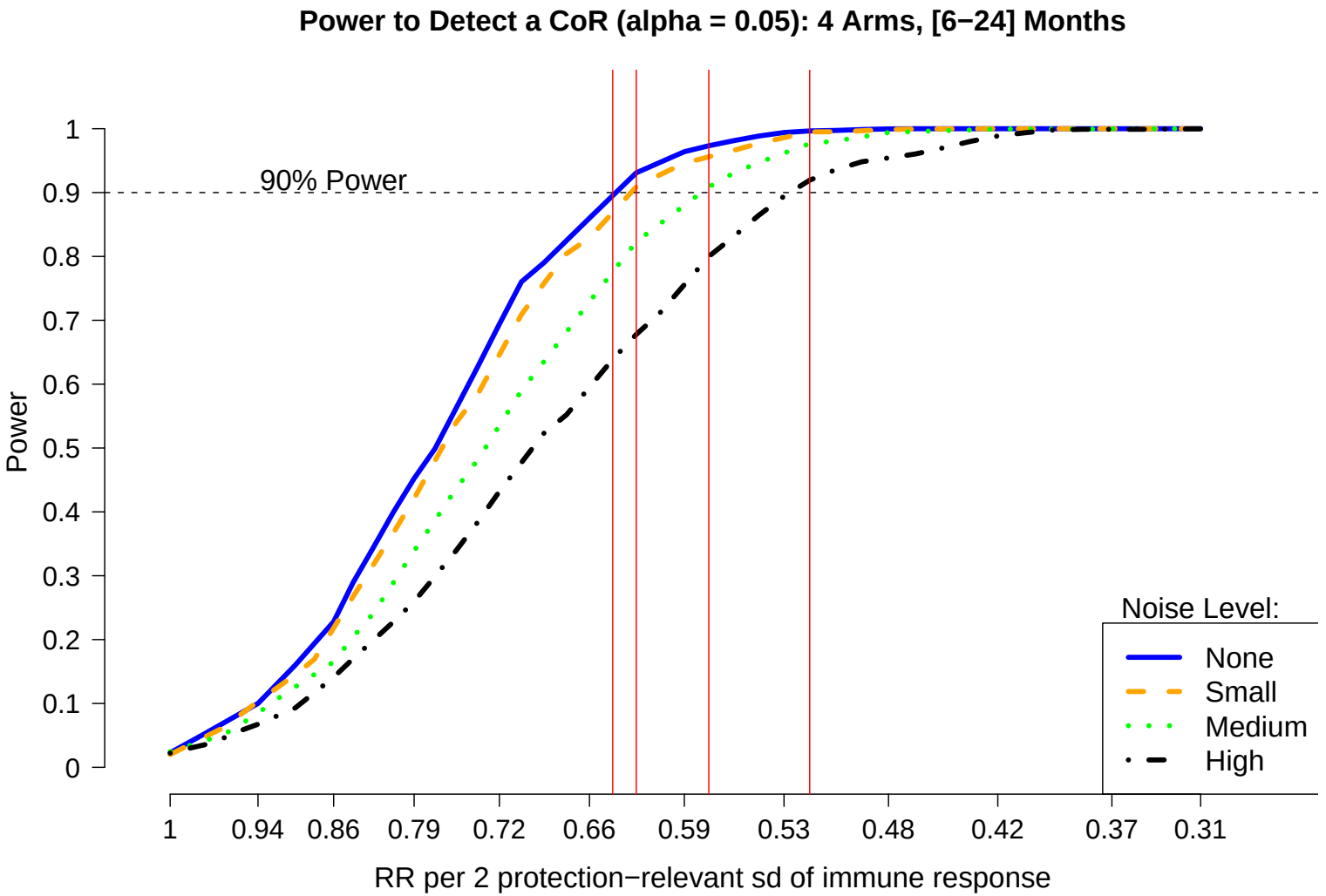
2 Arms; 6.5 – 24 Months



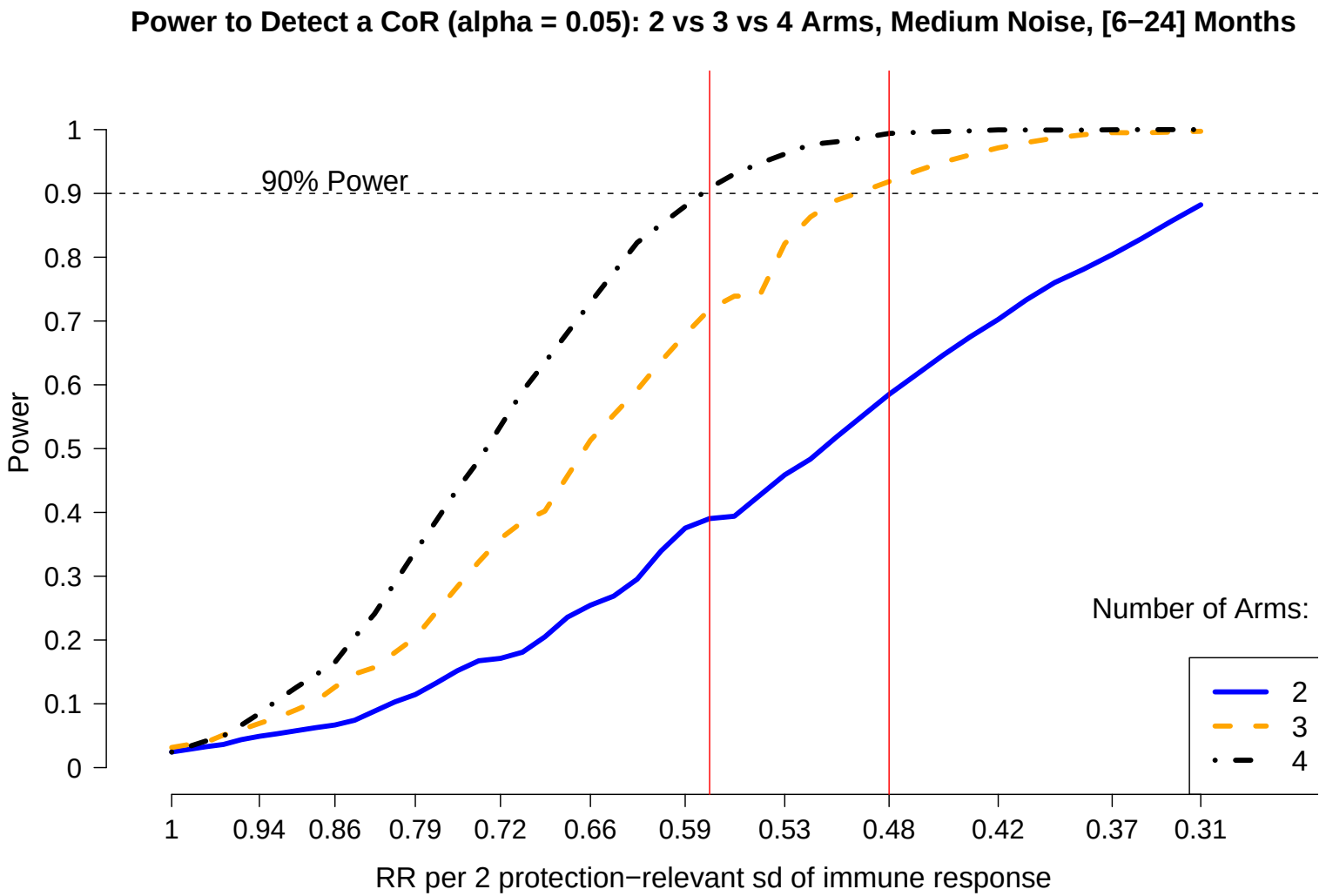
3 Arms; 6.5 – 24 Months



