Module 16:
Evaluating Vaccine Efficacy

Instructors:
Dean Follmann, Peter Gilbert, Erin Gabriel, Michael Sachs

Session 3: Introduction to Frameworks for Assessing Immune Correlates of Protection

Summer Institute in Statistics and Modeling in Infectious Diseases
University of Washington, Department of Biostatistics

Course materials at:
http://faculty.washington.edu/peterg/SISMID2017.html

July 24–26, 2017

Statistical Center for
HIV/AIDS Research & Prevention
SCHARP

FRED **HUTCHINSON**
CANCER RESEARCH **CENTER**
A LIFE OF SCIENCE

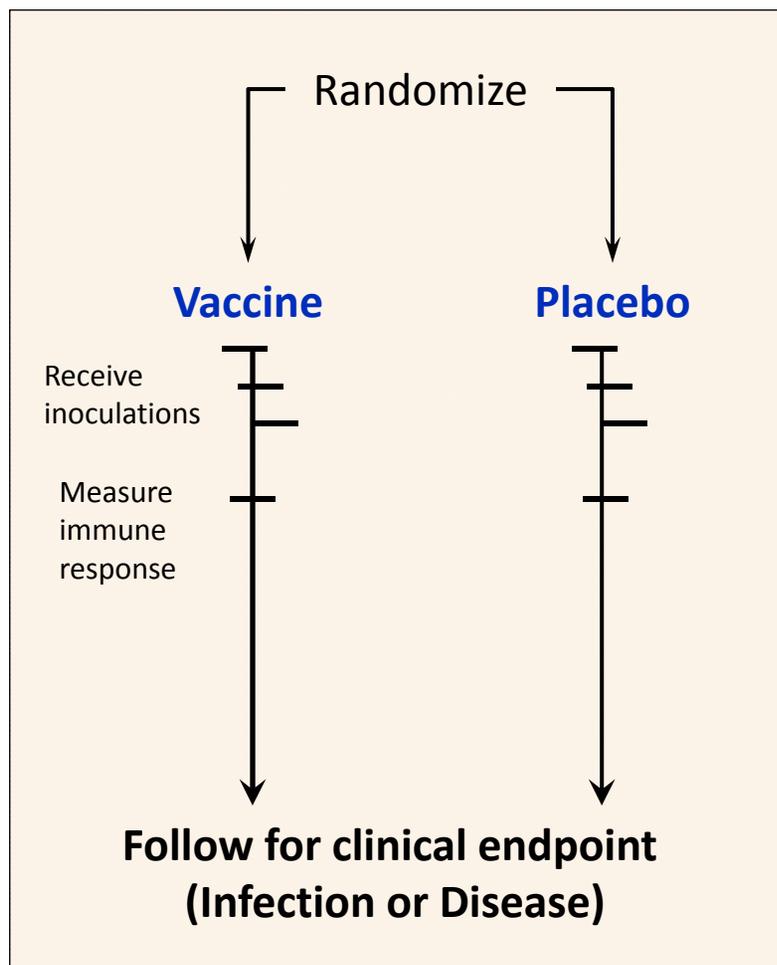# Outline of Module 8: Evaluating Vaccine Efficacy

Session 1 (Gabriel)   Introduction to Study Designs for Evaluating VE

Session 2 (Follmann)   Introduction to Vaccinology Assays and Immune Response

Session 3 (Gilbert)   **Introduction to Frameworks for Assessing Surrogate Endpoints/Immunological Correlates of VE**

Session 4 (Follmann)   Additional Study Designs for Evaluating VE

Session 5 (Gilbert)   Methods for Assessing Immunological Correlates of Risk and Optimal Surrogate Endpoints

Session 6 (Gilbert)   Effect Modifier Methods for Assessing Immunological Correlates of VE (Part I)

Session 7 (Gabriel)   Effect Modifier Methods for Assessing Immunological Correlates of VE (Part II)

Session 8 (Sachs)   Tutorial for the R Package *pseval* for Effect Modifier Methods for Assessing Immunological Correlates of VE

Session 9 (Gilbert)   Introduction to Sieve Analysis of Pathogen Sequences, for Assessing How VE Depends on Pathogen Genomics

Session 10 (Follmann)   Methods for VE and Sieve Analysis Accounting for Multiple Founders

Statistical Center for HIV/AIDS Research & Prevention
SCHARP

FRED **HUTCHINSON**
CANCER RESEARCH **CENTER**
A LIFE OF SCIENCE

Outline Session 3

- Introduction to immune correlates

- Prediction paradigm vs. mechanism of protection paradigm

- Frameworks for statistical assessment of immune response biomarkers as correlates of protection (CoPs)/surrogate endpoints

  1. Surrogate endpoint for the clinical endpoint providing reliable inferences about VE [valid replacement endpoint]

  2. Policy/predictors of VE [controlled effects] and mediators of VE [natural direct and indirect effects]

  3. Effect modifiers of VE [one or a few efficacy trials]

  4. Effect modifiers of VE [meta-analysis of a series of efficacy trials]

- Summary and conclusions

Preventive Vaccine Efficacy Trial

Randomize

Vaccine   Placebo

Receive inoculations

Measure immune response

Follow for clinical endpoint (Infection or Disease)

- **Primary Objective**
  - Assess **VE**: Vaccine Efficacy to prevent infection or disease with a pathogen

- **Secondary Objective**
  - Assess immune response biomarkers measured after vaccination as "immune correlates of protection" against infection or disease

Importance of an Immune Correlate

- Finding an immune correlate is a central goal of vaccine research
  - One of the 14 'Grand Challenges of Global Health' of the NIH & Gates Foundation (for HIV, TB, Malaria)
- Immune correlates useful for:
  - Shortening trials and reducing costs
  - Guiding iterative development of vaccines between basic and clinical research
  - Guiding regulatory decisions
  - Guiding immunization policy
  - **Bridging efficacy of a vaccine observed in a trial to a new setting**
    - ✓ Pearl (2011, *International Journal of Biostatistics*) suggests that bridging is the critical application

Regulatory Agencies Typically set Thresholds of Protection for Guiding Vaccine Licensure (this slide from Former FDA CBER Director, Dr. Norman Baylor)

Vaccine	Test	Correlate of Protection
Diphtheria	Toxin Neutralization	0.01-0.1 IU/mL
Hepatitis A	ELISA	10 mIU/mL
Hepatitis B	ELISA	10 mIU/mL
Hib Polysaccharides	ELISA	1 mcg/mL
Hib Conjugate	ELISA	0.15 mcg/mL
Influenza	HAI	1/40 dilution
Lyme	ELISA	1100 EIA U/mL
Measles	Microneutralization	120 mIU/mL
Pneumococcus	ELISA (Opsonophagocytosis)	0.20-0.35 mcg/mL (for children); 1/8 dilution
Polio	Serum Neutralization	1/4 - 1/8 dilution
Rabies	Serum Neutralization	0.5 IU/mL
Rubella	Immunoprecipitation	10-15 mIU/mL
Tetanus	Toxin Neutralization	0.1 IU/mL
Varicella	Serum Neutralization; gb ELISA	$\geq$ 1/64 dilution $\geq$ 5 IU/mL

Adapted from Plotkin S. Correlates of Vaccine Induced Immunity (Vaccines 2008:47)

Hard to Rigorously Identify Immune Correlates: Knowledge Level about Correlates for Licensed Vaccines

None/Low	Intermediate	High
1. Acellular Pertussis	1. Anthrax	1. Diphtheria & Tetanus Toxoids
2. BCG Live	2. Hepatitis B Recombinant	2. Haemophilus b Conjugate
3. Hepatitis A	3. Influenza Live	3. Meningococcal Polysaccharide Diphtheria
4. Japanese Encephalitis Invactivated	4. Measles Live	4. Rabies
5. Poliovirus Inactivated	5. Mumps Live	5. Tetanus & Diphtheria Toxoids
6. Rotavirus	6. MMR	6. Varicella Zoster Live
7. Rubella Live	7. Pneumococcal Polyvalent	7. Yellow Fever
8. Typhoid Live	8. Smallpox	
	9. Dengue	

But What Exactly is an Immune Correlate?

- Confusion in the meaning of the terms: "Immune correlate," "Correlate of protection," "Correlate of protective immunity"

- Generally "immune correlate" is connected to the concept of a surrogate endpoint, e.g. with definition:

  "A validated surrogate endpoint is an endpoint which allows prediction of a clinically important outcome."
  - International Conference on Harmonization, document E8

- Statistical methods for assessing the validity of surrogate endpoints are surprisingly subtle and not widely understood

- Many pitfalls for scientists to be misled about surrogate endpoints

Outline Session 3

- Introduction to immune correlates

- **Prediction paradigm vs. mechanism of protection paradigm**

- Frameworks for statistical assessment of immune response biomarkers as correlates of protection/surrogate endpoints

  1. Surrogate endpoint for the clinical endpoint providing reliable inferences about VE [valid replacement endpoint]

  2. Policy/predictors of VE [controlled effects] and mediators of VE [natural direct and indirect effects]

  3. Effect modifiers of VE [one or a few efficacy trials]

  4. Effect modifiers of VE [meta-analysis of a series of efficacy trials]

- Summary and conclusions

Two Major Concepts/Paradigms of Immune Correlates

- Causal agent paradigm (e.g., Plotkin, 2008, *Clin Infect Dis*)
  - **Causal agent of protection** = marker that **mechanistically causes** vaccine efficacy against the clinical endpoint

- Prediction paradigm (e.g., Qin et al., 2007, *J Infect Dis*)
  - **Predictor of protection** = marker that **reliably predicts** the level of vaccine efficacy against the clinical endpoint

- Both are extremely useful for vaccine development, but are assessed using different research techniques

