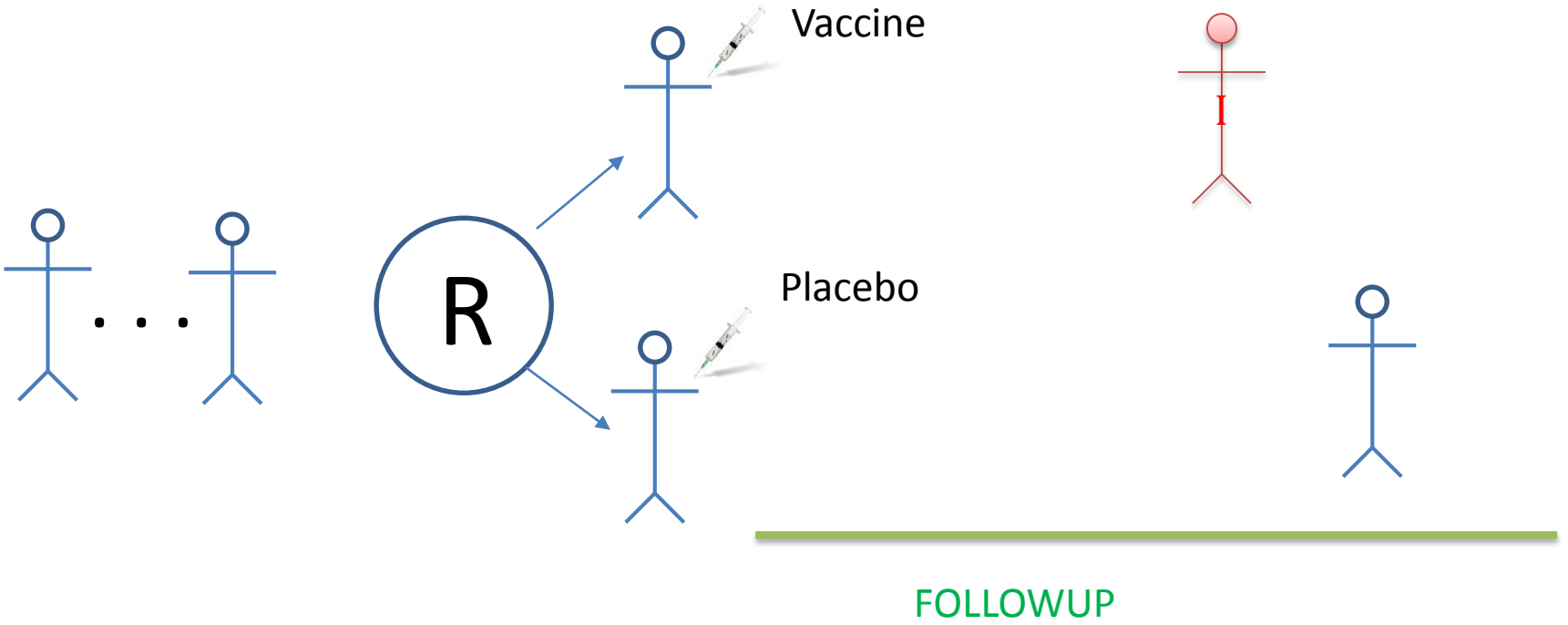
Incorporating Infecting Pathogen Counts In Vaccine Trials

Dean Follmann

National Institute of Allergy and Infectious Diseases

Vaccine Trial

- Randomize at risk healthy volunteers to vaccine or placebo
- Follow them & count significant infections

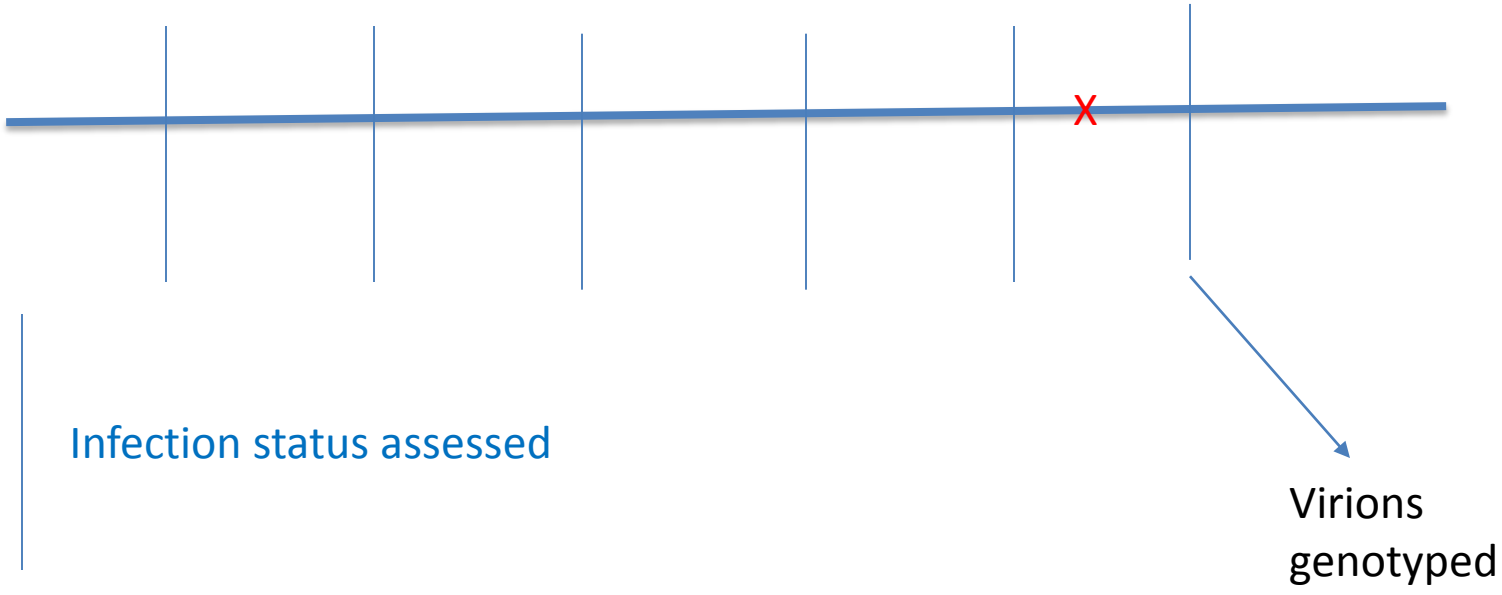


Vaccine Efficacy (VE)

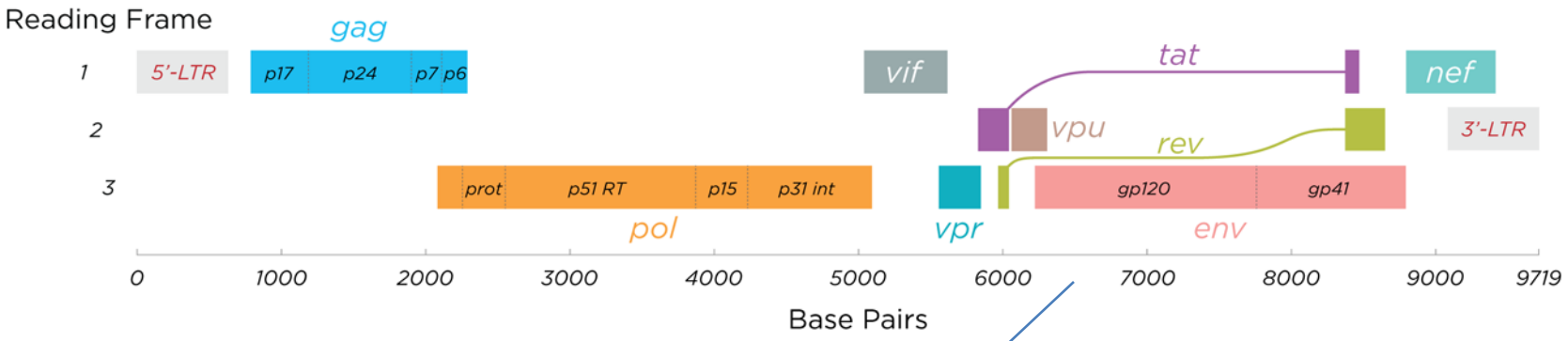
- What is the proportion reduction in some outcome on vaccine compared to placebo?
- $VE = 1 - \frac{Infection\ Rate\ on\ Vaccine}{Infection\ Rate\ on\ Placebo}$
- $VE = 1 - \frac{hazard\ rate\ on\ vaccine}{hazard\ rate\ on\ palcebo}$
- Based on human infection yes/no . . .

HIV Infection Detection

- Volunteers are followed at regular intervals (e.g. 6 months for infection)