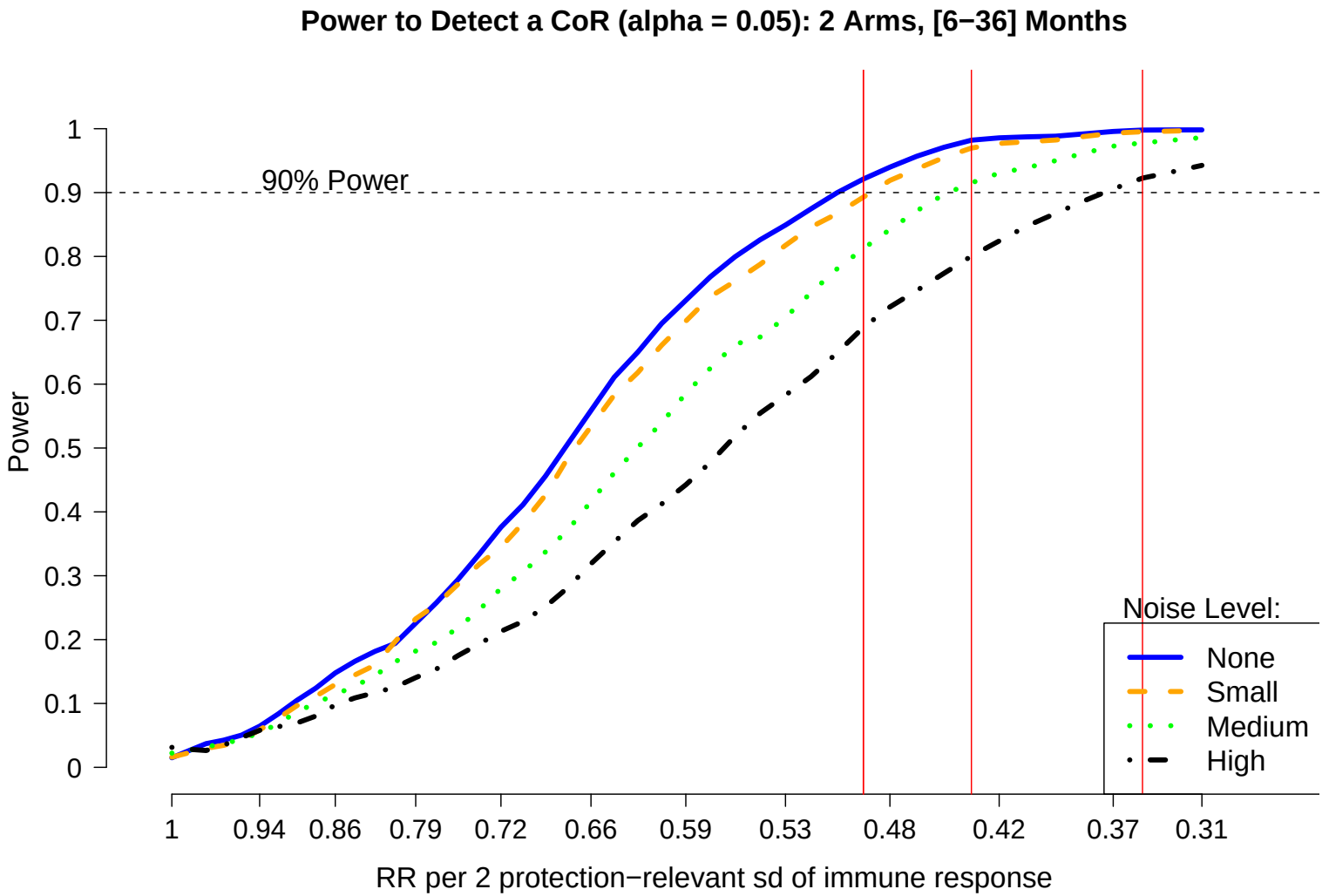
4 Arms; 6.5 – 24 Months



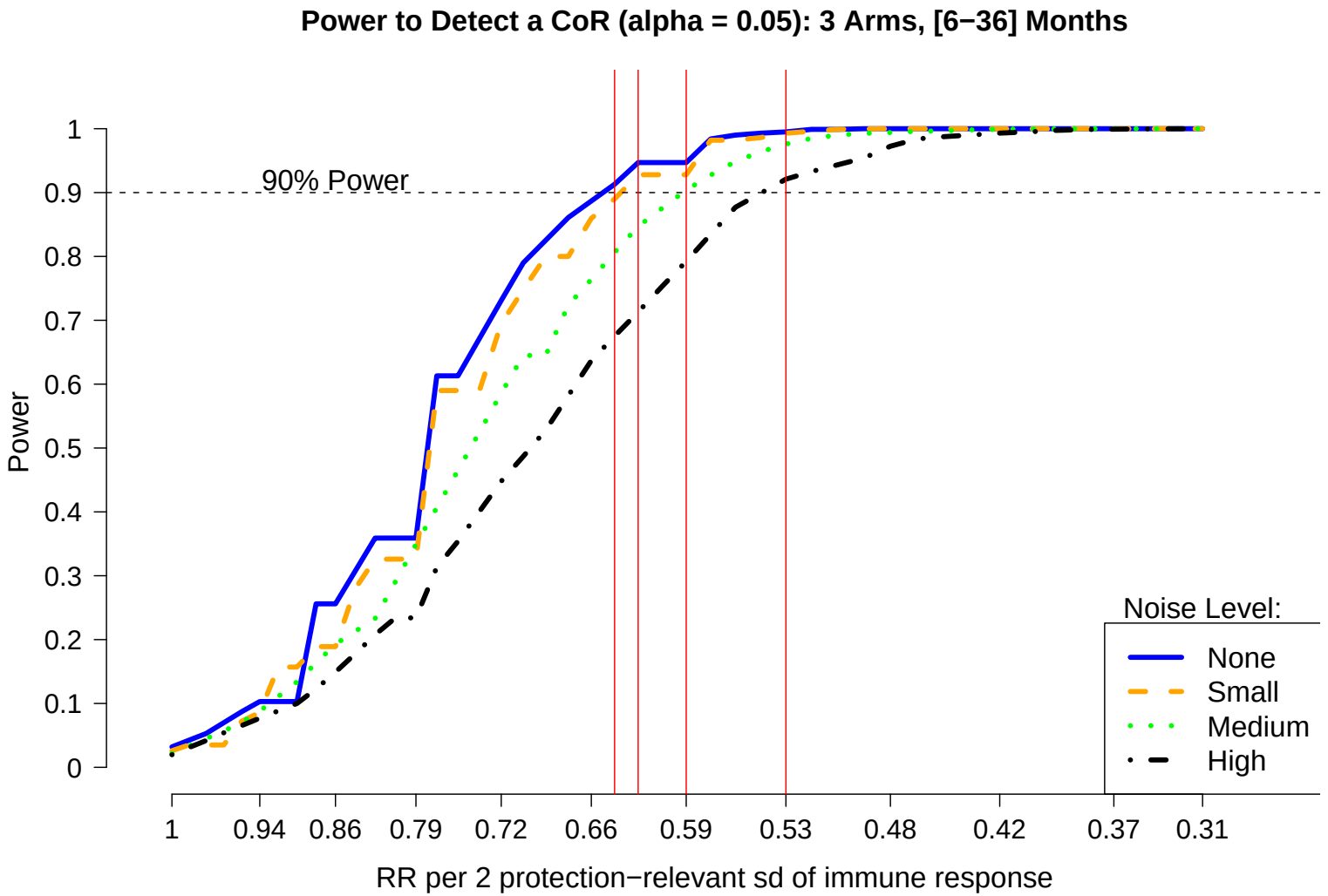
2 vs 3 vs 4 Arms; 6.5 – 24 Months; for Medium Noise Scenario