- Statistical assessment mostly focuses on the prediction paradigm

A Predictive Correlate May or May Not be a Mechanism of Protection*

- **Informal Definition of an Immune Correlate:** An endpoint that can be used to reliably predict the vaccine effect on the clinical endpoint
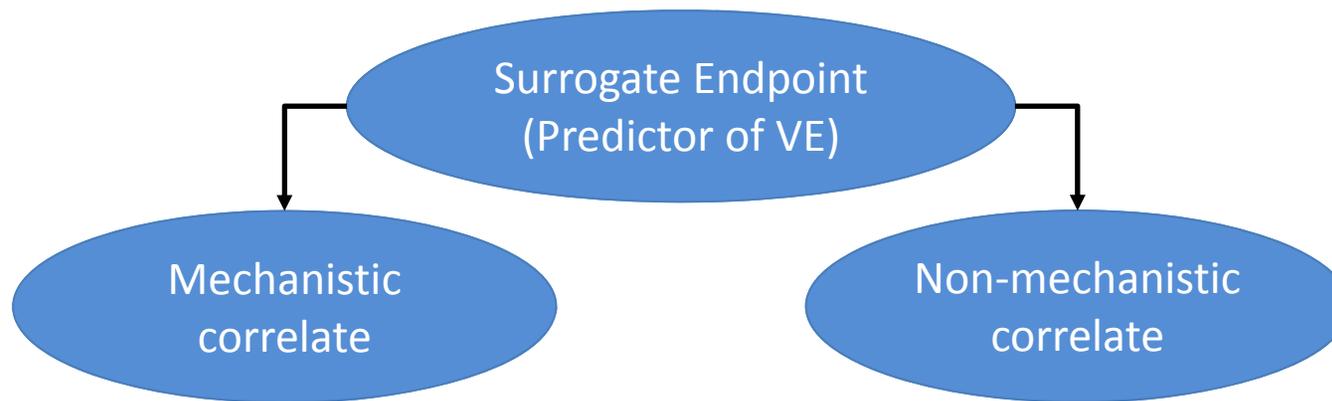
- **Example:** Meningococcal vaccine**
  - Mechanistic correlate: Bactericidal antibodies
  - Non-mechanistic correlate: Binding antibodies (ELISA)

* Plotkin and Gilbert (2012 *Clin Inf Dis*)
** Borrow et al. (2005, *Vaccine*)

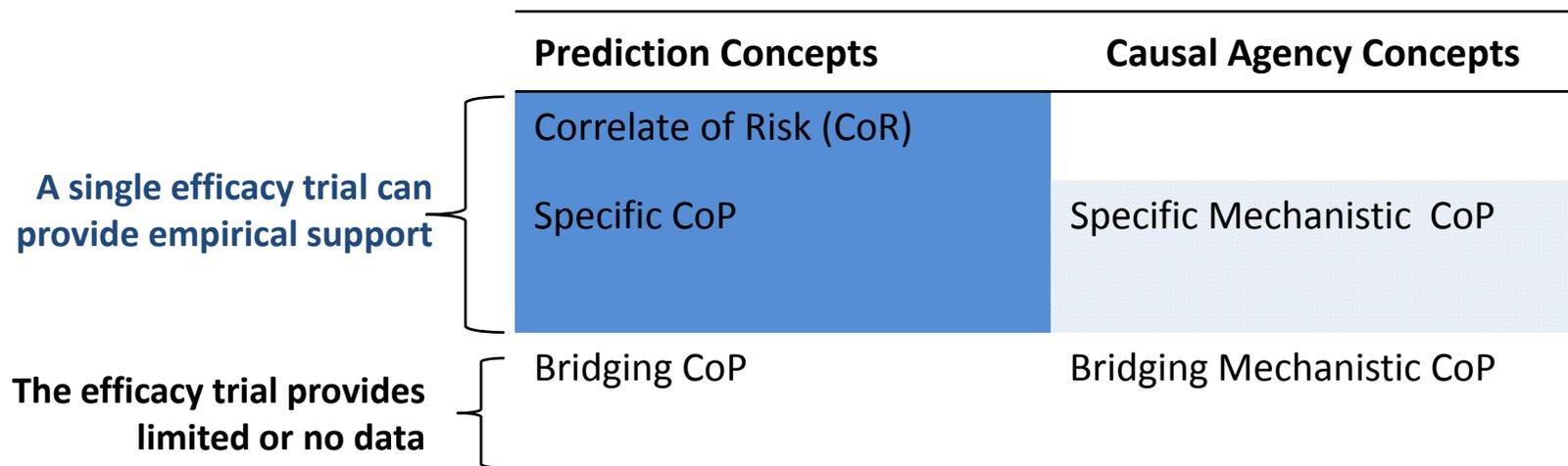# Examples of Mechanistic and Non-Mechanistic CoPs

- Meningococcal vaccine (Borrow et al., 2005, *Vaccine*)
    - mCoP = bactericidal antibodies
    - nCoP = binding antibodies (ELISA)

- Zoster vaccine (Weinberg et al., 2009, *J Infec Dis*)
    - mCoP = cellular response (IFN-γ ELISpot)
    - nCoP = binding antibodies to varicella-zoster virus (gpELISA)

- Rotavirus vaccines (Franco et al., 2006, *Vaccine*)
    - mCoP = none known
    - nCoP = total serum IgA antibody titers

Prediction Paradigm: Nested Hierarchy of Immune Correlates Definitions (Qin et al., 2007, *J Infect Dis*)

	Definition	Framework for Empirical Assessment
Correlate of Risk (CoR): Tier 1	The biomarker correlates with the clinical endpoint measuring vaccine efficacy	Vaccine efficacy trials/ epidemiological studies
Specific Correlate of Protection (CoP): Tier 2	Vaccine effects on the biomarker predict vaccine efficacy, for the **same setting** as the efficacy trial	Single large efficacy trial or multiple similar trials
General Correlate of Protection (CoP): Tier 3	A specific CoP that reliably predicts vaccine efficacy in **different settings** (e.g., across vaccine lots, vaccine formulations, human populations, viral populations)	Multiple diverse efficacy and/or post-licensure trials

- Hierarchy in scientific importance and degree of data requirements for statistical assessment
- General correlates (i.e., "bridging correlates") are for a particular new setting
  - E.g., new vaccine formulation, human population, viral population
  - Reliable prediction to one new setting may fail for a different new setting

Importance of Causal Agency for Credibility of Bridging Predictions of Vaccine Efficacy

Prediction Concepts	Causal Agency Concepts
Correlate of Risk (CoR)	
Specific CoP	Specific Mechanistic CoP
Bridging CoP	Bridging Mechanistic CoP

A single efficacy trial can provide empirical support

The efficacy trial provides limited or no data

- A single efficacy trial can provide direct data for assessing CoRs and specific CoPs, and perhaps supportive data for assessing causal agency, but typically provides scant direct information for assessing bridging correlates
- But, reliable bridging predictions is a central need for guiding research and deployment
- **Knowledge of the causal mechanism(s) of protection is core for building the rationale basis for bridging**

Assessing a Mechanistic CoP

- Many approaches outside of vaccine efficacy trials are needed

- Basic science
  - Understand specificity/functionality of biomarkers
  - Understand all the effects of vaccination (intended and unintended)
  - Understand the disease process

- Laboratory validation studies
  - Understand measurement/variability characteristics of immune biomarkers

- Causal manipulation studies in animal trials
  - Challenge efficacy trials comparing animals with and without induction of the immune biomarker(s)
  - E.g., passive antibody transfer repeated low-dose challenge studies in macaques

Nesting of CoR and CoP Concepts

Misleading for vaccine development: Does not predict vaccine efficacy

CoR that is not a CoP

CoP
(Non-mechanistic)

CoP
(Mechanistic)
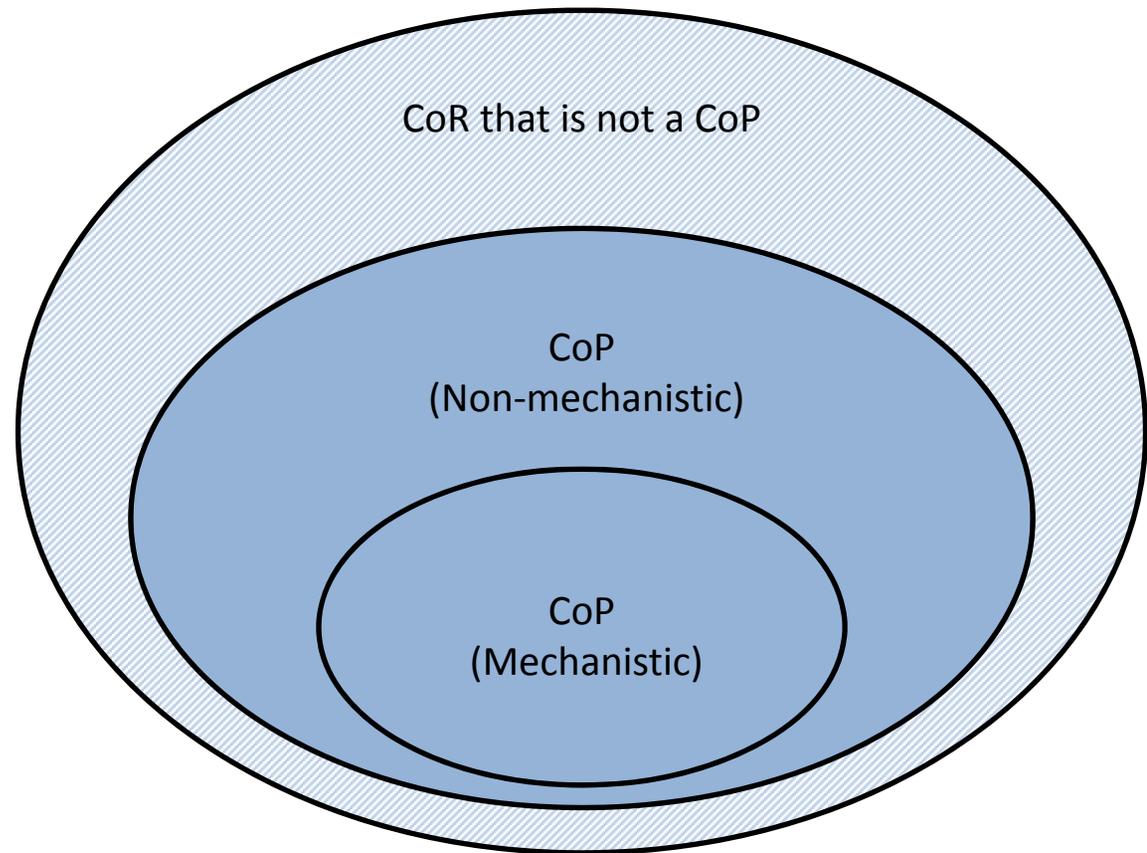
Useful for vaccine development: Predicts vaccine efficacy

Outline Session 3

- Introduction to immune correlates

- Prediction paradigm vs. mechanism of protection paradigm

- **Frameworks for statistical assessment of immune response biomarkers as correlates of protection (CoPs)/surrogate endpoints**

  1. **Surrogate endpoint for the clinical endpoint providing reliable inferences about VE [valid replacement endpoint]**

  2. **Policy/predictors of VE [controlled effects] and mediators of VE [natural direct and indirect effects]**

  3. **Effect modifiers of VE [one or a few efficacy trials]**

  4. **Effect modifiers of VE [meta-analysis of a series of efficacy trials]**

- Summary and conclusions

Four Frameworks for Surrogate Endpoints
(Joffe and Greene, 2008, *Biometrics*)