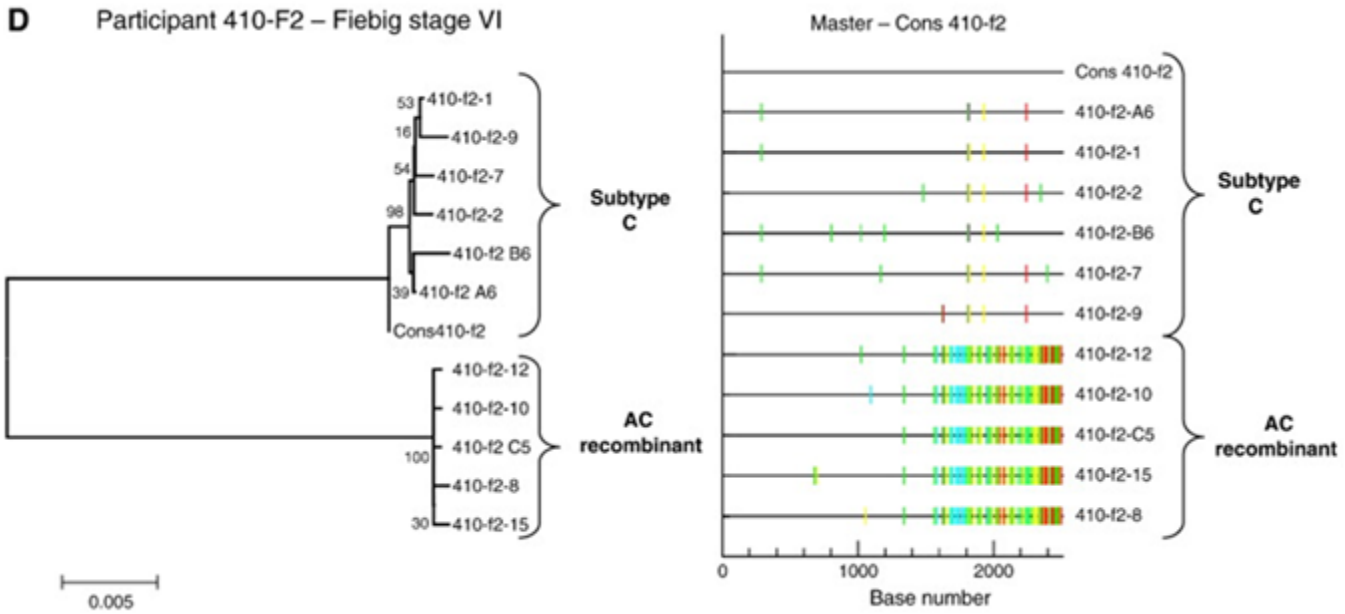


The swarm of HIV virions in an infected individual are not genetically identical



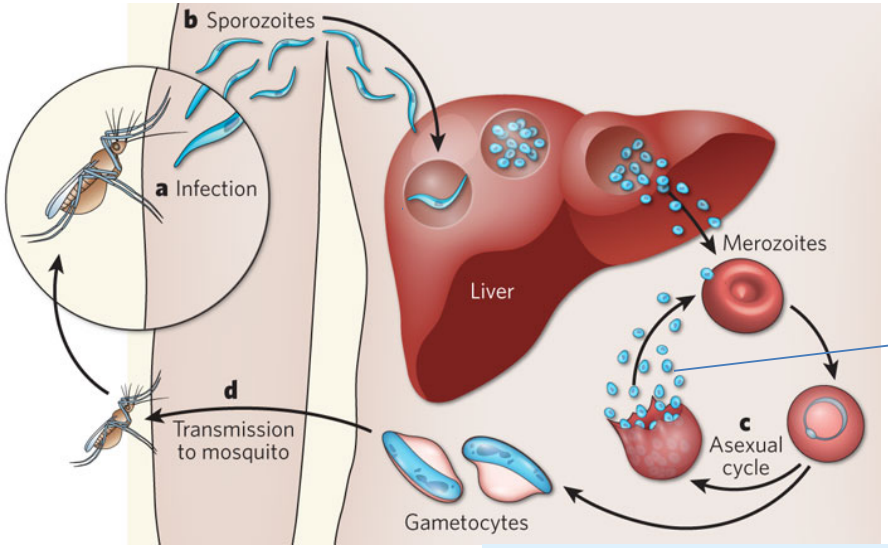
Virion 1	A T C T A T
Virion 2	A T G G C T
Virion 3	T T C T A T
CONSENSUS	A T C T A T

Founder Viruses Tell More Than Infection Yes/No



X=2

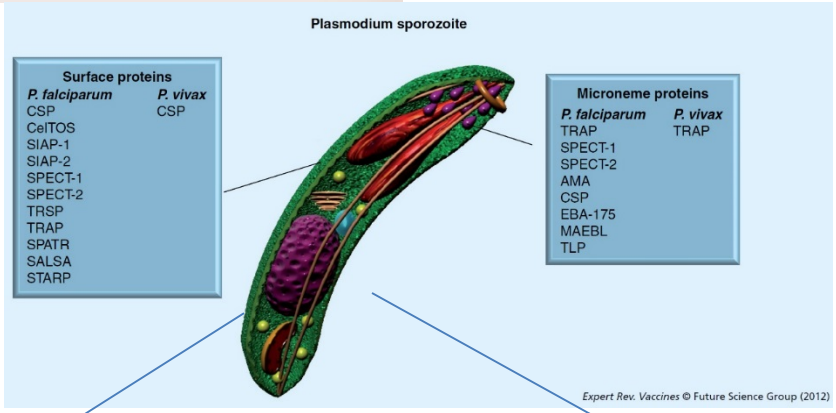
Malaria Sampling



Malaria life cycle

Sample blood stage parasites
PCR amplification of CS region
Then Next Gen sequencing.

NRNAN . . . EW
NRNEN . . . TW

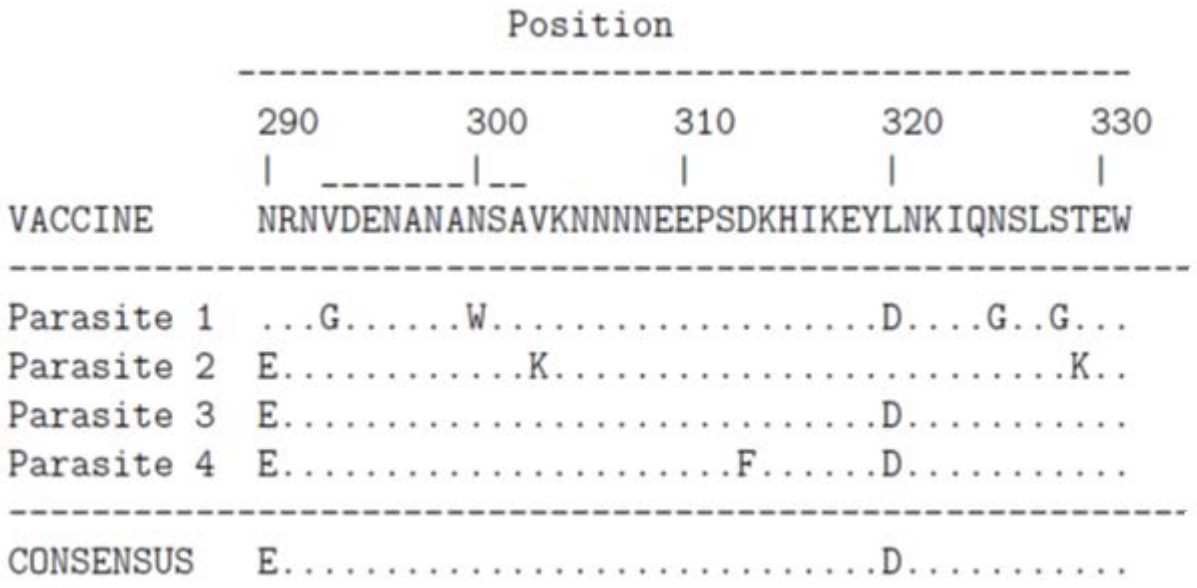


290 300 310 320 330

NRNVDENANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEW

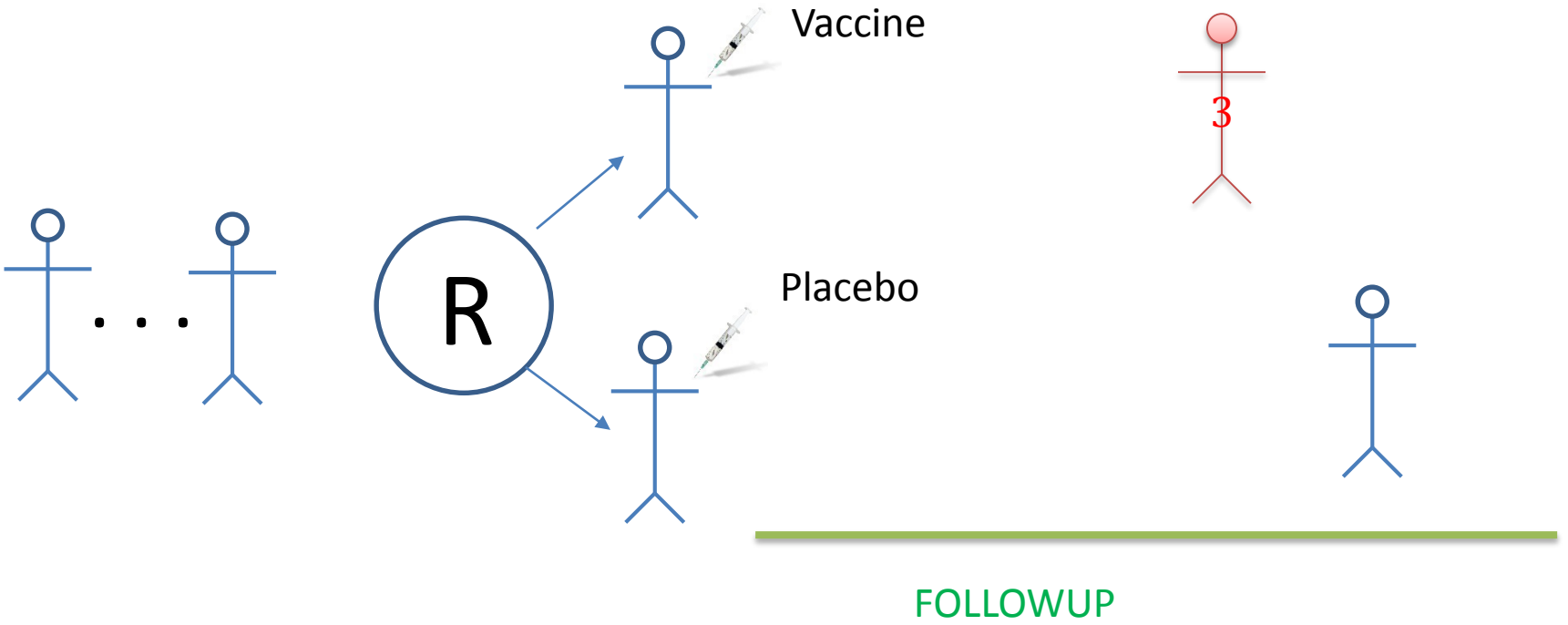
AA sequence of
Parasite used in
RTS,S/AS01 Vaccine

4 Founding Parasites

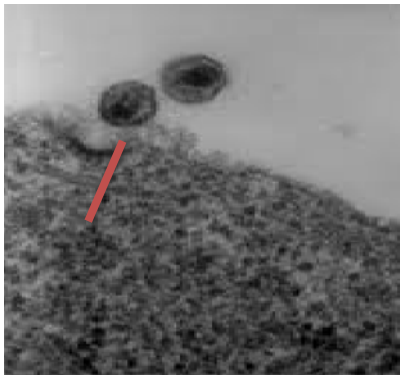


Vaccine Trial Redux

- Randomize at risk healthy volunteers to vaccine or placebo
- Follow them & count # infecting pathogens



Placebo Volunteer

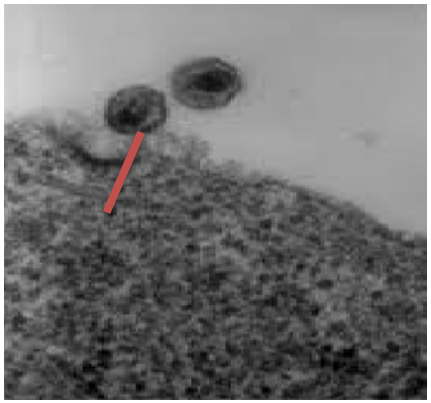


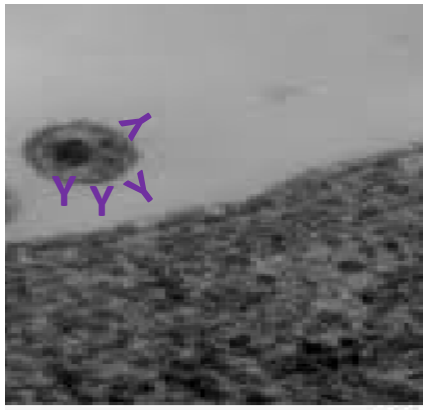
Cell infected

2 Virions infect cells

X = 2

Vaccine Volunteer





1 Virion infects a cell

Antibodies **Y** block infection

X=1

Both humans are infected, but the vaccine reduces founder viruses

Useful information that the vaccine is doing something

Mechanisms of Vaccine Protection

- All-or-none vaccine: a proportion of vaccinees are protected for *all* exposures.
- Leaky vaccine: chance of *human disease after exposure* is like flipping a coin w.p. Q
  - Q_v in vaccine arm Q_P in placebo arm
- Leaky leaky vaccine: chance of *pathogen infecting a cell* is like flipping a coin w.p. P
  - P_v in vaccine arm P_P in placebo arm