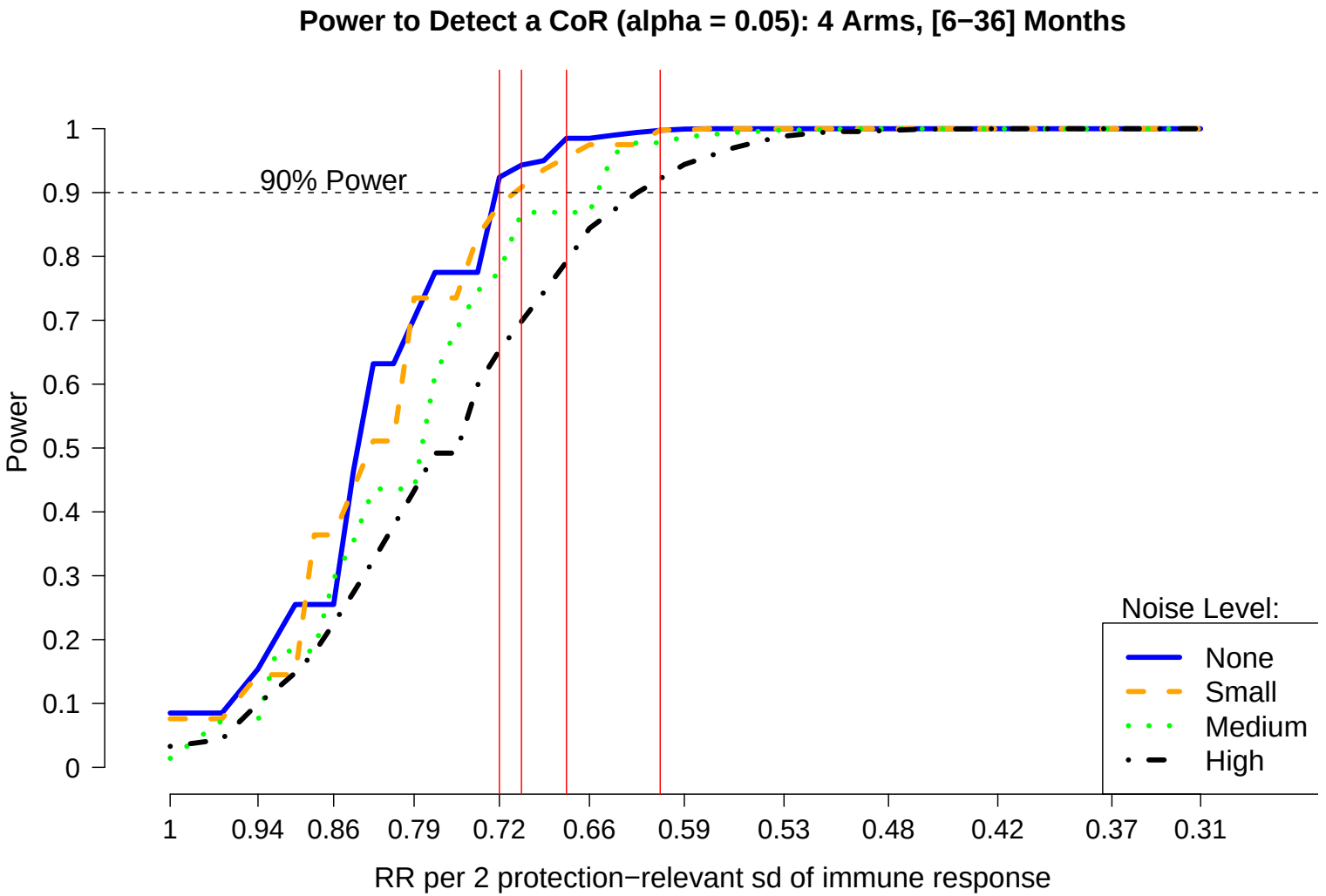
2 Arms; 6.5 – 36 Months



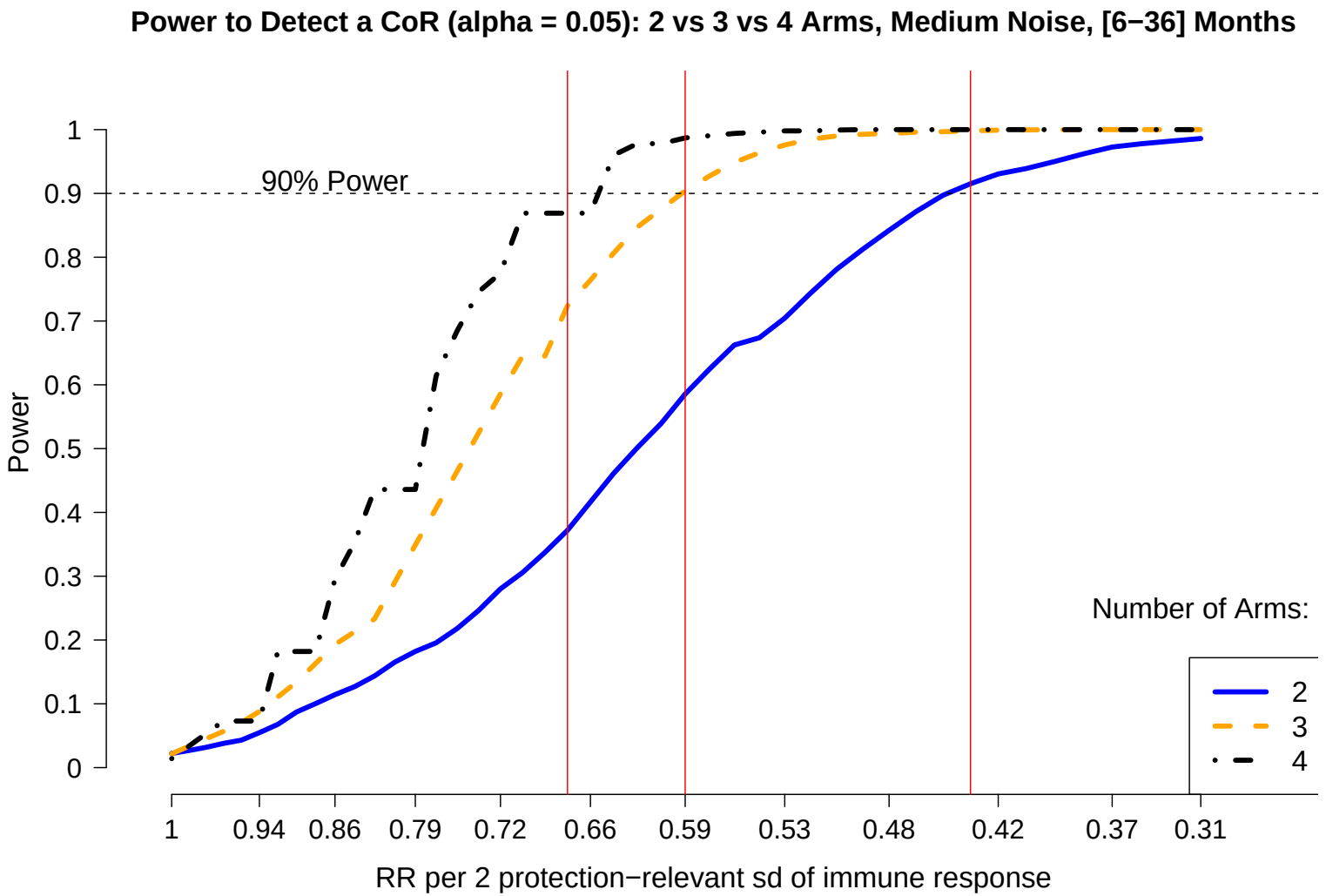
3 Arms; 6.5 – 36 Months



4 Arms; 6.5 – 36 Months



2 vs 3 vs 4 Arms; 6.5 – 36 Months; for Medium Noise Scenario



How to Apply the Power Curves

- For each immunological measurement that will be evaluated:
  1. Determine a plausible effect size in terms of RR per 2 sd change in protection-relevant variability of the immune response
  2. Determine the noise-ratio, or at least a range for plausible noise-ratios
  3. Read from the plot the power available
  4. If under-powered and assessment of a CoR is a priority, then consider a larger trial and iterate the power calculations

Summary of Power Calculations Exercise

- In 2-arm study (1 vaccine arm), low power to detect a CoR for infections in $6.5 - 24$ months; and power is still fairly low for infections in $6.5 - 36$ months
- To make a 2-arm study well-powered for detecting an immune correlate, would need to enlarge the trial (for example, by powering it for a smaller VE than 40%)
- Increasing the number of vaccine arms improves power

Summary of Power Calculations Exercise

- If an assay has a high noise ratio, then power is low, and recommend to not use the assay
- Important to standardize the implementation of an assay as much as possible, to minimize the noise-ratio
  - E.g., use a single central lab, and the same operator to the extent possible
- Important to estimate the components of variability that cannot be eliminated (e.g., technical measurement error, hour-to-hour or day-to-day within-subject variability)
- Advanced planning essential; may systematically estimate the variability components from previous studies or from designed samples from the efficacy trial
  - RV144 case study (later)

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Typical Correlates Assessments are Inefficient

- Broadly in epidemiology studies, biomarker-disease associations are commonly assessed ignoring much data collected in the study
- That is, only subjects with the biomarker measured (i.e., the Phase II sample) are included in the analysis
- Standard case-cohort analyses use inverse probability weighting of the subjects sampled in Phase 2, including all of the methods discussed so far
- These ubiquitously-used methods are implemented in the R package cch (Breslow and Lumley)

Typical Correlates Assessments are Inefficient

- Breslow et al.* urge epidemiologists to consider using the whole cohort in the analysis of case-cohort data
- Baseline data on demographics and potential confounders are typically collected in all subjects (the Phase I data measured in everyone)
- These Phase I data are most valuable when they predict "missing" data

*Breslow, Lumley et al. (2009, American Journal of Epidemiology; 2009, Statistical Biosciences)

How Leverage All of the Data?