- **Causal-effects paradigm**

    "for a good surrogate, the effect of treatment on the surrogate, combined with the effect of the surrogate on the clinical outcome, allow prediction of the effect of treatment on the clinical outcome"

    1. Prentice definition of a valid surrogate   *Valid replacement endpoint*

    2a. Controlled direct and indirect effects   *Policy/Prediction*

    2b. Natural direct and indirect effects   *Mediation*

- **Causal-association paradigm (i.e., effect modification)**

    "for a good surrogate, the effect of treatment on the surrogate is associated with its effect on the clinical outcome"

    3. Effect modifiers/Principal stratification   *Association of individual-level treatment effects*

    4. Effect modifiers/Meta-analysis   *Association of group-level treatment effects*

Statistical Center for
HIV/AIDS Research & Prevention
SCHARP

FRED **HUTCHINSON**
CANCER RESEARCH **CENTER**
A LIFE OF SCIENCE

18

Problem Set-Up: Randomized Placebo-Controlled Vaccine Efficacy Trial

- Fixed follow-up period of duration τ_1

Z = randomized treatment (1=vaccine; 0=placebo)

Y = clinical endpoint (1=event by τ_1; 0=event-free at τ_1)

S = candidate surrogate measured at a fixed time point
$$\tau < \tau_1$$

- VE = 1 – E[Y|Z=1]/E[Y|Z=0])

    = multiplicative reduction in risk by vaccine

Prentice (1989, Statistics in Medicine) Definition of a Valid Surrogate Endpoint

- $VE = 1 - E[Y|Z=1]/E[Y|Z=0])$

- Definition: S is a valid surrogate endpoint for Y if a valid test of H^Y_0: VE=0 is obtained by testing H^S_0: $E[S|Z=1] = E[S|Z=0]$ (i.e., H^Y_0 iff H^S_0)

[alternatively, H^S_0 could be based on whole distributions
$$H^S_0: P(S \leq s|Z=1) = P(S \leq s|Z=0) \text{ for all } s]$$

  – A highly useful definition, fitting to the ICH document E8 intent

Useful Property of the Prentice Definition: Guarantees Avoidance of the Surrogate Paradox

- The Prentice definition is about hypothesis testing and the correct sign of VE, not about estimation of VE

- Definition closely connected to the *consistent surrogate* property (= avoidance of the surrogate paradox)

> ## Surrogate Paradox*
> – Vaccine effect on S is positive ($E[S|Z=1] > E[S|Z=0]$)
> – S and Y are inversely correlated in each group $Z=z$, $z=0,1$
> – Yet VE < 0 [a harmful vaccine!]

*E.g., Fleming and DeMets (1996, *Ann Int Med*), Chen et al. (2007, *JRSS-B*), Ju and Geng (2010, *JRSS-B*), VanderWeele (2013, *Biometrics*)

Prentice (1989) Criteria for Checking the Definition of a Valid Surrogate

- **Main Prentice criteria**:

  1. Vaccination impacts the immunological marker

     - $E[S|Z=1] > E[S|Z=0]$

  2. The immunological marker is associated with the clinical outcome in each of the vaccine and placebo groups

     - $E[Y|S > s, Z=z] < E[Y|S \leq s, Z=z]$ for some s and both z=0,1

  3. The relationship between the immunological marker and the clinical outcome is the same in the vaccine and placebo groups

     - $E[Y|S = s, Z = 1] = E[Y|S = s, Z = 0]$ for all s

Example: 1943 Influenza Vaccine Field Trial (Salk, Menke, and Francis, 1945)

- Endpoint Y = Hospitalization with Weiss Strain A influenza during the winter flu season

- Incidence of Y

  - **Placebo group:**   75 of 888   (8.45%)
  - **Vaccine group:**   20 of 888   (2.25%)

  Estimated **VE** = $(1 - 2.25/8.45) \times 100\% = 73\%$
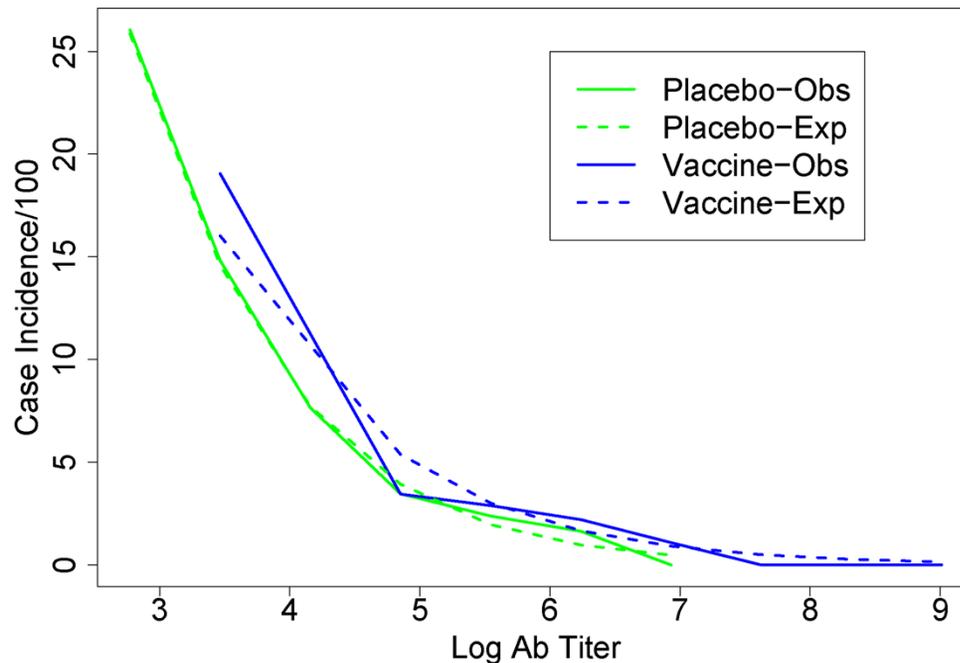
- **Goal:** Check the Prentice criteria for assessing Weiss Strain A antibody titer measured one week after vaccination as a valid replacement surrogate for Y

Criterion 1 Holds

Placebo Group
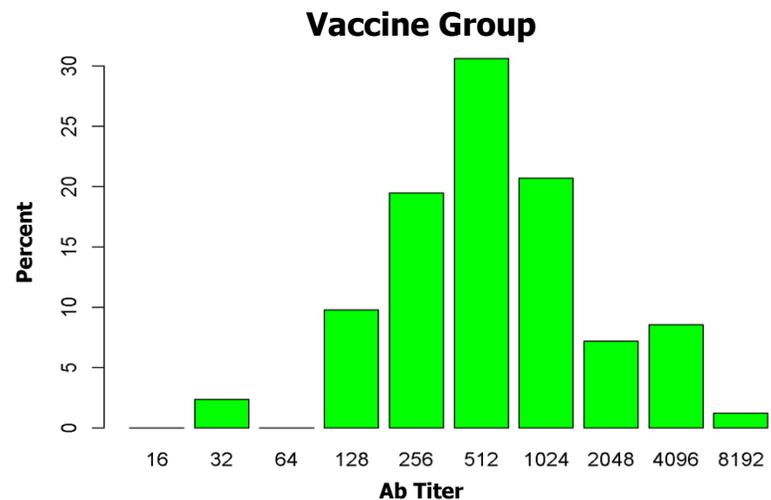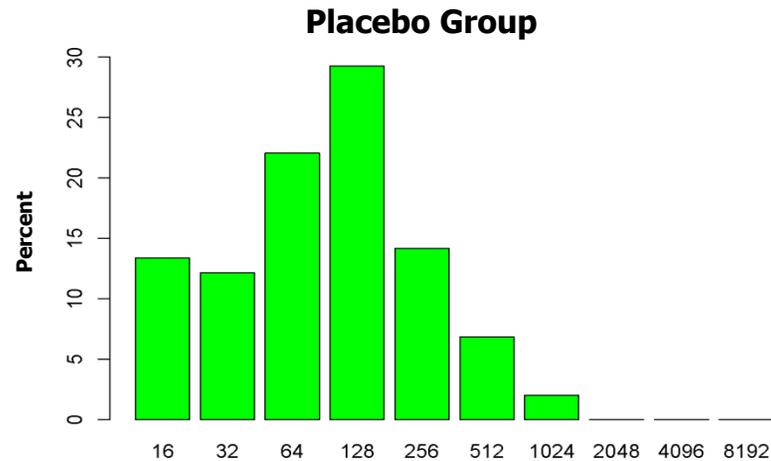
Vaccine Group

- Criterion 1: Weiss strain A Ab titers are higher in the vaccine than placebo group

- Criterion 1 holds ✓

Criteria 2 and 3 Hold*

*Based on logistic regression and on empirical fits

- Criterion 2: Check for an association of the marker with outcome in each treatment group √

- Criterion 3: Check for consistency between the association models in the vaccine and placebo √ groups

Prentice Criteria: 3 Challenges

1. For validity must include all dual predictors of the biomarker and the clinical outcome*

2. For validity must include all dual predictors of clinical risk before and after the biomarker is measured**

3. Cannot evaluate the Prentice operational criteria for vaccine trials for which there is ~no variability of the biomarker in the placebo group***

*E.g., Frangakis and Rubin (2002, *Biometrics*), Gilbert et al. (2009, *Stat Med*)

**E.g., Wolfson and Gilbert (2010, *Biometrics*)

***E.g., Chan et al. (2002, *Stat Med*)

# Challenge 1 with the Prentice Criteria: Post-Randomization Selection Bias

- Checking criterion 3 entails checking, for each immune response level s, equal clinical risk between the groups

  {Vaccinees w/ marker level s} vs. {Placebos w/ marker level s}

- However, S is measured after randomization
  - Therefore this comparison may reflect **selection bias, potentially misleading about the markers' value as a surrogate endpoint***

  *This phenomenon has been discussed broadly including in Rosenbaum (1984, *JRRS-B*), Robins (1986, *Math Modeling*), Pearl (2000, 2009, Causality textbook), Frangakis and Rubin (2002, *Biometrics*), Joffe and Greene (2008, *Biometrics*)

Illustration of Post-Randomization Selection Bias

- Binary immunologic measurement (positive or negative)