Smith et al 1984
Struchiner et al 1990
Halloran et al 1991

Vaccine Efficacy From the Virion's View

- Exposure has N virions. Each has probability p ($p\Delta$) of infecting a cell in a placebo (vaccine) recipient.
- Model X = # founder viruses
  - Vaccine $E(X) = N\,p\,\Delta = \mu\,\Delta$
  - Placebo $E(X) = N\,p = \mu$
- $VE_M = 1 - \frac{E(X|Z=1)}{E(X|Z=0)} = 1 - \Delta$

VE on the mean count

Per virion reduction in probability of infection
Holds for any mixture over μ

Efficiency gain using X in lieu of I(X>0)

- Suppose $X_1,\ldots,X_n \sim$ Poisson (μ)
- Dumb Method
  - Convert X to Y = I(X>0)
  - Estimate P(X>0) by avg(Y)
- Smart Method
  - Estimate $\hat{\mu}$ = avg(X)
  - Estimate P(X>0) by 1-exp(-$\hat{\mu}$)
- var (smart) /var(dumb) --- estimates of P(X>0)

μ = .25	μ= 1	μ= 3
1.1	1.7	5.8

Monkey Studies

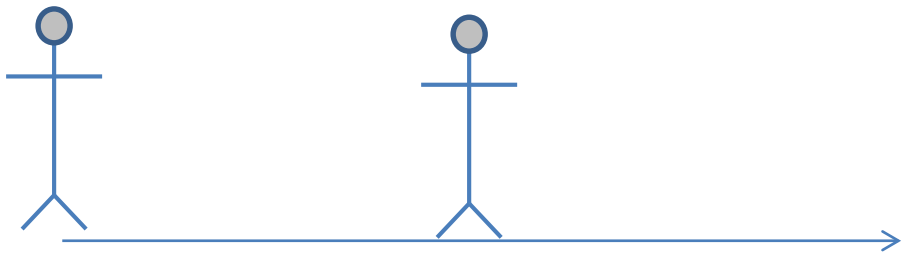
- Monkeys repeatedly *challenged* by exposing them to virus
- Assume X per challenge is Poisson($\mu\ \Delta^Z$)
- Likelihood contribution for a monkey infected on *third* challenge with 4 founder viruses.
  - P(X=0) P(X=0) P(X=4)
- Use maximum likelihood to estimate $\mu\ \Delta$
  - Form $\widehat{VE_M}$ =1-$\widehat{\Delta}$

Animal vs Human Experiments

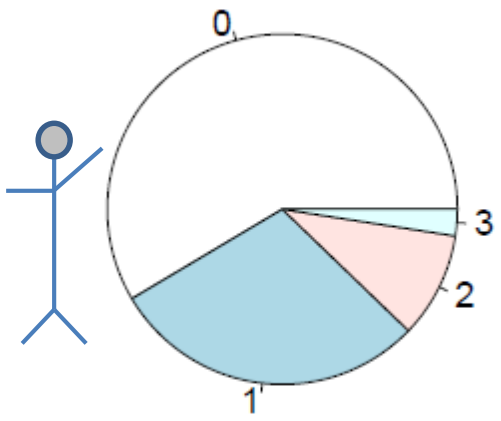
- Animal Experiments
  - Control exposure: N virions from known pool
  - Identify all Xs, even when X=0
- Human Field Trials
  - N=inoculum size uncontrolled and unknowable
  - Exposure not crisply defined
  - Exposures unknown unless infection occurs
    - *X=0 never seen*

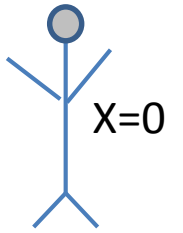
Casino Behavior

Placebo Queue



Placebo Roulette



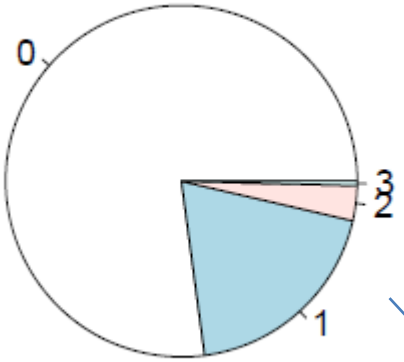


Uninfected

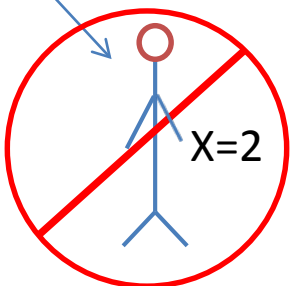
$\omega(t)$ = Instantaneous risk of gambling

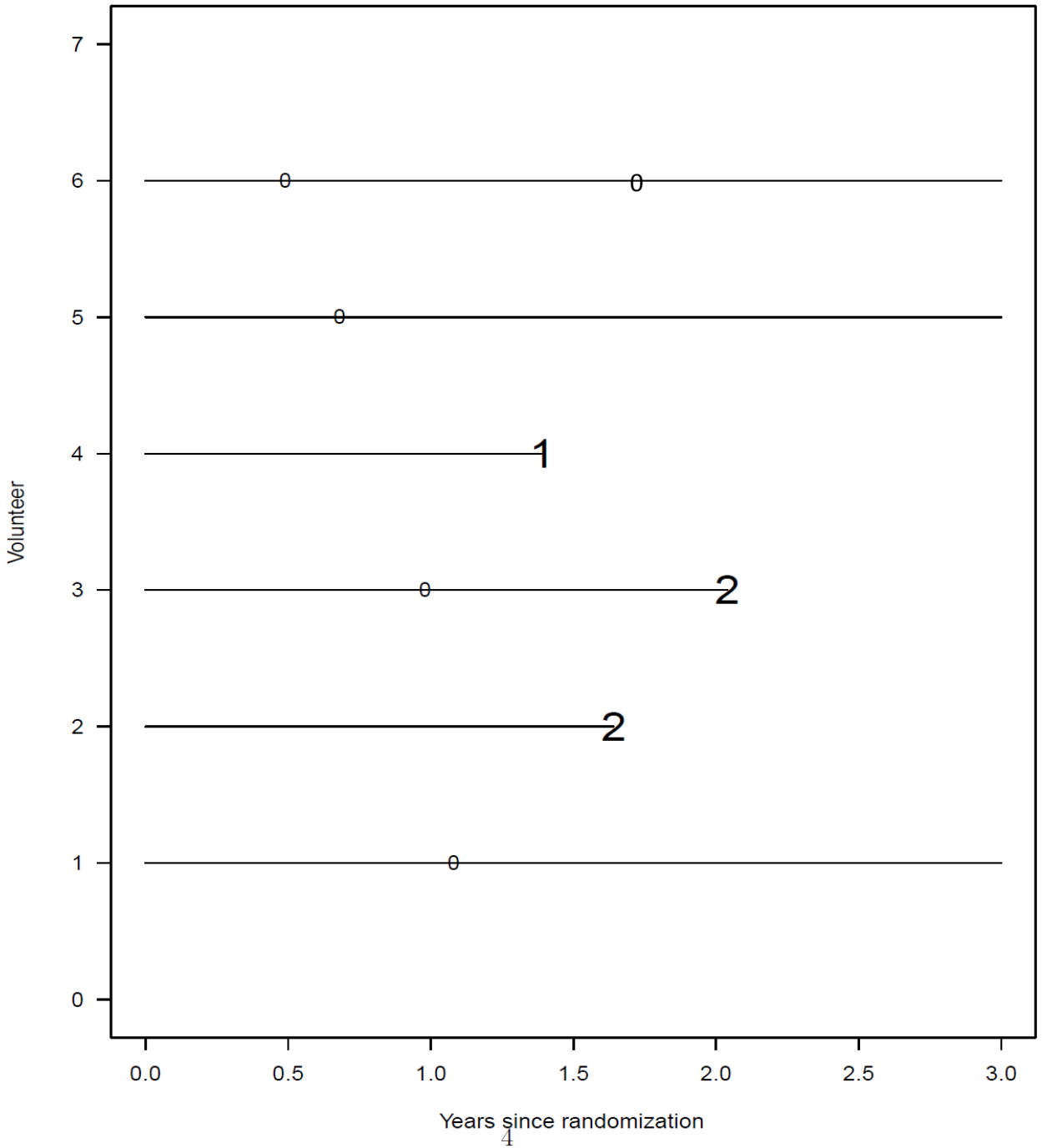
Vaccine Queue

Vaccine Roulette



Infected




Volunteer
7
6
5
4
3
2
1
0
0
0
1
2
2
0
0
0.0
0.5
1.0
1.5
2.0
2.5
3.0
Years since randomization

Cox Regression For Infection

- A model for the instantaneous risk of infection

$$h(t) = \omega(t)\, P(X>0 \mid Z=0) \quad \text{in placebo group}$$

$$h(t) = \omega(t)\, P(X>0 \mid Z=1) \quad \text{in vaccine group}$$

Risk of INFECTION

Risk of EXPOSURE
Same in both groups

Probability of infection, given exposure

Cox Regression 2

- No matter the distribution of X

$$h(t) = \omega(t)\ \{P_0\ (X>0)\}\ \exp\{\ \log\left(\frac{P_1(X>0)\}}{P_0(X>0)\}}\right) Z\ \}$$

$$= h_0(t)\ \exp\{\beta\ Z\}$$

- $\beta = \log\left(\frac{P_1(X>0)\}}{P_0(X>0)\}}\right)$

- $\exp(\beta)$ is the *per-exposure* reduction in the risk of infection

Truncated mean proportional to Untruncated mean

- $E(X) = \sum_{x=0}^{\infty} xP(X = x) = \sum_{x=1}^{\infty} xP(X = x)$

$$= \sum_{x=1}^{\infty} xP(X = x)\frac{P(X>0)}{P(X>0)}$$

$$= E(X|X>0)\,P(X > 0)$$

- Thus

$$E(X|X>0) = \frac{E(X)}{P(X>0)}$$

Multiply

- Multiplication produces a product estimate
- $e^{\widehat{\beta}} \frac{\overline{X_1}}{\overline{X_0}} \rightarrow \frac{P(X>0|Z=1)}{P(X>0|Z=0)} \frac{\frac{E(X|Z=1)}{P(X>0|Z=1)}}{\frac{E(X|Z=0)}{P(X>0|Z=1)}}$