- **Question:** How can we use the Phase I data to improve the assessment of CoRs?
- **Answer:** Adjust the sampling weights used in the conventional analyses
- The following 20 slides borrow from Professor Norman Breslow's Plenary Lecture at the World Congress of Epidemiology in Porto Alegre, Brazil, September 23, 2008. This lecture closely tracks with Breslow et al. (2009, AJE).
- We will skip these slides for the sake of time, and encourage workshop participants to read Breslow et al. (2009)

Illustration: ARIC Case-Cohort Study*

- *N*=12,345 in main cohort, followed 6-8 yrs
  - Plasma collected at second visit (start of FU)
  - Free of CHD, transient ischemic stroke
- *n*=1,336 at Phase II (604 CHD, 732 controls)
  - Plasma assayed for C-reactive protein (CRP) and lipoprotein-associated pholpholipase A_2 (Lp-PLA_2)
- Focus on association of Lp-PLA_2 with CHD after adjustment for traditional risk factors

* Ballantyne CM et al. *Circulation* **109**:837-42, 2004

ARIC Case-Cohort Study

	Non CHD cases (controls)								CHD cases	Totals
Race	Black				White					
Sex	Female		Male		Female		Male			
Age	<55	≥55	<55	≥55	<55	≥55	<55	≥55		
Stratum (k)	1	2	3	4	5	6	7	8	9	
Cohort N_k	1,133	719	598	393	2,782	2,213	1,959	1,818	730	N=12,345
Sample n_k	59	54	42	71	88	154	117	147	604	n=1,336
Weights N_k/n_k	19.2	13.3	14.2	5.5	31.6	14.4	16.7	12.4	1.2	

Available Data

- X = variables in Cox regression model
  - time to development of CHD or time followed
  - main risk factor Lp-PLA$_2$ (known only at Phase II)
  - adjustment variables: age, sex, race, SBP, DBP, HDL-C, LDL-C, …
- V = variables known for entire cohort
  - used to stratify Phase II sampling or adjust the weights
  - includes adjustment variables for ARIC CHD study
  - in general includes variables *not* in regression model

Horvitz-Thompson Estimator

- Inverse probability weighting (IPW) -- notation
  - ξ_i = 1 if i^{th} subject sampled at Phase II, 0 otherwise
  - π_i = known sampling probability

$$\pi_i = \frac{n_k}{N_k} \text{ if subject } i \text{ in stratum } k$$

- Probability model $P_{\theta,\eta}(X)$
  - θ = regression coefficients in Cox model
  - η = baseline hazard function

Horvitz-Thompson Estimator

- Solve IPW likelihood equations for $(\hat{\theta}_N, \hat{\eta}_N)$

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\xi_i}{\pi_i}\dot{\ell}_{\theta,\eta}(X_i) = 0 \quad (\dot{\ell}_{\theta,\eta}, \text{ scores for } \theta)$$

$$\frac{1}{N}\sum_{i=1}^{N}\frac{\xi_i}{\pi_i}B_{\theta,\eta}h(X_i) = 0 \quad (B_{\theta,\eta} \text{ is score operator})$$

$h \in \mathcal{H}$, directions from which η may approach η_0

- Yields Barlow's (1994) method of analysis of case-cohort data and "robust" variance
  - solve weighted partial-likelihood equations
  - most common method in current use

Sampling Properties of HT Estimator*

$$\sqrt{N}(\hat{\theta}_N - \theta_0) = \sqrt{N}(\tilde{\theta}_N - \theta_0) + \sqrt{N}(\hat{\theta}_N - \tilde{\theta}_N)$$

$$\approx \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tilde{\ell}_0(X_i) + \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \left(\frac{\xi_i}{\pi_i} - 1 \right) \tilde{\ell}_0(X_i)$$

$$\text{Var}_{\text{TOT}} = \text{Var}_{\text{PHASE I}} + \text{Var}_{\text{PHASE II}}$$

- $\tilde{\theta}_N$ is **unobserved** MLE based on complete data
- $\tilde{\ell}_0$ is semi parametric **efficient influence function**
- $\text{Var}_{\text{PHASE II}}$ is **design based**: normalized error in HT estimation of unknown finite pop. total $\tilde{\ell}_{\text{TOT}} = \sum_{i=1}^{N} \tilde{\ell}_0(X_i)$
- Phase I and Phase II contributions approx. independent

* Breslow & Wellner, *Scandinav J Statist*, 2007/8, and others

22

Two Components of Variance

- **Phase I** variance represents usual uncertainty in generalizing results for N cohort subjects to target population
  - only variance if complete data for all
  - cannot be reduced by adjustment of weights
- **Phase II** variance represents additional uncertainty from not having complete data for all N cohort members, but only for n
  - can be reduced by adjustment of weights

Improving Efficiency: Survey Techniques

- Construct auxiliary vars $\tilde{V} = \tilde{V}(V)$ to **adjust weights**
  - Post-stratification (finer than needed for biased sampling)
  - Calibration (generalized raking: Deville, Särndal, Sautory, *JASA* 93)
  - Estimation using correct parametric model: $\pi_i = \pi(\tilde{V}_i; \alpha)$

    (Robins, Rotnitzky, Zhao, *JASA* 94)
- One possibility for auxiliary variables $\tilde{V}$
  1. Impute missing X values using parametric model $[X|V]$
  2. Fit model $P_{\theta,\eta}(X)$ to main cohort using imputed data
  3. Construct $\tilde{V}_i$ as "delta-betas" for above model
     - surrogates for unknown $\tilde{\ell}_0(X_i)$
- Estimate θ using adjusted weights based on $\tilde{V}_i$

Calibration *vs* Estimation of Sampling Weights*

- Weights **calibrated** to Phase I totals solve

$$\sum_{i=1}^{N} \xi_i w_i \tilde{V}_i = \sum_{i=1}^{N} \tilde{V}_i$$

- Weights **estimated** using Phase I variables solve

$$\sum_{i=1}^{N} \xi_i \tilde{V}_i = \sum_{i=1}^{N} \tilde{V}_i / w_i$$

- For auxiliary stratum indicators, $\tilde{V}_i$ binary indicators of stratum membership, two sets of weights agree

$$w_i = \frac{1}{\pi_i} = \frac{N_j}{n_j} \quad \text{for} \quad i \in \text{stratum } j$$

- Both sets converge to true π_i^{-1} in large samples

* after Lumley (2007)

Survey Package

- Implements adjustment of weights in weighted Cox regression analysis of stratified case-cohort data (and a whole lot more)
- See author Thomas Lumley's website http://faculty.washington.edu/tlumley/survey/
- Datasets and sample R code used for NWTS simulations reported below are at my site http://faculty.washington.edu/norm/software.html

Adjustment of Sampling Weights

- If variables used for calibration or estimation are the **only** variables in the probability model, then weighted estimate same as estimate from fit to main cohort and Phase II variance component is zero
- Illustrate by exploring relationship between Lp-PLA_2 and HDL-C with ARIC data

Association between Lp-PLA$_2$ and HDL-C: Standard Sampling Weights

HDL-C (mg/L)	Lp-PLA$_2$ (µG/L)			
	0-0.309	0.310-0.421	0.422-1	Total
< 40	701.4 (105.7)	938.6 (111.3)	1,561.3 (138.9)	3,201.3 (185.4)
40-59.0	1,764.4 (166.9)	2,310.2 (187.5)	2,175.3 (170.0)	6,249.9 (234.1)
≥ 60.0	1,569.6 (164.2)	909.4 (124.1)	414.8 (81.9)	2,893.8 (197.7)
Total	4,035.4 (217.3)	4,158.2 (222.2)	4,151.4 (206.1)	12,345