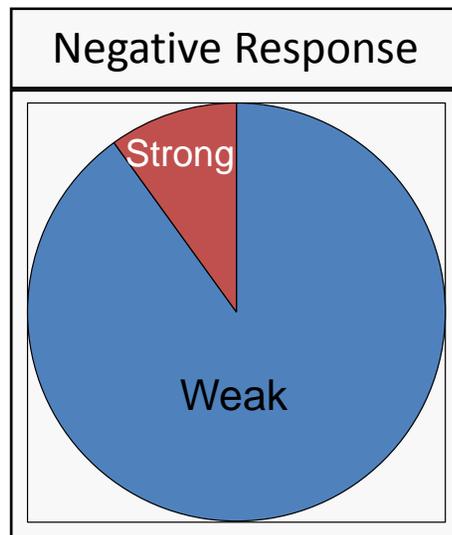- Consider an unmeasured covariate reflecting strength of immune system (strong or weak)
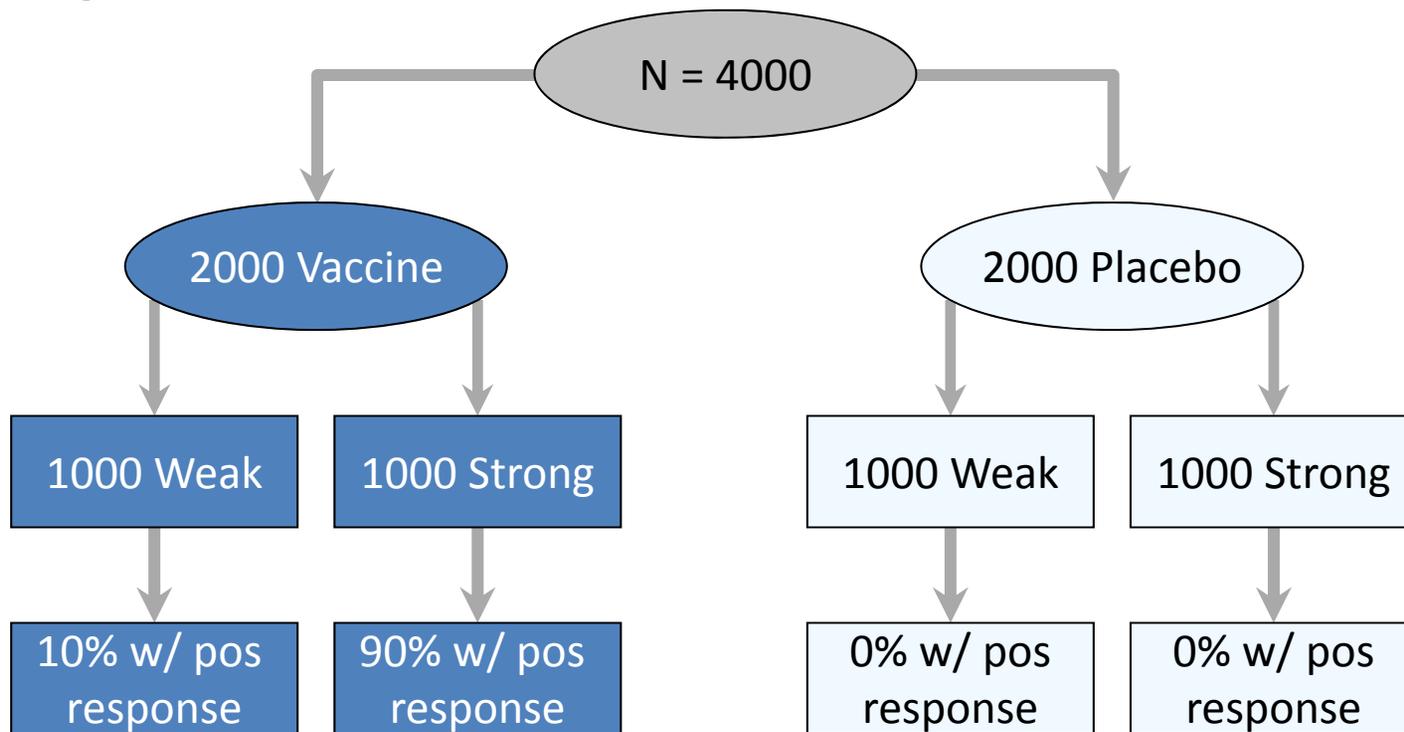
Prentice Criterion 3:
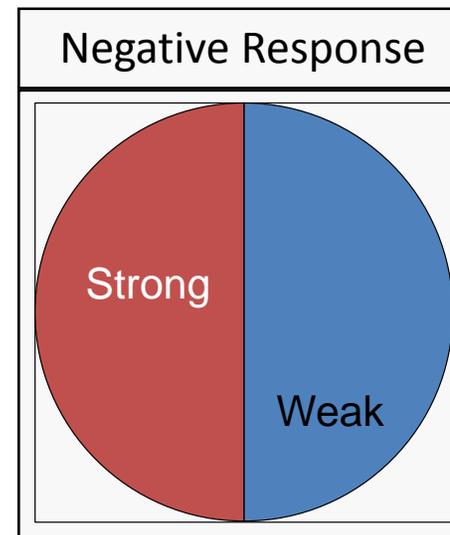# Compares Apples and Oranges

{Vaccinees w/ neg response} vs {Placebos w/ neg response}

compares clinical endpoint rates between the groups

Vaccine

Placebo

Compares a group with 90% weak immune systems to one with
50% weak immune systems: **Incomparable**

Challenge 2 with the Prentice Criteria: Unmeasured Dual Predictors of Early and Later Clinical Risk

- **Y^τ = Indicator of infection before the biomarker is measured**

- **For validity the statistical surrogate approach assumes no unmeasured simultaneous predictors V of Y^τ and Y**

  - That is, for validity there cannot be any unaccounted for subject characteristics that predict both early and later clinical risk

  - Practically speaking, this means that one must control for all clinical prognostic factors

Practical Implications of Unmeasured Dual Predictors of S and Y

- Update the third Prentice criterion to

  $$E[Y|S = s, Z = 1, W] = E[Y|S = s, Z = 0, W] \text{ for all } s$$

  with W measured baseline covariates, and W is assumed to include all dual predictors and all prognostic factors for Y

- In checking the criteria through regression models, always include all baseline covariates W that may predict both S and Y or predict both Y^τ and Y

- Acknowledge there may be residual confounding

  - Consider a sensitivity analysis to assess the sensitivity of inferences to unmeasured confounding

  - Consider making inference using estimated uncertainty intervals* that account for uncertainty due to partial non-identifiability as well as due to sampling variability

*E.g., Vansteelandt et al. (2006, *Statistica Sinica*); Richardson et al. (2015, *Stat Science*)

Challenge 3 with the Prentice Criteria: Cannot be Checked if No Variability of the Marker in the Control Arm*

- The immunologic measurement is a response to a pathogen-specific protein or proteins
  - If trial participants have never been infected with the pathogen, the immune response will be "non-response"/zero for (almost) all placebo recipients
  - E.g., HIV vaccines
- The Prentice framework does not apply, because criterion 3 cannot be checked and it cannot hold

*E.g., this challenge discussed by Chan et al. (2002, *Stat Med*), Gilbert and Hudgens (2008, *Biometrics*)

Surrogate Failure Can Happen in Many Ways

1. The biomarker is not in the pathway of the intervention's effect, or is insensitive to its effect (e.g., stemming from measurement error)

2. The biomarker is not in the causal pathway of the disease process

3. The intervention has mechanisms of action independent of the surrogate-mediated disease process

Example 1: CAST Trial*

- Encainide and flecainide in patients after a heart attack had a promising effect on the putative surrogate Arrhythmia suppression

- However, in a Phase 3 trial these drugs both increased the rate of mortality compared to placebo

- Classic example of the surrogate paradox

*Echt et al. (1991, *New Engl J Med)*

Example 2: Acellular Pertussis Vaccines with Mechanisms of Action Independent of the Surrogate-Mediated Disease Process*

## (Sweden I Trial with DT control: 10,000 subjects)

- Other relevant immune responses not captured by the assay (incomplete measurement of Ab responses)

Vaccine	VE	95%
SKB	58%	(51%, 66%)
Aventis Pasteur	85%	(81%, 89%)

- Biomarkers: *Filamentous Haemagglutinin (FHA)* and *Pertussis Toxoid (PT)* antibody responses were superior with the SKB vaccine

*Example from Tom Fleming

Example 2: Acellular Pertussis Vaccines with Mechanisms of Action Independent of the Surrogate-Mediated Disease Process
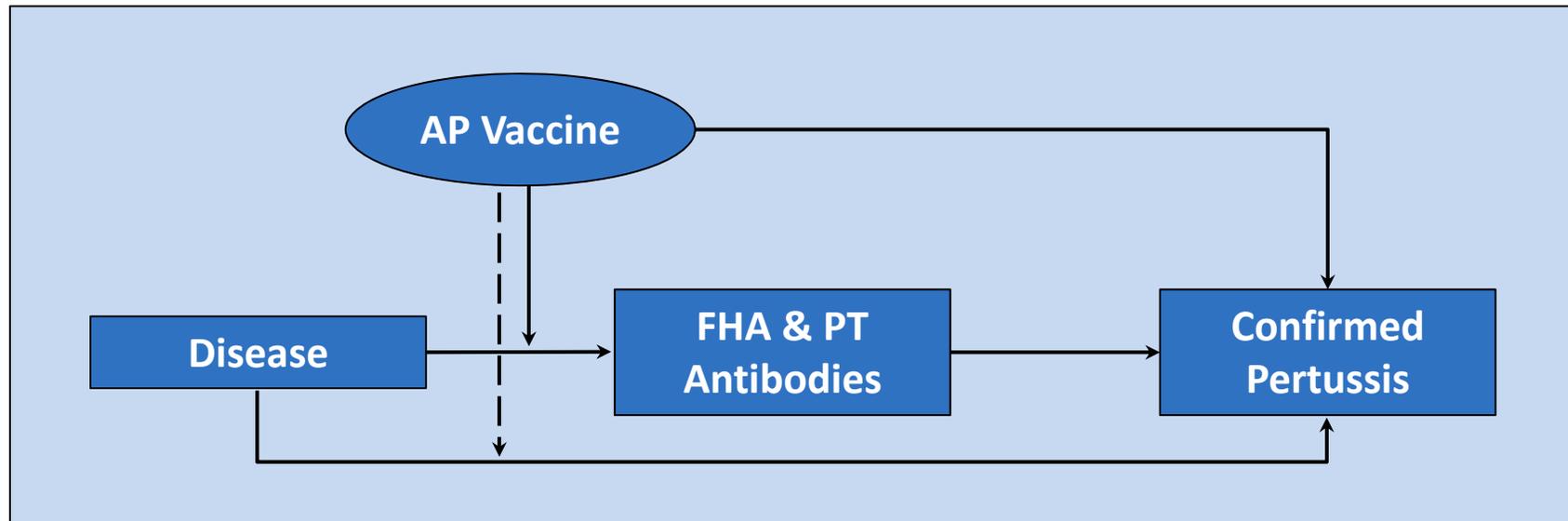
- Other immune responses, including those resulting from additional antigens in the vaccines:
  - Pertactin
  - Fimbriae (types 2 and 3)

3 Causes of the Surrogate Paradox*

1. **[Confounding]** Confounding of the association between the potential surrogate and the clinical endpoint (failure reason 1)**

2. **[Non-Transitivity]** The vaccine positively affects both the surrogate and the clinical endpoint, but for different sets of participants (failure reason 2)

3. **[Off-Target Effects]** The vaccine may have a negative clinical effect in ways not involving the potential surrogate (failure reason 3)

*VanderWeele (2013, *Biometrics*)

**Add to this confounding of the association between occurrence of the clinical endpoint before and after the surrogate is measured

FRED **HUTCHINSON**
CANCER RESEARCH **CENTER**
A LIFE OF SCIENCE

"There is a plague on Man, the opinion that he knows something."