$\bar{X}_Z$ mean number of virions on Z among infected (i.e. X>0)

The Product Method Estimate of Δ

- Multiplication produces a product estimate
- $e^{\hat{\beta}} \frac{\overline{X_1}}{\overline{X_0}} \rightarrow \frac{P(X>0|Z=1)}{P(X>0|Z=0)} \frac{\frac{E(X|Z=1)}{P(X>0|Z=1)}}{\frac{E(X|Z=0)}{P(X>0|Z=1)}} = \frac{E(X|Z=1)}{E(X|Z=0)} = \Delta$

$\bar{X}_Z$ mean number of virions on Z among infected (i.e. X>0)

- Truncated X data gets ratio of *untruncated* X^* means.
- X distribution unspecified
- Arbitrary intensity of exposure function ω (t)

Easy Asymptotics for Product Method

- $\log(\hat{\Delta}) = \log\left(e^{\hat{\beta}_{Cox}} \frac{\overline{X}_1}{\overline{X}_0}\right)$

$$\log\left(e^{\hat{\beta}_{Cox}} \frac{\overline{X}_1}{\overline{X}_0}\right) = \hat{\beta}_{Cox} + \log(\overline{X}_1) - \log(\overline{X}_0)$$

- Delta-method $\log(\overline{X}_Z) \approx N\left(\log(\mu_Z), \frac{\sigma_Z^2}{I_Z \mu_Z^2}\right)$

- $\log(\hat{\Delta}) \sim \mathrm{N}(\log(\Delta)\ ,\ \widehat{\mathrm{var}}(\hat{\beta}_{Cox}) + \frac{S_1^2}{I_1 \overline{X}_1^2} + \frac{S_0^2}{I_0 \overline{X}_0^2}$

Product Method w/ Exponential Dbn

- Product estimate under exponential time to infection

$$\widehat{\Delta} = \left(\frac{I_1}{T_1} \Big/ \frac{I_0}{T_0}\right) \frac{\bar{X}_1}{\bar{X}_0} = \left(\frac{X_{1+}}{T_1} \Big/ \frac{X_{0+}}{T_0}\right)$$

where I_Z total number of infections on Z

T_Z total follow-up time on Z

X_{Z+} total number of virions on Z

$\bar{X}_Z$ mean number of virions on Z

V:P Hazard ratio estimate

Monkey Studies-know all exposures

- 10 on placebo: 1, 2, … ,10

$$\hat{\mu} = \frac{8 + 0+0+2 + \dots + 0+0+7}{1+3+\dots 3} = \frac{179}{57} = \frac{X_{0+}}{N_0}$$

- 10 on vaccine 1, 2, … ,10

$$\widehat{\mu\,\Delta} = \frac{0+0+4 + 0+\dots+0 + \dots + 0+1}{3+8+\dots 2} = \frac{75}{113} = \frac{X_{1+}}{N_1}$$

- $\hat{\Delta} = \left(\frac{X_{1+}}{N_1} / \frac{X_{0+}}{N_0}\right)$

10^{th} vaccine monkey
Infected at 2^{nd} exposure
With 1 founding Pathogen

Product Method Analogous to Estimator from Monkey Studies

- Product estimate under exponential time to infection

$$\widehat{\Delta} = \left(\frac{I_1}{T_1} \Big/ \frac{I_0}{T_0}\right) \frac{\bar{X}_1}{\bar{X}_0} = \left(\frac{X_{1+}}{T_1} \Big/ \frac{X_{0+}}{T_0}\right)$$

N_1 N_0

where N_Z total number of challenges on Z

Product method replaces total number of challenges with total time at risk

V:P Hazard ratio estimate

Concerns

- Same ω (t) for all
  - Some may have more frequent exposures
- One dbn of X for all in same group
  - Some individuals have poorer mucosal barriers...more virions get in.
- Measured covariates can address concerns

Incorporating Covariates

- Covariates for time to exposure: W^E
  - e.g. I(>3 sexual partners last month at baseline)
  - $h(t) = h_0(t) \exp(Z \beta + \theta W^E)$ . . . *product method*

- Covariates that impact X: W^X
  - e.g. damaged cells, immune response to vaccine, closeness of infecting virus to vaccine insert
  - Natural to have $E(X^*) = e^{\varphi_0+\varphi_1 Z+\varphi_2 W+\varphi_2 WZ}$

Weighted Estimating Equations

- WEE = X-weighted Cox score equation

$$\sum_{i=1}^{n} \int_0^{\infty} X_i \left\{ Z_i - \frac{\sum_{i=1}^{n} Z_i \Delta^{Z_i} I(Y_i \geq t)}{\sum_{i=1}^{n} \Delta^{Z_i} I(Y_i \geq t)} \right\} dN_i(t)$$

- Above a functional of empirical processes. Asymptotics for $\hat{\Delta}$ from functional delta method.
- . . . but generalizes to handle both W^E & W^X

Example HIV

- VAX003 randomized 2,546 Thai IDUs to HIV vaccine AIDSVAXB/E or placebo
  - 211 infections reported 105:106 V:P
- $VE_I = 1 - e^{-.00245} = .002$

Product Method Estimate of VE_M

- 39 volunteers, # founder viruses determined
  - High risk (IDU) volunteers
  - Infection detection within 100 days
- Mean X in vaccine 1.33, placebo 1.67

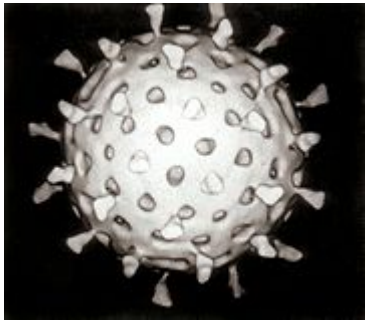
$$VE_M = 1 - e^{-.00245} \frac{1.33}{1.67} = .21$$

95% delta-method CI(-.33, .52)

Sieve Methods With Pathogen Counts

- A pathogen species can have distinct strains
  - Serotypes ---- different surface antigens
  - Genetics ---- different DNA or RNA
- Vaccines may protect differentially against the different strains
  - Vaccine induced antibodies may protect well against some strains but not others.
  - Vaccines may induce CD4 & CD8 T-cells with differential protection
    - HIV, malaria, Ebola