28

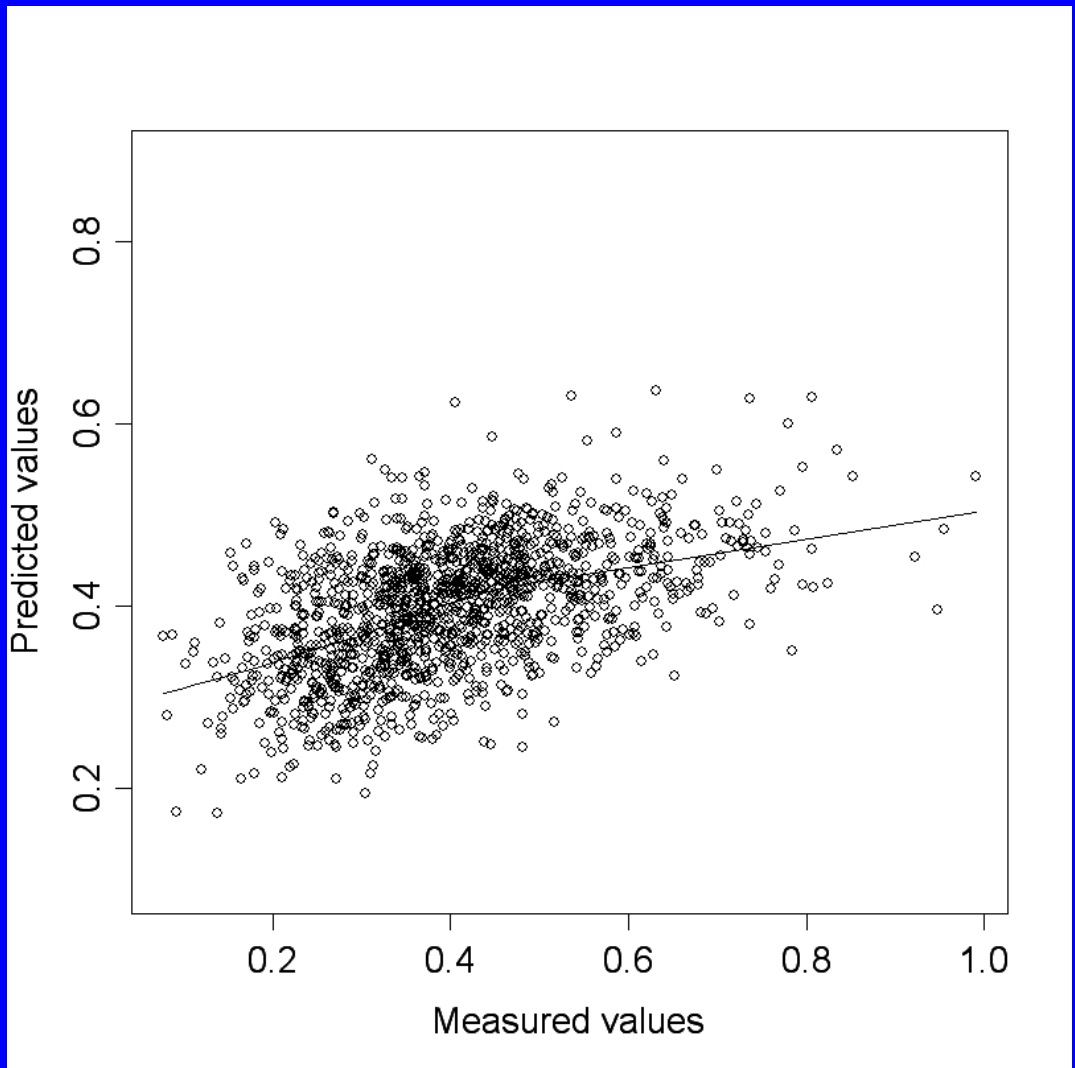
Association between Lp-PLA$_2$ and HDL-C: Calibrated (to HDL-C) Weights

HDL-C (mg/L)	Lp-PLA$_2$ (µG/L)			
	0.0.309	0.310-0.421	0.422-1	Total
< 40	739.0 (99.7)	988.9 (105.0)	1,645.0 (117.7)	3,373 (49.5)
40-59.0	1,665 (144.4)	2,180.1 (155.8)	2,059.9 (146.7)	5,898 (55.5)
≥ 60.0	1,667.4 (128.0)	966.0 (117.0)	440.6 (83.1)	3,074 (48.0)
Total	4,071.4 (212.9)	4,135.1 (220.3)	4,138.5 (201.8)	12,345

Horvitz-Thompson Estimation of Hazard Ratios by Lp-PLA$_2$ Tertile

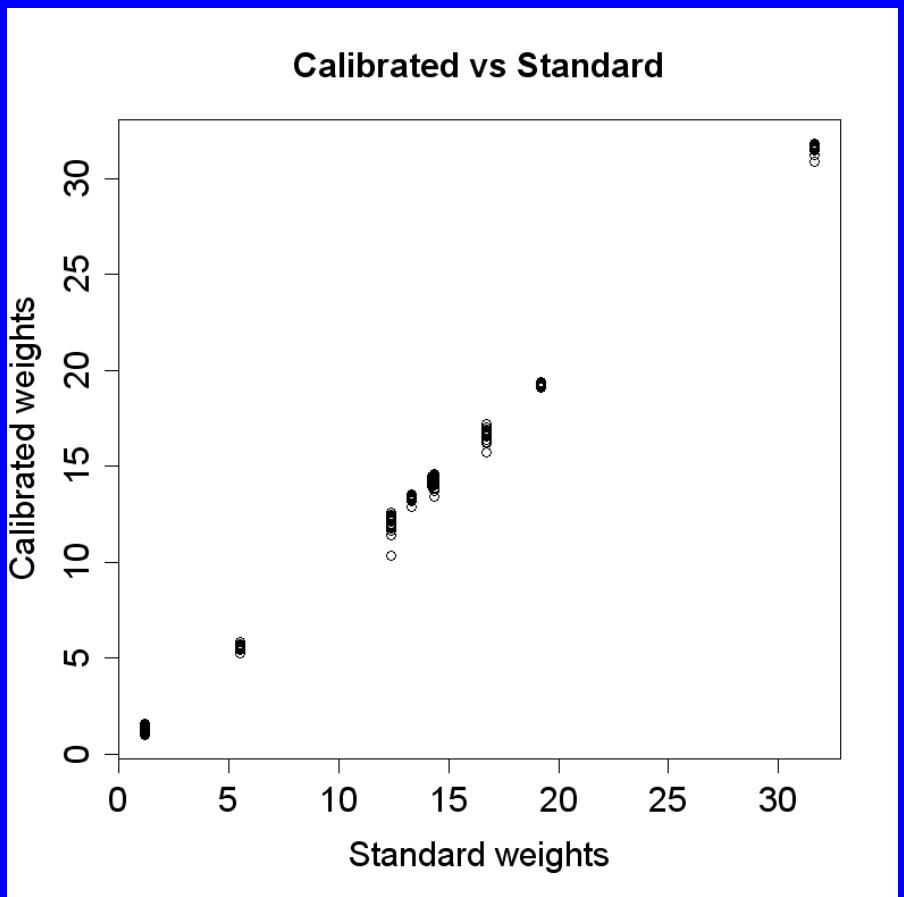
- Compare with Table 4, Model 2 of Ballantyne
  - results using standard weights virtually identical
- Prediction model for Lp-PLA$_2$: (weighted) linear regression on sex*race, LDL-C, HDL-C, SBP and DBP using Phase II data
- Impute values of Lp-PLA$_2$ for all in main cohort
- Fit Cox model to cohort using imputed values
- Extract "delta-betas" and use for calibration or estimation of weights