Michel de Montaigne (1580, *Essays*)

"Medicine is something a doctor prescribes while waiting for nature to cure the disease."

Voltaire (mid-18th Century)

Inference to Assess the Prentice Criteria in the General Setting of Variability of S in Placebo Recipients

- Specify a regression model
  - E.g., $\text{logit}(E[Y|Z=z,S=s]) = b_0 + b_1 z + b_2 s + b_3\, z \times s$

- First examine the interaction of $Z \times S$ (b_3)
  - Interaction ($b_3 \neq 0$) implies Prentice criteria violated

- If no interaction (($b_3 = 0$), examine the main effects of Z and S
  - Z main effect ($b_1 \neq 0$) implies Prentice criteria violated
  - No S main effect ($b_2 = 0$) implies Prentice criteria violated
  - $b_1 = 0$ and $b_2 \neq 0$ implies Prentice criteria

- The regression models must marginalize over potential baseline confounders W and must account for the marker sampling design (e.g., Breslow and Holubkov, 1997, *JRSS-B* for nested case-control studies)

- Example summary measure: Kobayashi and Kuroki (2014, *Stat Med*)

2. Controlled and Natural Causal Effects Framework

- Literature with key contributors Judea Pearl, James Robins, Sanders Greenland, Tyler VanderWeele

    – Robins and Greenland (1992, *Epidemiology*)

    – Pearl (2001, *Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence*)

    – VanderWeele (2015, *Explanation in Causal Inference: Methods for Mediation and Effect Modification*)

- Need potential outcomes notation (Neyman, 1923; Rubin, 1974)

Potential Outcomes/Counterfactuals

- **Potential Outcomes/Counterfactuals Framework**

  – Use of counterfactuals by the Father of Probability: Blaise Pascal, *The Pensees (1660),* No. 413, Lafuma Edition

  "Cleopatra's nose: if it had been shorter the whole face of the earth would have been different."

- Many historians acknowledge that Marc Antony's falling in love with Cleopatra played a major role in the fall of the Roman Republic

Blaise Pascal [1623-1662]

- For many, counterfactuals are a natural way of thinking

Potential Outcomes/Counterfactuals

- **Notation**
  - $S_i(Z)$ = potential immune response endpoint under assignment Z; for $Z = 0, 1$
  - $Y_i(Z)$ = potential clinical endpoint under assignment Z; for $Z = 0, 1$

- **Individual Causal Effects**
  - $S_i(1) - S_i(0)$ = causal effect on S for subject i
  - $Y_i(1) - Y_i(0)$ = causal effect on Y for subject i

- **Average causal effects:**
  - $E[S_i(1) - S_i(0)] = E[S|Z=1] - E[S|Z=0]$
  - $E[Y_i(1) - Y_i(0)] = E[Y|Z=1] - E[Y|Z=0] = P[Y=1|Z=1] - P[Y=1|Z=0]$

By randomization

For a binary outcome Y

2. Controlled and Natural Direct and Indirect Causal Effects

- **Two types of causal effects: controlled and natural**

- **Concept of Controlled**
  - ➢ Envisage an experiment where individuals are simultaneously assigned to treatment (vaccine vs. placebo) and to the immune response biomarker S

- **Concept of Natural**
  - ➢ Envisage an experiment where individuals are assigned to treatment (vaccine vs. placebo), and, for one of the treatments, are assigned to the immune response level they would have had under the unassigned treatment

Controlled Causal Effects

- **Notation**
  - $Y_i(Z, S)$ = potential clinical endpoint under <u>assignment to both</u> Z and S

- **Individual Causal Effects**
  - Total effect $\qquad\qquad\qquad Y_i(1) - Y_i(0)$
  - Direct effect $\qquad$ **(at s):** $Y_i(1, s) - Y_i(0, s)$ $\qquad$ **[hold S fixed at s]**
  - Indirect effect $\qquad$ **(at s):** Total effect – direct effect
    $$[Y_i(1) - Y_i(0)] - [Y_i(1, s) - Y_i(0, s)]$$

- **Average Causal Effects**
  - Total effect $\qquad\qquad\qquad$ TE $= E[Y_i(1) - Y_i(0)]$
  - Controlled direct effect $\;$ **(at s):** CDE(s) $= E[Y_i(1, s) - Y_i(0, s)]$
  - Controlled indirect effect **(at s):** CIE(s) $= E[Y_i(1) - Y_i(0)] - E[Y_i(1, s) - Y_i(0, s)]$

Controlled Causal Effects

- Average Causal Effects

$$TE = E[Y(1)] - E[Y(0)]$$

$$CDE(s) = E[Y(1, s)] - E[Y(0, s)]$$

- $CDE(s) = 0$ for a fixed level s means that the vaccine effect on Y is eliminated by intervening to set the immune response to s for all individuals

- $CDE(0) = 0$ is of special interest (set everyone to have no immune response)

- Proportion eliminated $PE(s) = \dfrac{TE - CDE(s)}{TE}$

- $PE(s)$ captures what would happen to the vaccine effect on Y if we were to fix the immune response to the same fixed value $S = s$ for all persons

- **Application: Predict VE for hypothetical vaccines that would generate certain fixed immune response levels**

Natural Causal Effects (Mediation)

- Definitions of the Natural Direct Effect (NDE) and the Natural Indirect Effect (NIE)
  - NDE: $E[Y(1,S(0))] - E[Y(0,S(0))] = E[Y(1,S(0))] - E[Y(0)]$
  - NIE: $E[Y(0,S(1))] - E[Y(0,S(0))] = E[Y(0,S(1))] - E[Y(0)]$

- Interpretation:
  - NDE expresses how much the outcome Y would change under vaccination vs. under placebo if for each individual the immune response were kept at the level it would have taken, for that individual, under placebo
    - I.e., what is VE if the immune response pathway is 'deactivated'?
  - NIE captures the effect of vaccination on the outcome that operates by changing the immune response
    - The vaccine has an effect on the immune response which in turn then effects the clinical outcome Y

Natural Causal Effects (Mediation)

- TE decomposes: TE = NDE + NIE

- Proportion mediated $PM = \dfrac{TE - NDE}{TE}$

- PM measures how much of TE is through the immune response S (a higher value indicates a stronger mechanism of protection)
- PM = 1.0 (i.e., NDE = 0) means S is a complete mechanism of protection
- Comparing markers by PM ranks them by strength of involvement as mechanisms of protection

Assumptions for Identifying the Average Causal Effects*

- **Controlled Effects**

  A1: No unmeasured confounding of the treatment-outcome relationship (Z, Y)

  A2: No unmeasured confounding of the mediator-outcome relationship (S, Y)

- **Natural Effects: Need A1+A2 and**

  A3: No unmeasured confounding of the treatment-mediator relationship (Z, S)

  A4: There is no mediator-outcome confounder that is affected by treatment

  – A1, A3 hold in randomized trials
  – A2 may not hold. Therefore, include in the analysis all baseline covariates that may be dual predictors of S and Y, and consider sensitivity analysis and estimated uncertainty intervals.
  – A4 may not hold. It may be more plausible if the mediator is measured shortly after randomization.

*As defined in VanderWeele textbook (2015, pages 24, 25)

Are Controlled and Natural Causal Effect Parameters Conceivable*?

- This framework requires **conceivability of manipulations/ interventions**
  - **Controlled:** CDE(s) requires setting a placebo recipient's immune response biomarker S to level s of the vaccine-induced immune response
  - **Natural:** NDE requires setting a vaccine recipient's immune response biomarker S to any possible level in the range of a placebo response

- Two scenarios:
  - Case Constant Biomarker (Case CB): No variability of S in the placebo group (occurs in trials of individuals without prior exposure to the pathogen, e.g., HIV)
    - CDE(s) inconceivable for s > 0: the framework is not applicable
    - NDE may be conceivable; some may debate conceivability
  - General case where S varies in the placebo group (occurs in trials of subjects with prior exposure to the pathogen)
    - CDE(s) and NDE may be conceivable; some may debate this

*Discussed in Gilbert, Hudgens, and Wolfson (2011, *J Int Biost*)

Example Method of Inference for a Continuous Mediator $S \sim N(\mu, \sigma^2)$ and a Binary Outcome Y*

- Specify a regression model for the mediator
  - E.g., $E[S|Z=z, W=w] = a_0 + a_1 z + a^T_2 w$
- Specify a regression model for the binary outcome
  - E.g., $\text{logit}(E[Y|S=s, Z=z, W=w]) = b_0 + b_1 z + b_2 s + b_3 z \times s + b^T_4 w$ (*)
- Assuming the 4 assumptions (A2.1−A2.4) and Y is relatively rare (e.g., $E[Y] < 0.10$), causal effects on the odds ratio scale for any 2 fixed immune response levels a and a*:
  - $OR^{CDE}(s) = \exp\{(b_1 z + b_3 s)(a-a^*)\}$
  - $OR^{NDE} = \exp\{(b_1 + b_3 a_0 + b_3 a_1 a^* + b_3 a^T_2 w + b_3 b_2 \sigma^2)(a - a^*) + 0.5 b^2_3 \sigma^2 (a^2 - a^{*2})\}$
  - $OR^{NIE} = \exp\{(b_2 a_1 + b_3 a_1 a)(a-a^*)\}$
  - $TE = OR^{NDE} \times OR^{NIE}$
- The outcome regression model (*) must account for the marker sampling design (e.g., with inverse probability weighting; Session 5)
- The outcome regression model (*) is commonly used for checking the Prentice criteria, partly explaining why the 2 frameworks are sometimes conflated

*Baron and Kenny (1986, *J Personality and Social Psychology*), VanderWeele and Vansteelandt (2010, *AJE,* pages 21, 27)