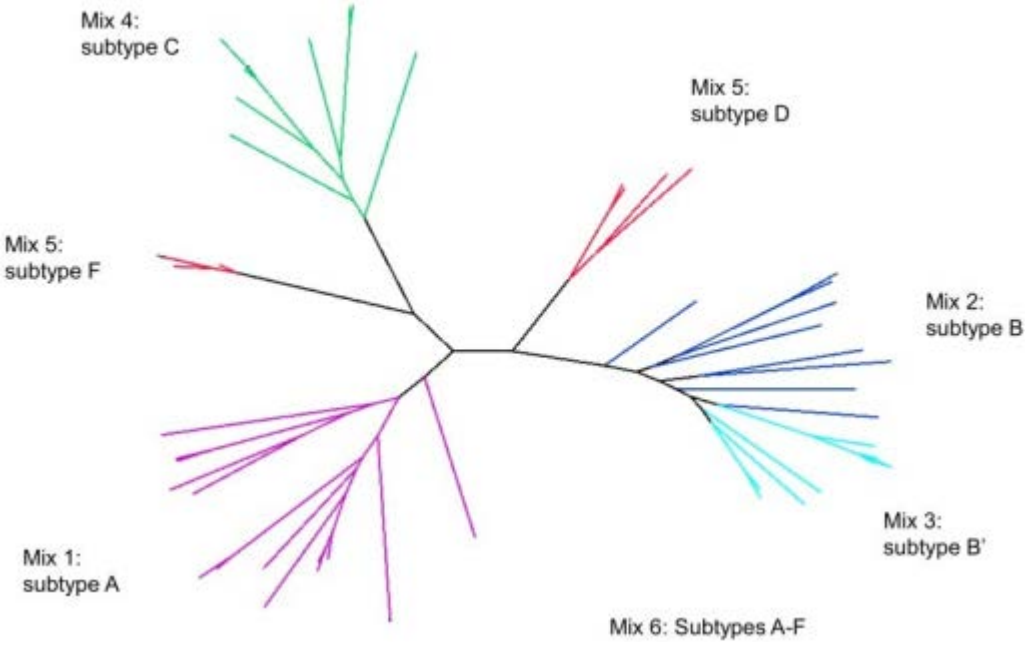
Pathogens are diverse



Rotavirus

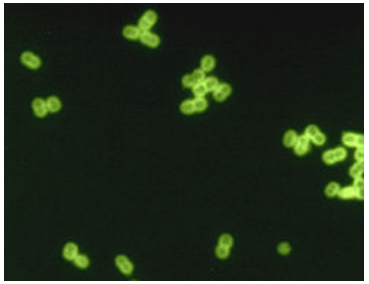
5 major serotypes

HIV multiple genotypes



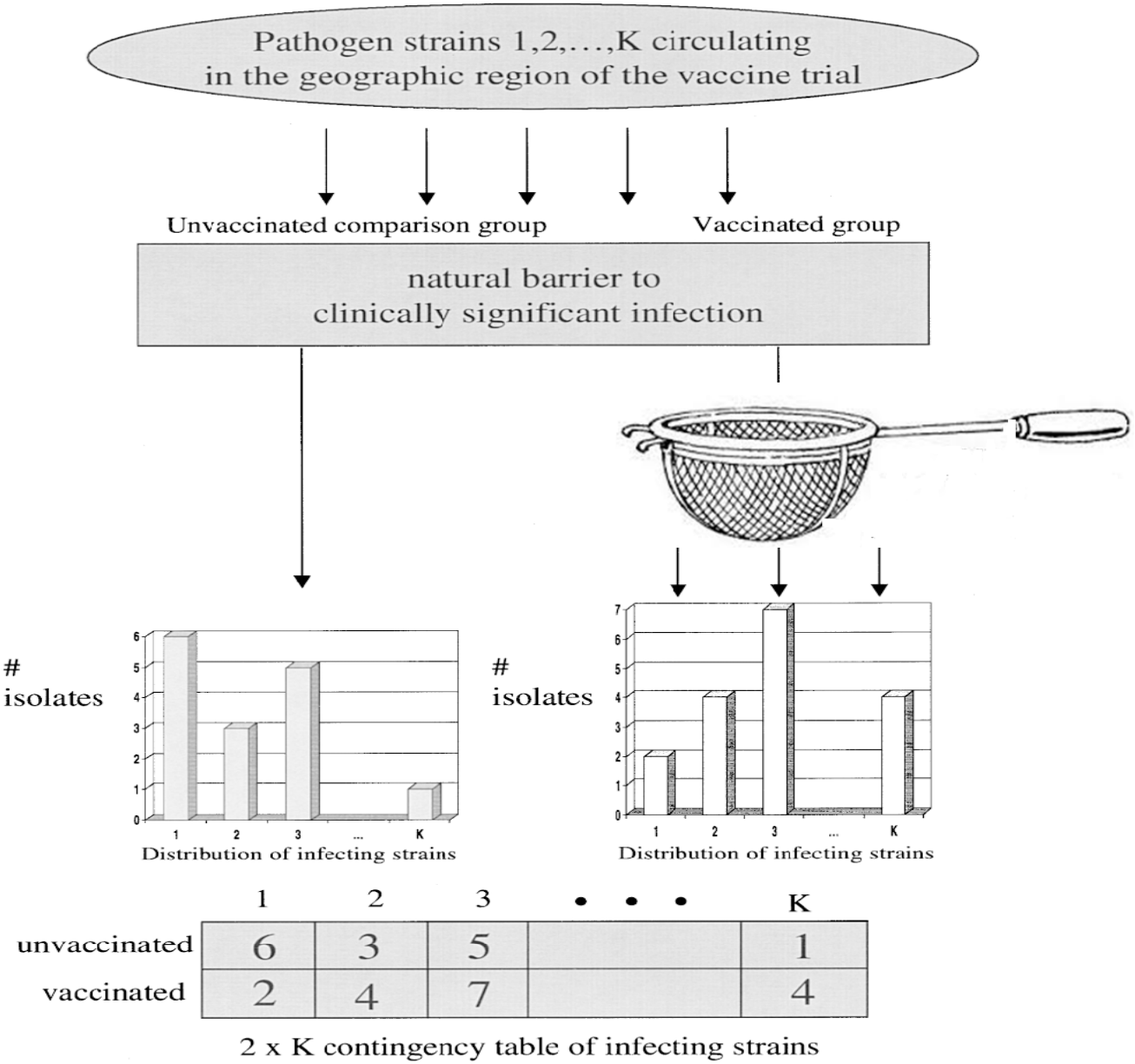
Bowles et al PLoS One 2014

Streptococcus pneumoniae



90+ serotypes


THE
SIEVE

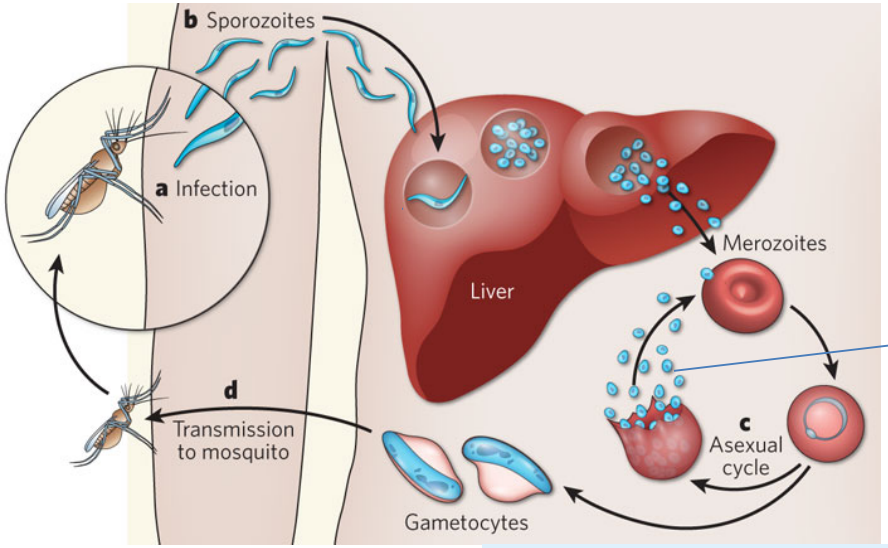


	1	2	3	• • •	K
unvaccinated	6	3	5		1
vaccinated	2	4	7		4

2 x K contingency table of infecting strains

Ref: Gilbert et al 2001

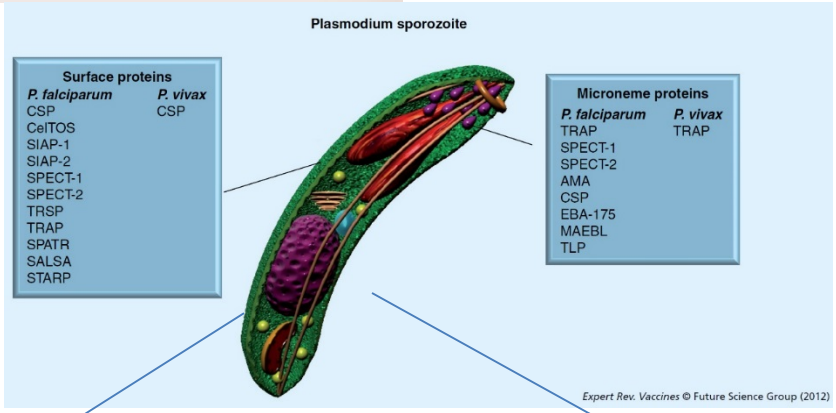
Malaria Sampling



Malaria life cycle

Sample blood stage parasites
PCR amplification of CS region
Then Next Gen sequencing.

NRNAN . . . EW
NRNEN . . . TW



290 300 310 320 330

NRNVDENANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEW

AA sequence of
Parasite used in
RTS,S/AS01 Vaccine

of Founding Parasites

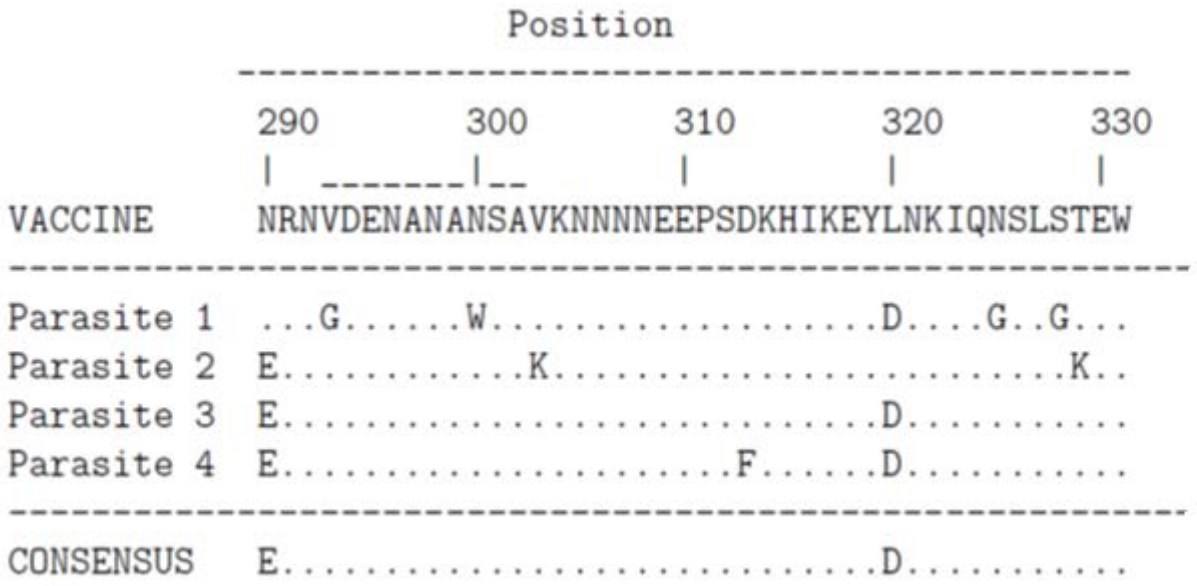


Table 1

Partial amino acid sequence of the RTS,S vaccine immunogen along with an illustration of 4 infecting parasites that could have been recovered from an infected volunteer in a vaccine trial along with the position-wise most popular sequence, i.e. the consensus sequence.

	Position							
	290	300	310	320	330	match at	match in	total mismatches
VACCINE	NRNVDENANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEW					320	293-302	290-331
Parasite 1	...G.....W.....D....G..G...					0	0	5
Parasite 2	E.....K.....K..					1	1	3
Parasite 3	E.....D.....					0	1	2
Parasite 4	E.....F.....D.....					0	1	3
CONSENSUS	E.....D.....					0	1	2

X_1, X_2=(# match at 320, # mismatch at 320) = (1,3)

Table 1

Partial amino acid sequence of the RTS,S vaccine immunogen along with an illustration of 4 infecting parasites that could have been recovered from an infected volunteer in a vaccine trial along with the position-wise most popular sequence, i.e. the consensus sequence.

	Position							
	290	300	310	320	330	match at	match in	total mismatches
VACCINE	NRNVDENANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEW					320	293-302	290-331
Parasite 1	...G.....W.....D....G..G...					0	0	5
Parasite 2	E.....K.....K..					1	1	3
Parasite 3	E.....D.....					0	1	2
Parasite 4	E.....F.....D.....					0	1	3
CONSENSUS	E.....D.....					0	1	2

X_a = # of infecting pathogens with 'a' total mismatches in 290-331

$X_0, X_1, X_2, X_3, X_4, X_5, \ldots$ = (0,0,1,2,0,1,0,0,0

0 1 2 3 4 5 . . . ..

Table 1

Partial amino acid sequence of the RTS,S vaccine immunogen along with an illustration of 4 infecting parasites that could have been recovered from an infected volunteer in a vaccine trial along with the position-wise most popular sequence, i.e. the consensus sequence.

	Position							
	290	300	310	320	330	match at	match in	total mismatches
	DV10 Region						DV10 Region	
VACCINE	NRNVDENANANSAVKNNNNEEPSDKHIKEYLNKIQNSLSTEW					320	293-302	290-331
Parasite 1	...G.....W.....D....G..G...					0	0	5
Parasite 2	E.....K.....K..					1	1	3
Parasite 3	E.....D.....					0	1	2
Parasite 4	E.....F.....D.....					0	1	3
CONSENSUS	E.....D.....					0	1	2

DV10 Region

$X_1 X_2$=(# match DV10 region, # mismatch DV10 region) = (3,1)

New type of data

- Before, used the consensus strain
  - Y_a = 1 if infected by `strain` a, else 0
  - e.g. (Y_1, Y_2) = (1,0) or (0,1)

- Now, get # infecting pathogens of each type
  - X_f = number of infecting pathogens with feature f

    e.g. (X_1, X_2) = (2,0) or (3,1)

Analysis of New Data

- Can we *shoehorn* this data with multiple infecting strains into existing methods for a single infecting strain?
- Can we *develop* new methods that explicitly account for multiple infecting strains?

Shoehorn: Within Cluster Resampling aka Multiple Outputation

1) Randomly pick a single pathogen for each infected person
   - Fred 4 unique strains: 1 match 3 mismatch
   - Pick a strain at random e.g. mismatch
2) Run a standard sieve analysis
   - VE(match) = .65 VE(mismatch) = .51
3) Repeat many many many times and average.