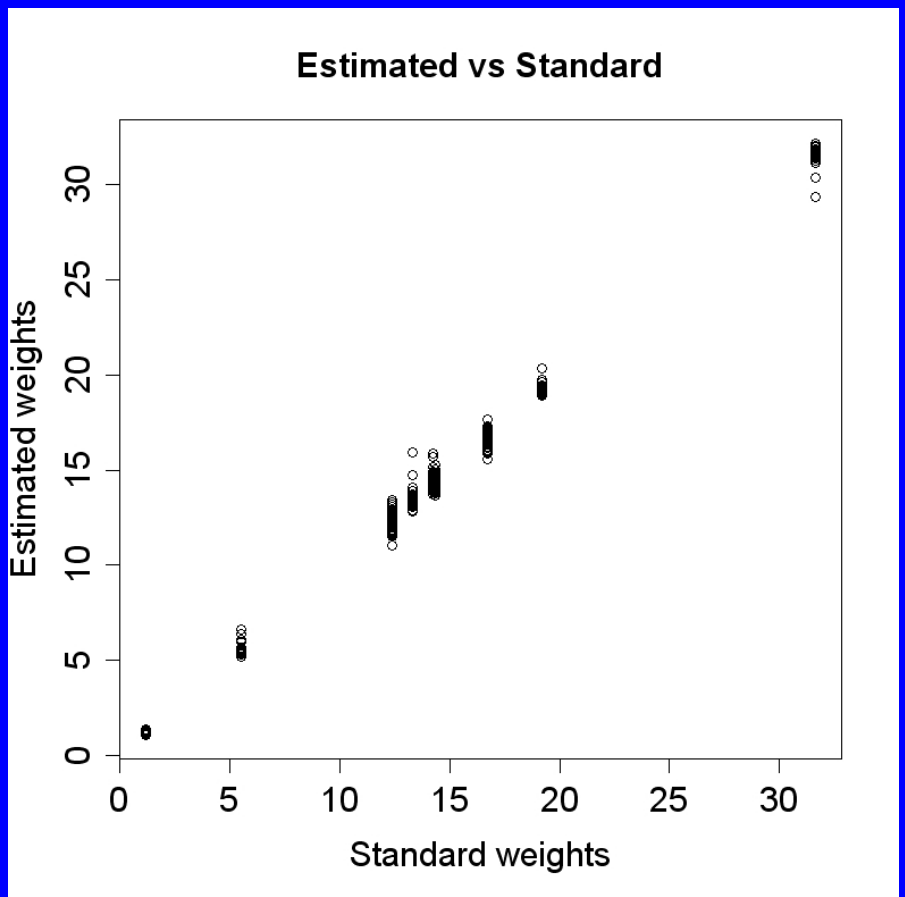
Prediction of Lp-PLA$_2$: $R^2 = 0.28$



31

Comparison of Weights





ARIC Case-Cohort Study of Lp-PLA$_2$ Results of Cox Regression Analyses

Model term*	Standard weights			Calibrated weights			Estimated weights		
	Coef.	**SE1**	SE2	Coef.	**SE1**	SE2	Coef.	**SE1**	SE2
Age/10	0.420	**.073**	.075	0.393	**.073**	.012	0.431	**.073**	.015
Male	0.762	**.088**	.091	0.791	**.088**	.019	0.742	**.088**	.022
White	0.037	**.098**	.090	0.159	**.099**	.016	0.101	**.100**	.029
Frmr smoker	-0.421	**.093**	.126	-0.464	**.092**	.017	-0.459	**.092**	.020
Never smoke	-0.552	**.099**	.129	-0.557	**.099**	.016	-0.622	**.099**	.020
SBP/100	1.554	**.207**	.267	1.539	**.208**	.046	1.580	**.207**	.048
LDL-C/100	0.777	**.106**	.151	0.786	**.106**	.045	0.748	**.107**	.048
HDL-C/100	-2.539	**.329**	.392	-2.361	**.329**	.052	-2.736	**.334**	.060
Diabetes	0.572	**.092**	.127	0.738	**.090**	.019	0.531	**.093**	.026
Lp-PLA$_2$ (2)*	0.052	**.110**	.126	0.054	**.111**	.127	0.050	**.111**	.127
Lp-PLA$_2$ (3)*	0.163	**.108**	.129	0.182	**.108**	.130	0.154	**.108**	.130

* 2nd and 3rd tertiles of lipoprotein-associated phospholipase A$_2$

ARIC Case-Cohort Study of Lp-PLA$_2$ Results of Cox Regression Analyses

Model term*	Standard weights			Calibrated weights			Estimated weights		
	Coef.	SE1	**SE2**	Coef.	SE1	**SE2**	Coef.	SE1	**SE2**
Age/10	0.420	.073	**.075**	0.393	.073	**.012**	0.431	.073	**.015**
Male	0.762	.088	**.091**	0.791	.088	**.019**	0.742	.088	**.022**
White	0.037	.098	**.090**	0.159	.099	**.016**	0.101	.100	**.029**
Frmr smoker	-0.421	.093	**.126**	-0.464	.092	**.017**	-0.459	.092	**.020**
Never smoke	-0.552	.099	**.129**	-0.557	.099	**.016**	-0.622	.099	**.020**
SBP/100	1.554	.207	**.267**	1.539	.208	**.046**	1.580	.207	**.048**
LDL-C/100	0.777	.106	**.151**	0.786	.106	**.045**	0.748	.107	**.048**
HDL-C/100	-2.539	.329	**.392**	-2.361	.329	**.052**	-2.736	.334	**.060**
Diabetes	0.572	.092	**.127**	0.738	.090	**.019**	0.531	.093	**.026**
Lp-PLA$_2$ (2)*	0.052	.110	**.126**	0.054	.111	**.127**	0.050	.111	**.127**
Lp-PLA$_2$ (3)*	0.163	.108	**.129**	0.182	.108	**.130**	0.154	.108	**.130**