# 3. Effect Modification "VE Curve" Framework [Heuristic, Math in Sessions 6−7]

- Introduced by Follmann (2006, *Biometrics*)

- Define

$$VE(s) = 1 - \frac{\text{Risk of infection for Vaccinees with immune response s to Vaccine}}{\text{Risk of infection for Placebos with immune response s to Vaccine}}$$

- Interpretation: Percent reduction in clinical risk for groups of vaccinees with Ab titer compared to if they had not been vaccinated

- Definition (Gilbert and Hudgens, 2008, *Biometrics*): A Principal surrogate is an immune response biomarker S satisfying

  1. **VE**(negative response s = 0) = 0   **[Average Causal Necessity]**
  2. Large variability of **VE**(s) in s   **[Strong Effect Modifier]**

- Gilbert, Gabriel, Huang, Chan (2015, *J Causal Inference*) suggested criterion 2. alone makes a marker useful, 1. not being crucial

The VE Curve Framework Provides a Way to Compare the Ability of Different Markers to Predict VE
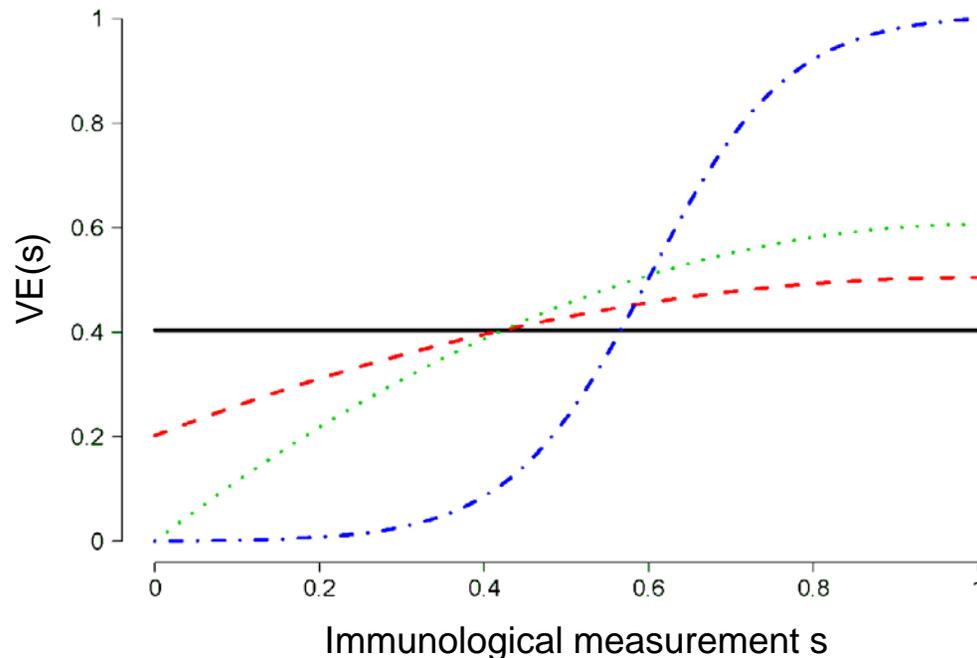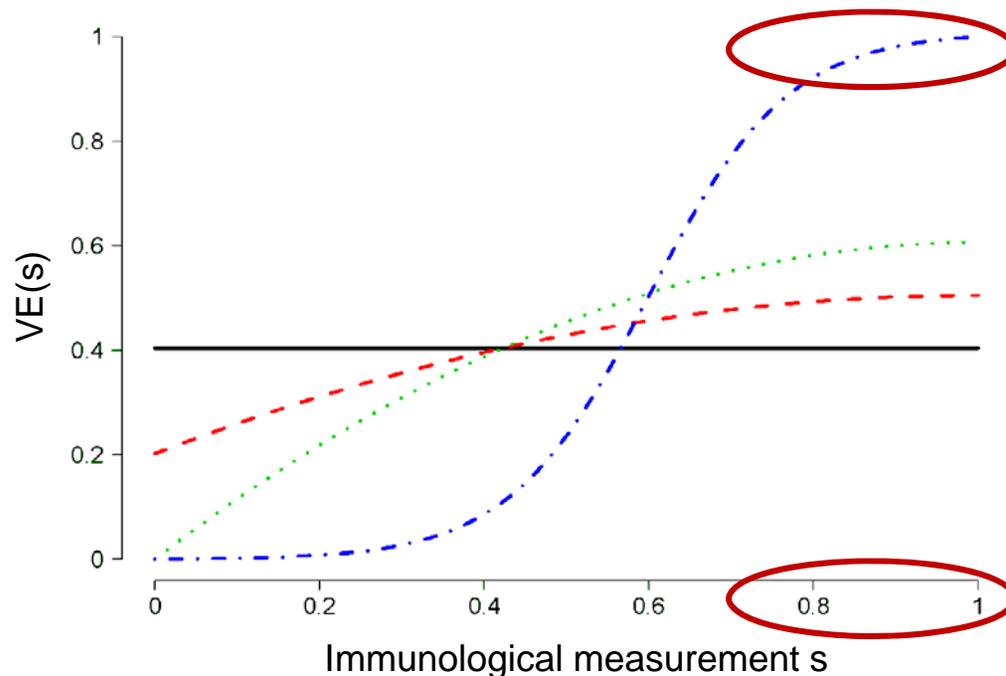

Immunological measurement s

- Black marker: Useless marker: no effect modification

- Green and blue markers: Useful marker, moderate effect modification, satisfy ACN

- Blue marker: Very useful marker, strong effect modification, satisfies ACN

ACN = Average Causal Necessity [VE(negative response s = 0) = 0]

Strong Effect Modifier:
Sets the Target for Improving the Vaccine

Target: Improve the vaccine regimen by increasing the percentage of vaccinees with high immune responses

- Black marker: Useless marker: no effect modification
- Green and blue markers: Useful marker, moderate effect modification, satisfy ACN
- Blue marker: Very useful marker, strong effect modification, satisfies ACN

ACN = Average Causal Necessity [VE(negative response s = 0) = 0]

VE Curve Framework Is About Effect Modification

- Evaluate if and how VE varies across subgroups of vaccine recipients defined by S measured at time point τ

- Conceptually identical to studying effect modification in baseline covariate subgroups (the essence of principal stratification)

- Some papers claim this 'Principal surrogate' framework is about mediation, but this is only 'half true'*

  - Average causal necessity (ACN) can hold yet NDE ≠ 0

 - But a rejection of ACN does imply NDE ≠ 0 (i.e., an incomplete mediator)

    *VanderWeele (2001, *Int J Biostatistics*)

# VE Curve Analysis May Guide Future Research to Develop Improved Vaccines

- Immune response biomarkers with stronger VE modification may be prioritized as study endpoints in follow-up Phase I/II trials of refined vaccines

- Generates a **bridging hypothesis:** If a future vaccine is identified that generates higher marker levels in more vaccinees, then it will have greater overall VE in a future efficacy trial

Using a Strong VE Modifier for Improving the Vaccine Regimen

Marker levels S
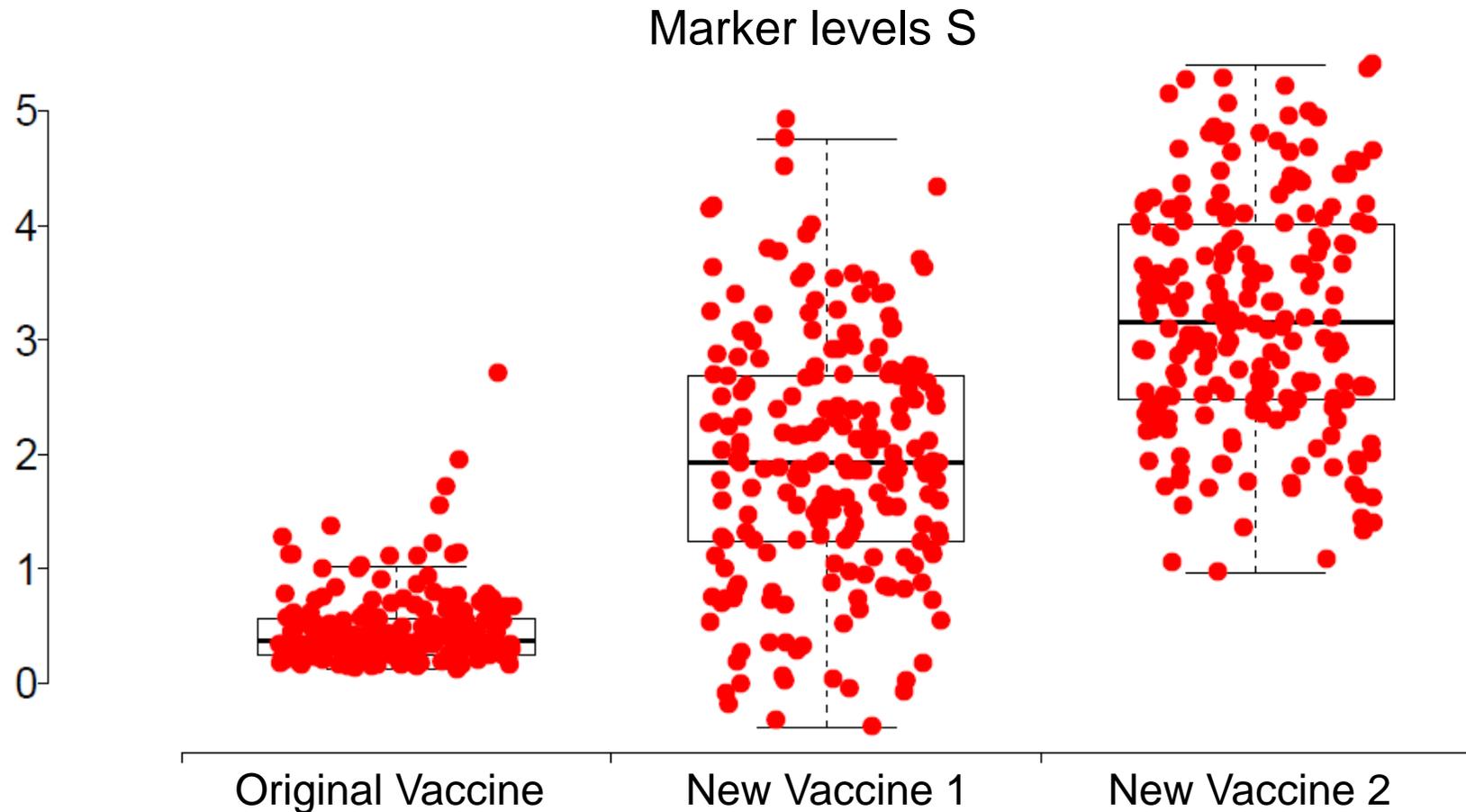
Using a Strong VE Modifier for Improving the Vaccine Regimen