Within Cluster Resampling Schematic

Resample #	Dataset	VE(match)	VE(mismatch)
1	D_1 →	65.1	42.1
2	D_2 →	51.2	53.4
3	D_3 →	71.3	38.1
4	D_4 →	61.3	47.8
9999	D_{9999} →	52.1	38.9
10000	D_{10000} →	63.2	54.1
AVERAGE		63.1	53.9

There is an easy way to get a p-value for within cluster resampling.

Easy Inference With WCR

- Each resample gives estimates of the parameter and its variance
  - $P_1\ V_1$, $P_2\ V_2$, . . . , $P_{10000}\ V_{10000}$
- Calculate 3 Statistics
  - Average the P_i, → $\bar{P}$
  - Average the V_i → $\bar{V}$
  - Sample variance of the P_i → S^2

$$\frac{\bar{P}}{\sqrt{\bar{V}-S^2}}$$ is standard normal on the null

Easy Inference With WCR

- Each resample gives estimates of the parameter and its variance
  - $P_1\ V_1$, $P_2\ V_2$, . . . , $P_{10000}\ V_{10000}$
- Calculate 3 Statistics
  - Average the P_i, → $\bar{P}$
  - Average the V_i → $\bar{V}$
  - Sample variance of the P_i → S^2

$$\frac{\bar{P}}{\sqrt{\bar{V}-S^2}}$$ is standard normal on the null

WCR

- WCR can be used whenever you have a statistical procedure P that requires 1 outcome per person, but you have multiple outcomes.
- Can be used in lieu of GEE
  - Like exchangeable with rho -> 1
    - One person, one vote
  - Opposite of working independence rho=0
    - One pathogen, one vote

WCR = t-test on cluster means

- Test means of two groups X vs Y

$x_{11}\ x_{12}\ x_{13}$ → $\bar{x}_1$

$x_{21}\ x_{22}$ → $\bar{x}_2$

x_{31} → $\bar{x}_3$

$y_{11}\ y_{12}\ y_{13}\ y_{14}$ → $\bar{y}_1$

$y_{21}\ y_{22}$ → $\bar{y}_2$

Sieving at DV10 Region

DV10 Region	RTS,S Vaccine # Events	Control Vaccine (% Incidence)	VE
Match	90 (2.5)	86 (5.6)	63.1
Mismatch	1091 (30.8)	822 (53.7)	53.9

Averaged over 1000s of synthetic data sets with 1 Strain per person

- Test of equal VE has p=.04
- Some evidence of sieving.

Neafsey et al 2015

New Methods

- Let's develop new methods that explicitly uses the counts
- Passive surveillance
  - Get $(X_1, X_2) = (0,0)$ or $(3,1)$ or $(2,0)$ at end of study
- Active surveillance
  - Get time of infection detection and
  - Get $(X_1, X_2) = (0,0)$ or $(3,1)$ or $(2,0)$

Passive and active surveillance

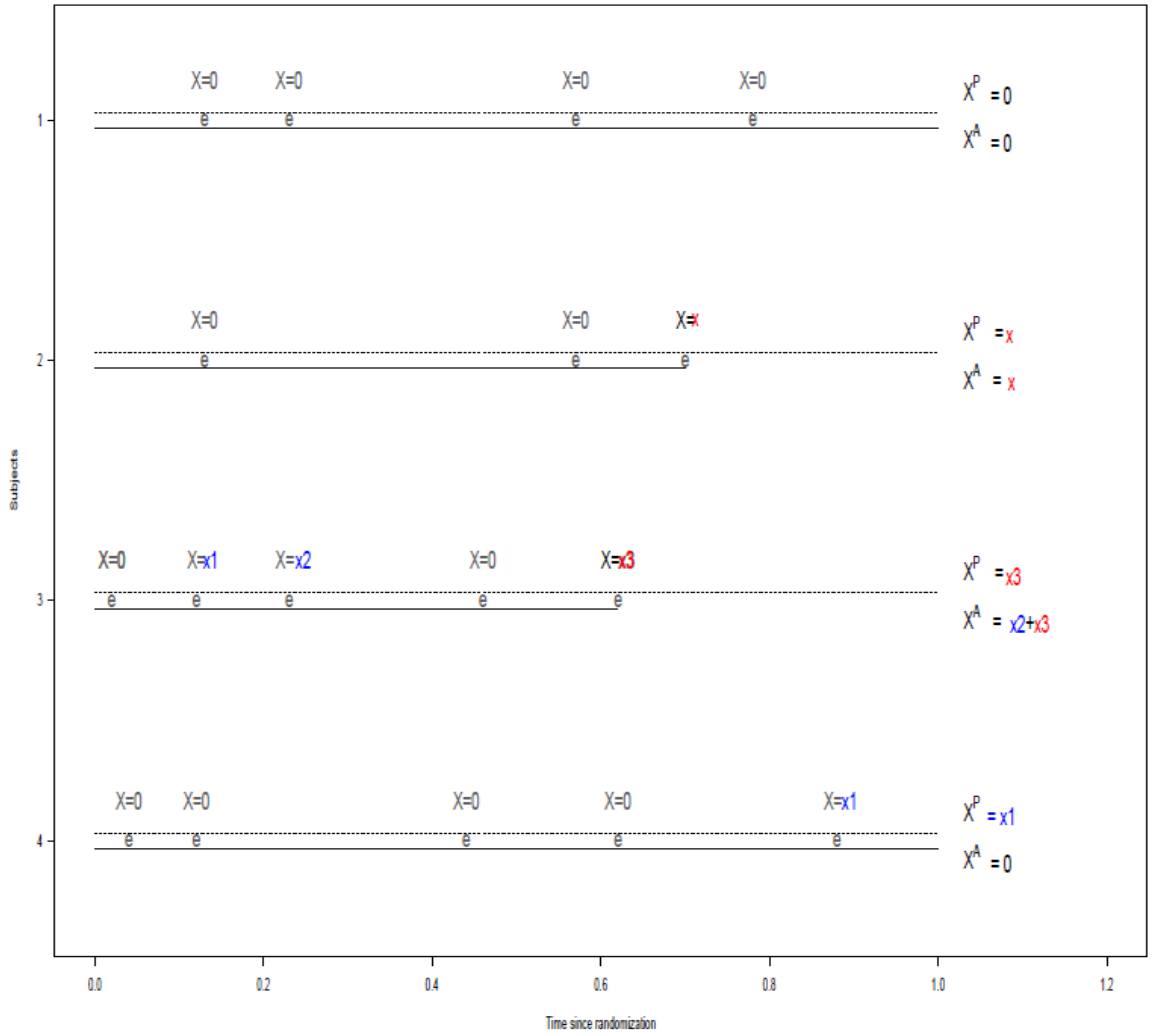


Figure 1. Trajectories of the exposure, infection (possibly subclinical), and terminal infection processes for four individuals Exposure=e at which time $\boldsymbol{X}$ is drawn from $F_Z()$. $\boldsymbol{X}^A$ and $\boldsymbol{X}^P$ are the vector of counts obtained under active (solid line) and passive (dashed line) surveillance. $\boldsymbol{X}$s which result in subclinical infections are blue while $\boldsymbol{X}$s that result in terminal infections are red. Trajectory 2 corresponds to a disease where all infections are terminal (e.g. HIV) while trajectories 3 and 4 correspond to a disease with subclinical infections (e.g. malaria). Note that we allow that old sub-clinical infections may be cleared (e.g. $x1$ from trajectory 3 under active surveillance).

Passive Surveillance: Modern Data & Analysis

Group	X_1	X_2
Vaccine	1	0
Vaccine	0	0
Placebo	3	0
Placebo	2	4
Vaccine	0	2
Placebo	0	0

Placebo group **5 mismatched out of 9**
Vaccine group **1 mismatched out of 3**

Passive Surveillance
Single Pathogen Data & Analysis

Group	X_1	X_2
Vaccine	1	0
Vaccine	0	0
Placebo	1	0
Placebo	0	1
Vaccine	0	1
Placebo	0	0

Placebo group **1 mismatched out of 2**
Vaccine group **1 mismatched out of 2**

Passive Surveillance

- Use *bivariate negative binomial*
  - X_{si} Poisson exp{b_i + B0 + B1 Z + B2 I(s=1) + B3 Z I(s=1) }
  - s=1,2 i=1,…n subjects exp(b_i) ~ Gamma (μ, V)
  - Z= vaccine indicator
- Estimation
  - GEE with working independence
  - Single Pathogen
  - Exhaustive WCR

Sieve effect if B3 is nonzero

Simulation

- X~ bivariate negative binomial
  - $\exp(b_i)$ ~ Gamma(.5,v) v=0,1,2
- Counts: Binomial (= GEE-I), WCR
- Infection: Bernoulli

SIMULATION VARIANCE OF Sieve effect B3					
v= Variance	GEE (new)	Single Pathogen	WCR (shoehorn)	Variance Ratio Single/GEE	WCR/GEE
0	.066	.139	.083	2.1	1.3
1	.072	.170	.109	2.4	1.5
2	.047	.201	.090	4.2	1.9

Active Surveillance

- Record T – time to infection or censoring
- X_{is} - # of 's' pathogen s=1,…,S
- Assume risk of exposure is
  - $\omega(t)\ \exp(\alpha_I W^E_i)$
- Assume *per exposure* mean is
  - $E(X_{is}) = \exp(\alpha_X W^X_{is})$ $\quad\quad$ $\dim(\alpha = \alpha_X + \alpha_I) = p$
  - W^X includes vaccine indicator

WEE estimation

- WEE solution equivalent to Cox regression with preprocessing and weighting:
  - Stack S datasets one for each infection type
  - Weight *failure* i by X_{is} in dataset s=1… S
- Generalizes the method of Lunn & McNeil (1995)
  - From 2 competing risks to S *concurrent* risks
  - Allows for failure weights other than 1 (use X_{is})
  - Allow covariates to parsimoniously model effect on S events

COUNT ENDPOINT WITH TIME CONSTANT VE_{Mf}
ONE DATASET FOR EACH EVENT

$$E(X_s|Z=1)/E(X_s|Z=0) = \exp(\beta_0 + \beta_1(s-1))$$

```
coxph(Surv(T,D)  ~ W1+W2 + strata(FailCount))
```

ID	T	Δ	Z	X_1	X_2	X_3
1	1.1	1	0	2	0	1
2	2.4	1	1	0	1	1
3	3.8	1	0	1	0	0
4	5.0	0	1	0	0	0

ID	T	D	Z	Feature	W_1	W_2	FailCount
1	1.1	1	0	1	0	0	1
2	1.1	0	1	1	0	0	1
3	1.1	0	0	1	0	0	1
4	1.1	0	1	1	0	0	1
1	1.1	1	0	1	0	0	2
2	1.1	0	1	1	0	0	2
3	1.1	0	0	1	0	0	2
4	1.1	0	1	1	0	0	2
1	1.1	1	0	3	0	0	3
2	1.1	0	1	3	0	0	3
3	1.1	0	0	3	0	0	3
4	1.1	0	1	3	0	0	3
2	2.4	1	1	2	1	1	4
3	2.4	0	0	2	0	0	4
4	2.4	0	1	2	0	0	4
2	2.4	1	1	3	1	2	5
3	2.4	0	0	3	0	0	5
4	2.4	0	1	3	0	0	5
3	3.8	1	0	1	0	0	6
4	3.8	0	1	1	0	0	6

RTS,S vaccine trial

- No Malaria vaccine yet.
- RTS,S vaccine targets circumsporozoite protein & has partial efficacy.
- GSK conducted a phase III trial in 11 sites across 7 African countries in 8922 children randomized 2:1 V:P, N=8922.
- Does vaccine show a sieve effect?