* 2nd and 3rd tertiles of lipoprotein-associated phospholipase A$_2$

34

ARIC Case-Cohort Study: Interaction of Lp-PLA$_2$ and SBP

Model term*	Standard weights			Calibrated weights			Estimated weights		
	Coef.	SE1	**SE2**	Coef.	SE1	**SE2**	Coef.	SE1	**SE2**
Lp-PLA$_2$(2)	0.137	.118	**.130**	0.139	.119	**.131**	0.138	.118	**.131**
Lp-PLA$_2$(3)	0.303	.121	**.132**	0.306	.122	**.131**	0.299	.121	**.131**
Lp-PLA-SBP	**-0.672**	**.204**	**.302**	**-0.681**	**.205**	**.274**	**-0.692**	**.205**	**.274**

* Results for adjustment variables not shown

Summary of ARIC Analyses

- Weak prediction of Phase II variable (Lp-PLA_2)
- No improvement in precision of main effect
- Dramatic improvement in precision of coefficients of adjustment variables
- Modest but significant improvement in precision of interaction between Phase I variable known for all and Phase II variables
  - Reduction of 10% in standard error
- **Adjustment of weights adds value to analysis**

Take Home Messages from Breslow et al., 2009

1. **Rule of thumb:** Obtain 'worthwhile' efficiency gain for the CoR assessment if baseline covariates can explain at least 40% of the variation in the immunological biomarker ($R^2 \geq 0.40$)
2. If interested in interactions (evaluation of whether a baseline covariate measured in everyone modifies the association of the biomarker and the clinical endpoint), can obtain worthwhile efficiency gain with a lower R^2
3. Even if no gain for the CoR assessment, will usually dramatically improve efficiency for assessing the associations of the Phase I covariates with outcome
4. Therefore it may often be the preferred method, and all practicing statisticians and epidemiologists should have the Breslow et al. method in their analytic toolkit
5. However, Breslow et al. (2009) currently only applies for a single immune response of interest measured at phase two, and does not handle a time-dependent immune response (serious practical limitationis that need more research to resolve)

Outline

1. Design of Vax004 for assessing immunological correlates of risk (CoRs)
2. Methods: Case-cohort sampling design Cox proportional hazards model
   - Continuous time
   - Discrete time
3. Application to Vax004
4. Key issues
   - Sampling design
   - Measurement error
   - Power calculations accounting for measurement error
5. Improved analysis method (Breslow et al., 2009)
6. R tutorial (cch and Breslow et al., 2009 method)

Outline of R Tutorial for Evaluating a CoR

- An HIV vaccine efficacy trial data-set and R code are posted at http://faculty.washington.edu/peterg/SISMID2013.html
- The data-set consists of pseudo-real data modeled after the Vax004 HIV vaccine efficacy trial discussed earlier, modified such that the vaccine efficacy is about 40%
  - The setting with partial vaccine efficacy is of greatest interest for evaluating a CoR and a specific SoP

2 Immune Responses to Evaluate as Potential CoRs/SoPs*

1. 50% titer for neutralization of HIV-1 MN (**MN Neuts**)
2. Level at which sera block the binding of HIV-1 GNE8 to soluble CD4 (**CD4 Blocking**)

- The 2 immune responses are measured at month 6.5

*Assays described in Gilbert et al. (2005, JID)

R Tutorial for Evaluating a CoR

- The goal of the tutorial is to provide hands-on experience with leading practical methods for case-cohort proportional hazards evaluation of a CoR, as implemented in cch and calibrated.weights.coxph, and to make practicable these tools so that you may apply them in your own research

Excercises and Questions

1. Assess MN neuts and CD4 blocking as CoRs adjusting for baseline covariates such as sex and behavioral risk score, using cch
2. Repeat the assessment using calibrated.weights.coxph, adding the 'infectivity assay' variable as an auxiliary for potentially improving efficiency
3. What impact does leveraging the whole Phase 1 data have on the CoR assessment?

- This data-set will be re-visited for assessing a specific SoP

More on Accessing the R Survey Package

- The R survey package was developed by Thomas Lumley and colleagues
- This package and some useful background materials themselves can be accessed via the Comprehensive R Archive Network (CRAN) as follows:

1. Navigate to the R Project home page: http/www.r-project.org/
2. Select a nearby CRAN mirror site (these are listed by country)
3. From the CRAN site select "Packages" under "Software"
4. Click on "S" and then on "survey"
5. Download the Reference manual "survey.pdf" and the vignette "Two-phase designs in epidemiology". It is best to install the package itself from within R following instructions.
6. Another relevant package available from the CRAN site is the NestedCohort package developed by Hormuzd Katki. This also contains a Reference Manual and a tutorial.

Some Results of R Case-Cohort Tutorial

- Results for Borgan estimator II

	Coef	HR	(95%	CI)	p
neut	-0.578	0.561	0.418	0.753	0.000
sex	0.582	1.790	0.566	5.665	0.312
white	0.569	1.766	0.841	3.709	0.125
riskscore	0.501	1.650	1.260	2.161	0.000

- Results for Breslow et al., calibrated weights

	Coef	HR	(95%	CI)	p
neut	-0.857	0.424	0.355	0.508	0.000
sex	0.431	1.539	0.764	3.101	0.23
white	0.061	1.063	0.693	1.630	0.78
riskscore	0.499	1.646	1.474	1.838	0.00

Breslow et al. provides much narrower confidence intervals for all coefficients, especially for Phase 1 covariates

Variance Estimates from Breslow et al.

- Phase 1 variances

	neut	sex	white	riskscore
neut	2.715825e-03	-0.003575300	-0.0002898980	7.394128e-05
sex	-3.575300e-03	0.105254478	0.0334805912	1.394625e-03
white	-2.898980e-04	0.033480591	0.0400124358	1.322364e-04
riskscore	7.394128e-05	0.001394625	0.0001322364	1.925909e-03

- Phase 2 variances

	neut	sex	white	riskscore
neut	0.0056388789	-0.0006799573	-0.0004831076	0.0003232830
sex	-0.0006799573	0.0225479072	0.0047606002	0.0009726579
white	-0.0004831076	0.0047606002	0.0076025710	-0.0006092518
riskscore	0.0003232830	0.0009726579	-0.0006092518	0.0012367984