- Suppose each new vaccine is tested in an efficacy trial
- Under the bridging hypothesis we expect the following efficacy results:
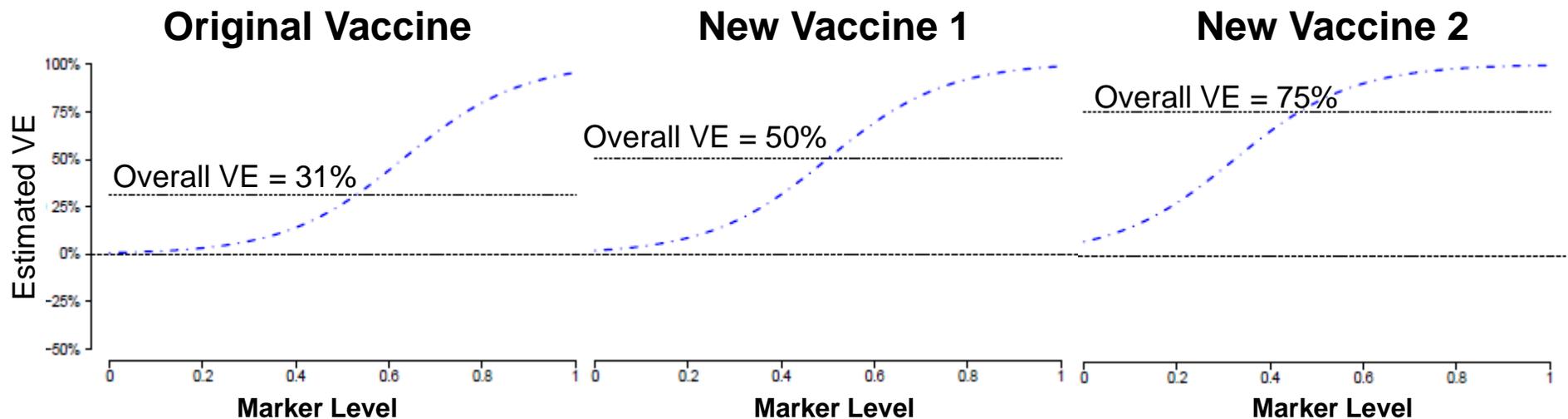
- This is the idealized model for using a VE modifier to iteratively improve a vaccine regimen
  - Gilbert and Huang (2016, *Epidemiologic Methods*) developed transport formulas for predicting VE in a new setting

Inference for the VE Curve Framework

- Challenge to inference: The immune responses to vaccine S(1) are missing for placebo recipients

  - Accurately filling in the unknown immune responses S is needed to precisely estimate the VE curve

  - Approaches to fill in the missing data:

    - At baseline, measure a predictor(s) of the immune response in both vaccine and placebo recipients (Follmann, 2006, *Biometrics*)

    - Close-out placebo vaccination (Follmann, 2006, *Biometrics*)

  - A literature has developed employing these techniques for inference on VE curves, all of these methods account for the marker sampling design [elaborated in Sessions 6–8]

4. Meta-Analysis Framework

- **Goal**: Develop models such that vaccine effects on the immune response biomarker S can be used to reliably predict VE in different settings (e.g., across vaccine lots, vaccine formulations, human populations, viral populations)

- **Approach to Evaluation:** Meta-Analysis
  - N pairs of immunologic and clinical endpoint assessments among vaccine and placebo recipients
  - Pairs chosen to reflect specific target of prediction
    - **Example:** Predict efficacy of new vaccine formulation: N vaccine efficacy trials of comparable vaccines but with different formulations
  - ➡ **Evaluation:** Study the relationship between the estimated **VEs** and the estimated vaccine effects on the immune response

Surrogate Endpoint Evaluation from Multiple Trials: Meta-Analysis (N = 25)*

Treatment Effect on AIDs vs Treatment Effect on VL
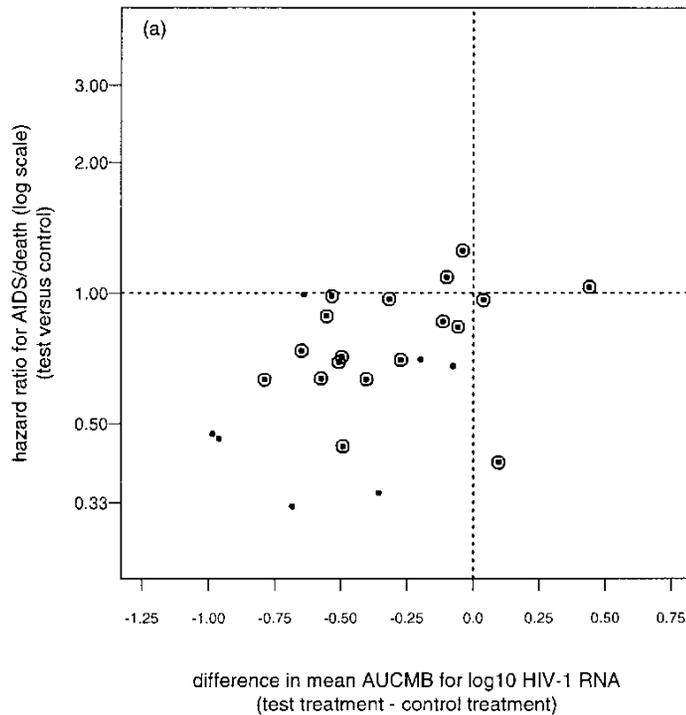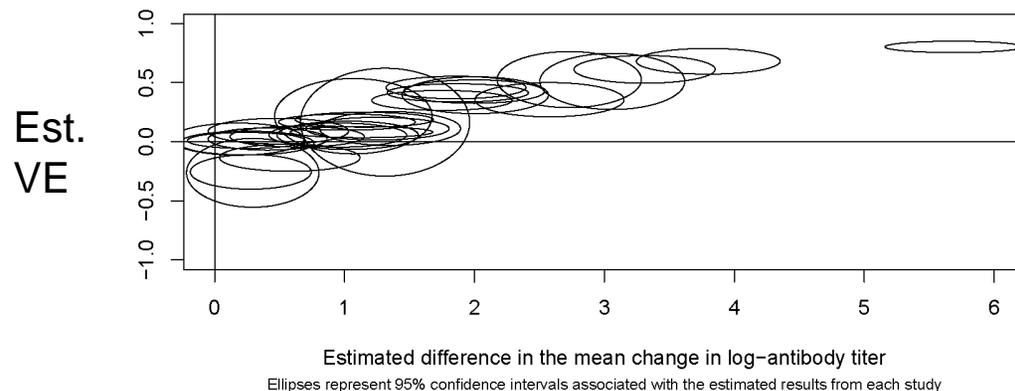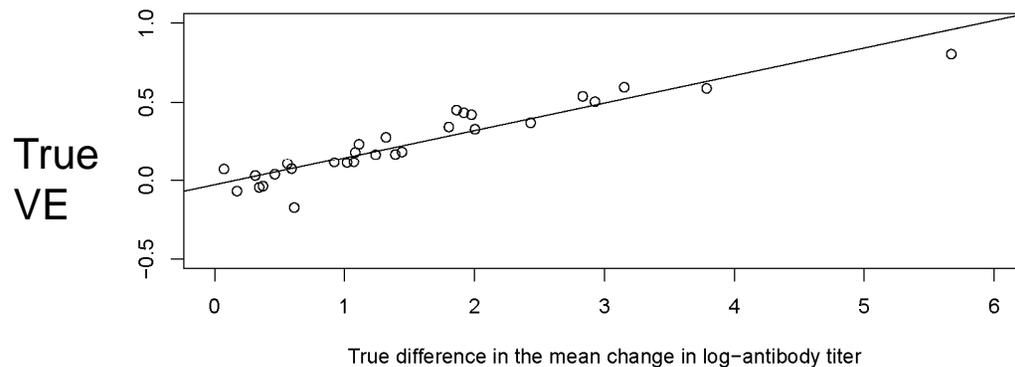
Treatment Effect on AIDs vs Treatment Effect on CD4
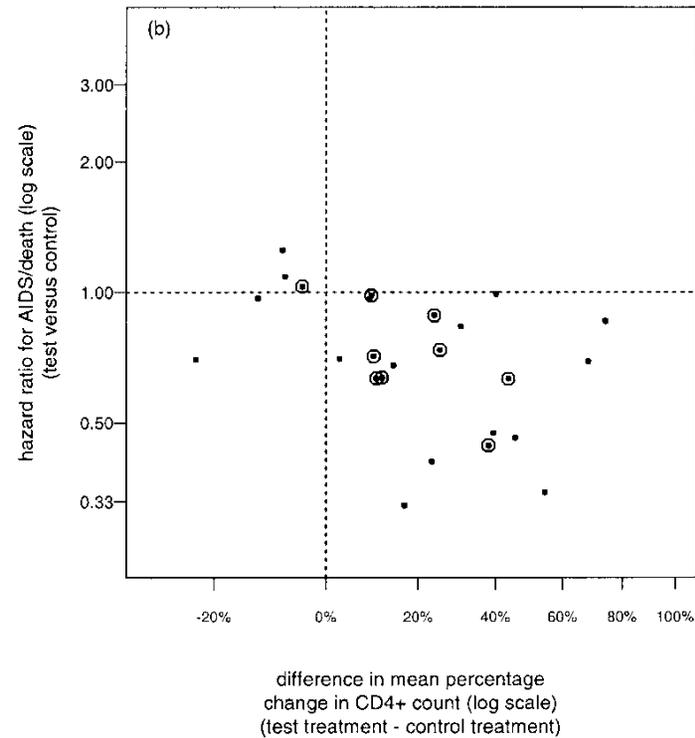
*HIV Surrogate Marker Collaborative Group (2000, *AIDS Res Hum Retroviruses*)

Simulated Meta-Analysis* Based on 29 Influenza Vaccine Trials (Villari et al., 2004, Vaccine)

Estimated VE vs Average difference in log Ab titer (Vaccine - Placebo)

Association between the clinical effect and the surrogate effect of vaccination

True VE

True difference in the mean change in log−antibody titer

Est. VE

Estimated difference in the mean change in log−antibody titer

Ellipses represent 95% confidence intervals associated with the estimated results from each study

Selected all placebo-controlled influenza vaccine trials of parainfluenza virus vaccines with ≥ 3 placebo cases