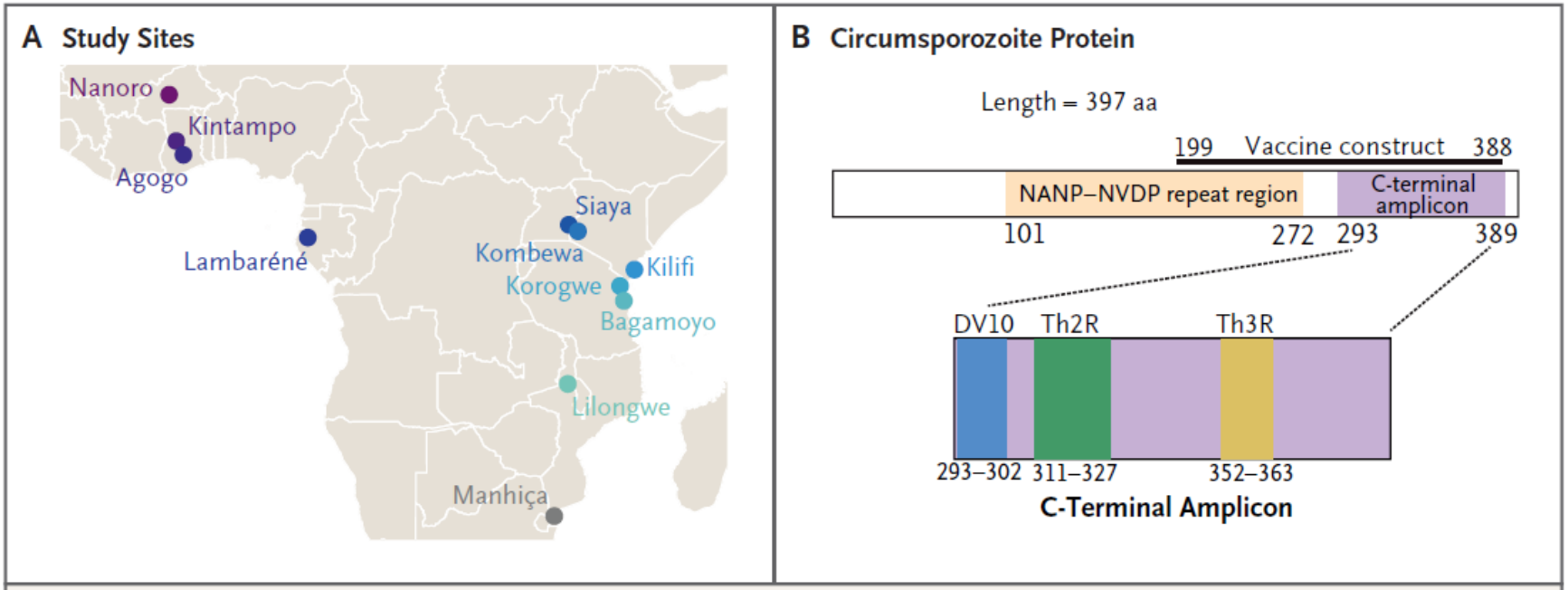
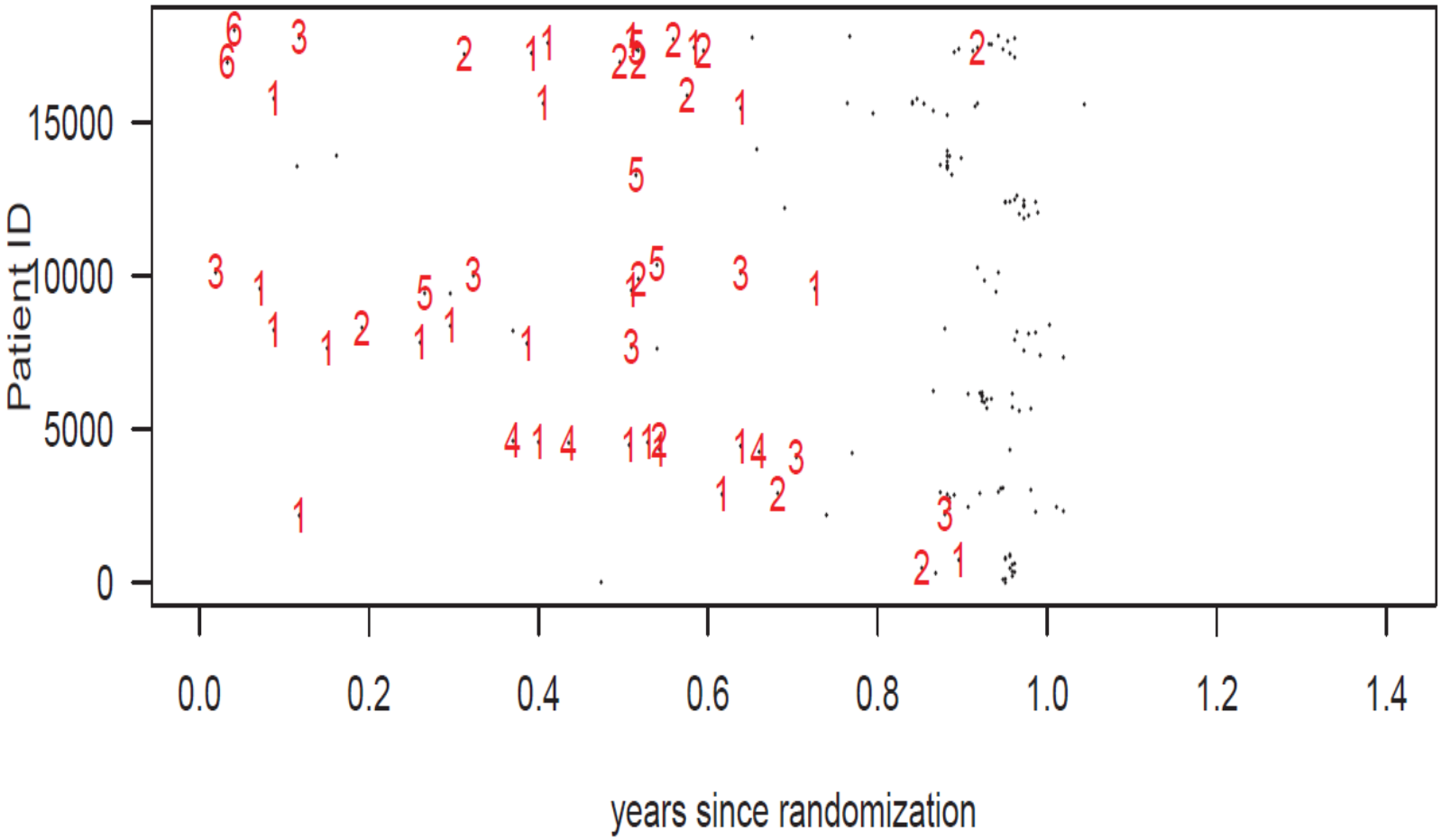


Figure 1. Study Sites and Genomic Units.


Patient ID
15000
10000
5000
0
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
years since randomization

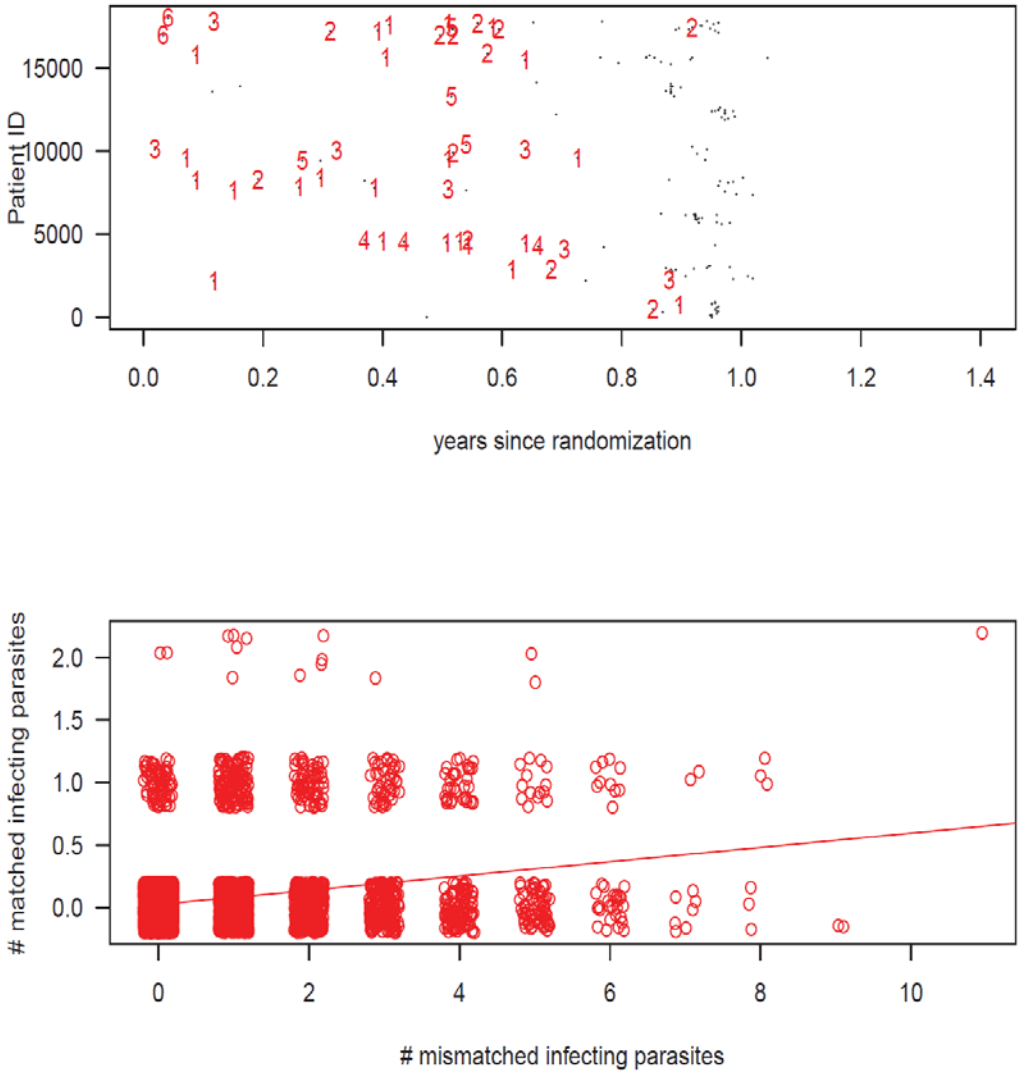


Figure 2. Top Panel: Years to censoring or first/only episode of clinical malaria for 165 randomly selected children. Small black dots denote censoring, red numbers provide the number of infecting pathogens at the time of the detection of infection. Bottom Panel: a scatterplot of the count of infecting pathogens by mismatch/match to the DV10 region.