The N = 29 studies span different flu seasons over 30-40 years

Objective: Predict VE for next year's flu season

Clinical endpoint = clinically Confirmed influenza infection

Marker endpoint = log Ab titer to the dominant strain circulating in the trial region

*Qin et al. (2007, JID)

Simulated Meta-Analysis Based on 29 Influenza Vaccine Trials (Villari et al., 2004, Vaccine)

Assessing ability to predict protection across simulated studies
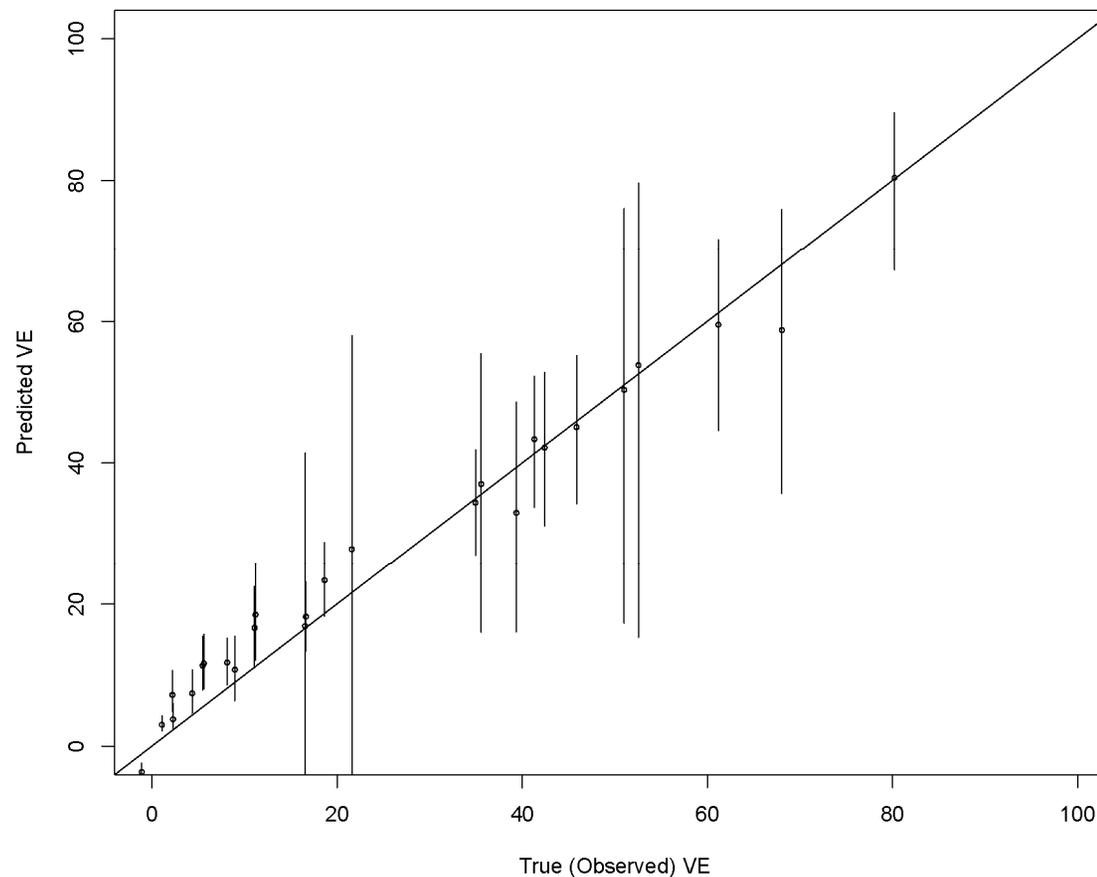
Predicting VE in a New Trial

- Building on Daniels and Hughes (1997, *Stat Med*), Gail et al. (2000, *Biostatistics*) developed methods for predicting VE with a bootstrap confidence interval in a new trial from
  - A series of N trials with estimated vaccine effects on the biomarker and on the clinical endpoint
  - A new trial with data on the biomarker only
- Summary of Gail et al. conclusions:
  - The strength of correlation of vaccine effects has a large effect on the precision for predicting VE
  - Need at least N=10 studies that have reasonably precise estimates of VE
    - Even with this, prediction of VE is quite imprecise
  - **Fundamental Challenge:** Do the N studies constitute an appropriate basis for extrapolating results to the setting of the new trial?

Summary of the Meta-Analysis Approach

- Meta-analysis directly assesses how well vaccine effects on the marker predict vaccine effects on the clinical endpoint

  - The VE curve framework for individual-level associations is conceptually similar, yet the meta-analysis framework has technical advantage that causal effects can be estimated with standard assumptions of randomized trials

  - However, meta-analysis requires large data resources (a series of diverse efficacy trials)

  - And predicting VE in a new setting requires extrapolation beyond the data support; hence knowledge of biological mechanisms may be needed for credibility of the bridging

Summary: Applications of the 4 Surrogate Endpoint Frameworks

Immune Correlate of Protection (CoP)/Surrogate Framework	Application
1. Prentice definition (Valid replacement endpoint)	Reliable inferences on VE based on the immune response surrogate alone (without measuring the true clinical endpoint)
2a. Controlled direct and indirect effects (Policy/Prediction)	Predict VE for a new vaccine that sets the immune response to certain levels
2b. Natural direct and indirect effects (Mediation)	Insights into mechanisms/pathways of protection
3. VE curve (Response type effect modification)	Inferences on how VE varies over subgroups of vaccine recipients defined by their immune response to vaccination
4. Meta-analysis of a series of vaccine efficacy trials	Associate causal effects on the immune response with causal effects on the clinical outcome, for predicting VE in new settings

Two Concepts: Individual- vs. Trial-Level Surrogacy

- **Trial level:** Both the surrogate and the VE are studied at the trial level
  - E.g., the difference in average immune response (vaccine vs. placebo) and overall VE
- **Individual subgroup level:** Both the surrogate and the clinical outcome or VE are studied at the individual level
  - E.g., the vaccine effect on my immune system and my personal probability of protection due to being vaccinated

Immune Correlate of Protection/Surrogate Framework	Application	Individual Level vs. Trial Level
1. Prentice definition (Valid replacement endpoint)	Reliable inferences on VE based on the immune response surrogate alone (without measuring the true clinical endpoint)	Definition is trial level; research has used it at both levels
2a. Controlled direct and indirect effects (Policy/Prediction)	Predict VE for a new vaccine that sets the immune response to certain levels	Individual level
2b. Natural direct and indirect effects (Mediation)	Insights into mechanisms/pathways of protection	Trial level
3. VE curve (Response type effect modification)	Inferences on how VE varies over subgroups of vaccine recipients defined by their immune response to vaccination	Individual level
4. Meta-analysis of a series of vaccine efficacy trials	Associate causal effects on the immune response with causal effects on the clinical outcome, for predicting VE in new settings	Trial level; Trial level + individual level [Literature led by Geert Molenberghs, Marc Buyse, et al.]

Outline Session 3

- Introduction to immune correlates
- Prediction paradigm vs. mechanism of protection paradigm
- Frameworks for statistical assessment of immune response biomarkers as correlates of protection (CoPs)/surrogate endpoints
    1. Surrogate endpoint for the clinical endpoint providing reliable inferences about VE [valid replacement endpoint]
    2. Policy/predictors of VE [controlled effects] and mediators of VE [natural direct and indirect effects]
    3. Effect modifiers of VE [one or a few efficacy trials]
    4. Effect modifiers of VE [meta-analysis of a series of efficacy trials]
- **Summary and conclusions**

Conclusions and Discussion (1)

- **What about 'modern data?'**

    1. Microarrays, proteomics, microbiomics, etc.

        - The concepts and principles are the same

    2. Promise of modern data:

        - May yield more comprehensive understanding of vaccine effects and of the infection and disease process, yielding mechanistic correlates that could not be uncovered with lower-dimensional techniques
        - May yield earlier predictive correlates (closer to randomization), greatly aiding surrogate assessment methods and improving their practical utility
        - E.g., 2-4 subgroups may be defined based on high-dimensional expression array or cell subpopulation frequency profiles, and the categorical subgroup variable may be assessed as a correlate of protection

Statistical Center for
HIV/AIDS Research & Prevention
SCHARP

FRED **HUTCHINSON**
CANCER RESEARCH **CENTER**
A LIFE OF SCIENCE

Conclusions and Discussion (2)

- **All of the surrogate endpoint frameworks are useful, with distinct applications**

- Settings where disease-experienced individuals are vaccinated
  - All frameworks apply and may bear fruit in a complementary way

- Settings where vaccine is given to disease-naïve persons
  1. Prentice criteria do not apply; the Prentice definition can be checked by other means
  2. Controlled and natural effects frameworks arguably do not apply
  3. Effect modification/VE curve framework appealing because identifiability easier
 4. Meta-analysis applies

Conclusions and Discussion (3)

- **Understanding surrogate endpoint validity is highly inter-disciplinary and requires synthesis of many experimental and data sources**

  - Basic science, pre-clinical research, clinical research work iteratively and in parallel to generate and test hypotheses

  - Knowledge of mechanism is particularly important for building credibility of valid surrogate endpoints, especially for bridging efficacy to new settings
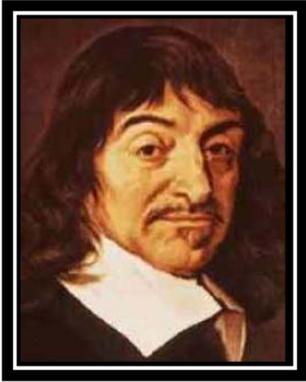
Descartes vs. Pascal: Complexity of the Surrogate Endpoint Problem
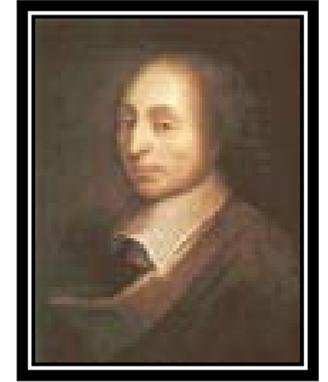
- Cartesian scientific method for discovering scientific truth:

  - Accepting as "truth" only clear, distinct ideas that could not be doubted

  - Breaking a problem down into parts

  - Deducing one conclusion from another

  - Conducting a systematic synthesis of all things

- Limited success for biology: overly-optimistic about what 'reductionist science' could deduce

- Pascal was skeptical about what Descartes' method could deliver:

  "But the parts of the world are all so related and linked together that I think it is impossible to know one without the other and without the whole"

  Blaise Pascal
  (1670, *The Pensees*, Lafuma Edition, No. 199)

- Correctly anticipated that complexity is too great to understand via reductionist methods