Table 3

Different methods of estimating differential vaccine efficacy applied to the DV10 region of the circumsporozoite protein. Data from a phase 3 trial of the RTS,S/AS01 malaria vaccine in African infants. The differential VE parameter $\alpha_{2U} = \log\{(1\text{-}VE_{U1})/(1\text{-}VE_{U2})\}$, where $U = I$ for infection indicator or M for mean. 95% confidences in parentheses.

	VE_{If} on Infection			VE_{Mf} on Count
Parameter	One Parasite	10,000 Monte Carlo WCR	Product Method on $I(X_f > 0)$	Product Method on X_f
Matched: VE_1	.55 (.39,.66)	.56 (.44,.65)	.60 (.51,.67)	.61 (.52,.68)
Mismatched:VE_2	.43 (.38,.48)	.43 (.38,.48)	.44 (.39,.49)	.52 (.46,.56)
		Sieving Effect		
$\hat{\alpha}_2$	-.219	-.245	-.324	-.211
$\hat{\text{var}}(\hat{\alpha}_2)$	.0244	.0150	.0097	.0100
$\hat{\alpha}_2/\sqrt{\hat{\text{var}}(\hat{\alpha}_2)}$	-1.40	-2.00	-3.29	-2.10

Malaria Conclusions

- Different methods WCR WEE-I WEE-M estimate different things
  - all useful
- If only power matters.
  - WEE-I more significant than WEE-M than X_1 ,X_2
  - WEE-I more significant than WCR
    - Later simulations show this is generally true,

      WEE-I bigger estimand with smaller variance than WCR

Summary

- New technology allows us to count the number of clonally unique infecting pathogens
- Leads to a new vaccine metric, reduction in the mean number of infecting pathogens
- Can be used for overall VE and for strain-specific VE.

Poliomyleitis caused by *poliovirus*

- Poliomyleitis is a viral disease that can infect the central nervous system and cause lasting disabilities in a small number of infected individuals.
- Polio infection is most common in children but adults are at risk too
  - Franklin Roosevelt developed polio
- Polio was greatly feared.
  - Outbreaks are unpredictable
  - Paralyzed children are a visual reminder

- National Foundation for Infantile Paralysis was formed in

  1938 to develop a vaccine.

Key developments

- Virus was isolated in infected subjects 1908
- Identification of three serotypes of polovirus, each serotype has a distinctive surface and a specific antibody works against a specific type.
- Confirmation that neutralizing (blocking) antibodies protect against disease
  - At risk children who received antibodies from polio survivors saw 80% reduction in paralytic poliomyelitis compared to children with gelatin
- Growth of virus in cell culture
  - Allows production of vaccine—germ bits

Vaccine developments

- Inactivated polio vaccine (IPV):
  - Three serotypes grown in cell culture and then killed by formalin
  - Developed by Jonas Salk, injected
  - Can’t cause disease
- Oral polio vaccine (OPV):
  - Three serotypes were weakened by repeated passage in cold non-human cells
  - Replicates in the gut. Very rarely causes disease or mutates to a more virulent form
  - Developed by Sabin, swallowed

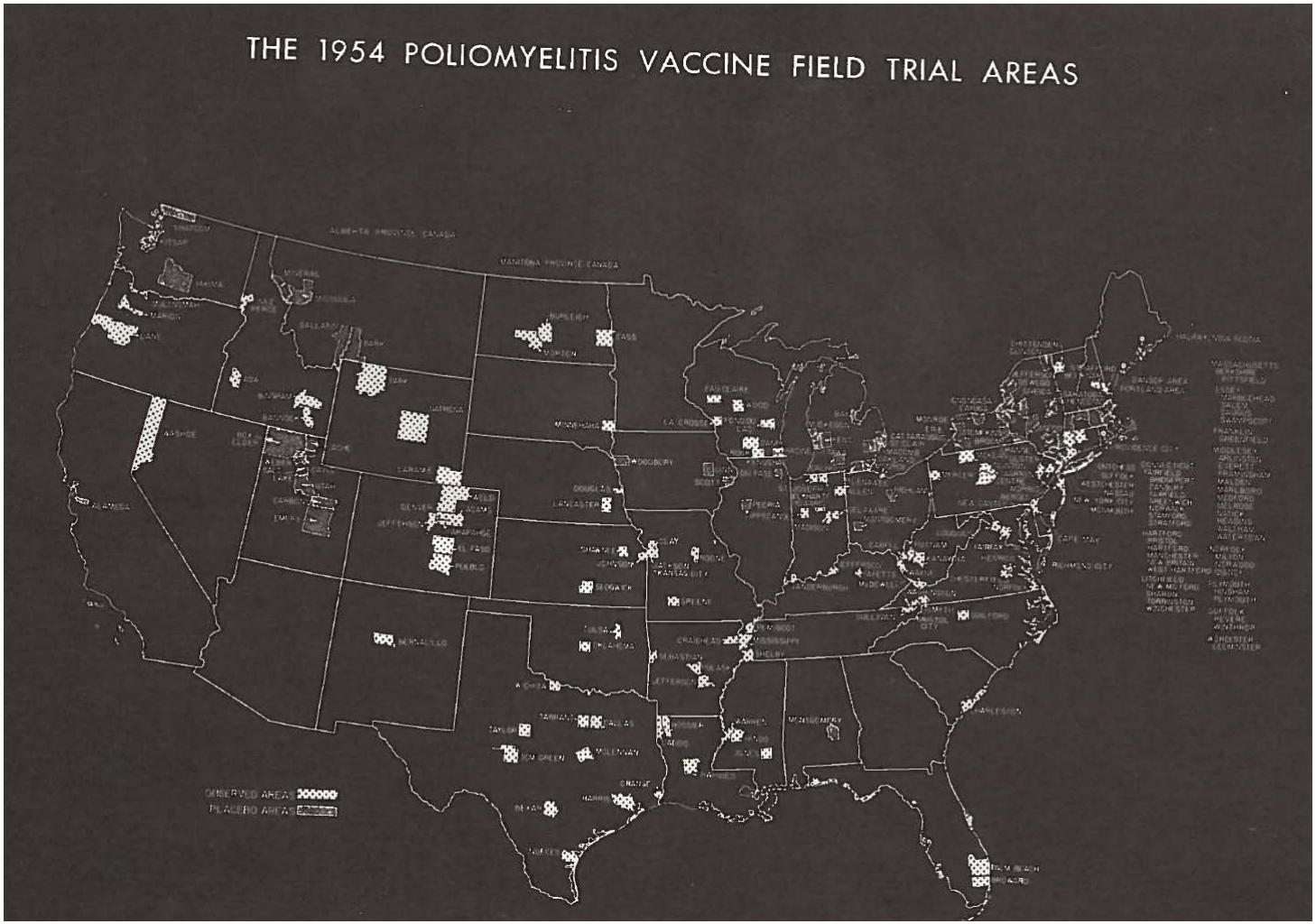
1954 Polio Field Trial of Salk Vaccine

- Salk Vaccine was promising but unproven.
- A field trial was essential. Earlier killed vaccines had some unkilled virus that lead to disease
- Intense publicity about the vaccine. Trial needed to be done in a single season
- Rate of paralytic polio by region was highly variable.

Key Features of Trial

- Two studies
  - Blinded placebo controlled individually randomized study in 84 areas in 11 states. Children in grades 1-3 randomized.
  - Observational trial 127 areas in 33 states. Children in grade 2 vaccinated

    grades 1 and 3 received nothing. Helped public support
- Conducted in spring and summer of 1954
  - Enrollment took long---vaccinations into mid June
  - Antibodies measured


THE 1954 POLIOMYELITIS VACCINE FIELD TRIAL AREAS
OBSERVED AREAS
PLACEBO AREAS

All Local Teams Are

Mother Objects To Son's Girl

Journal American

7 SPORTS WALL ST. SPECIAL

SALK'S VACCINE WORKS!

Official Count

Polio Vaccine Reported 80 to 90% Effective

Table 5

DIAGNOSTIC CLASS BY VACCINATION STATUS OF STUDY CASES PLACEBO AND OBSERVED AREAS

Vaccination Status	Study Population	Total Study Cases		Poliomyelitis						Doubtful Poliomyelitis		Not Poliomyelitis	
				Total		Paralytic		Nonparalytic					
		Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate	Number	Rate
All Areas - Total	1,829,916	1,012	55	858	47	682	37	176	10	66	4	88	5
Placebo Areas - Total	749,236	428	57	355	47	267	36	88	12	24	3	49	7
Vaccinated	200,745	81	40	56	28	33	16	23	11	10	5	15	7
Placebo	201,229	162	81	138	69	110	55	28	14	7	3	17	8
Incomplete Vaccinations	8,484	2	24	2	24	2	24	-	-	-	-	-	-
Incomplete Placebo Injections	8,577	6	70	6	70	4	47	2	23	-	-	-	-
Not Inoculated	330,201	177	54	153	46	118	36	35	11	7	2	17	5
Observed Areas - Total	1,080,680	584	54	503	47	415	38	88	8	42	4	39	4
Vaccinated	221,998	75	34	55	25	38	17	17	8	12	5	8	4
Controls	725,173	440	61	391	54	331	46	60	8	24	3	25	3
Incomplete Vaccinations	9,904	4	40	4	40	4	40	-	-	-	-	-	-
Second Grade Not Inoculated	123,605	65	53	53	43	42	34	11	9	6	5	6	5

1-16/55 = .71 Vaccine efficacy

After the 1954 Field Trial

- Cutter incident of Salks inactivated polio vaccine (IPV)
  - One manufacturer didn't properly kill the virus
  - 260 cases were caused: 94 vaccinees, 126 family, 40 community
- Sabin's oral attenuated vaccine (OPV) worked well in Soviet Union
  - Licensed in US 1960
  - Widely used in US 1961-89, simpler & worked bett IPV but
  - Causes paralysis in 1 of 2.9 million vaccinations
- By 2000 US had switched from OPV to IPV

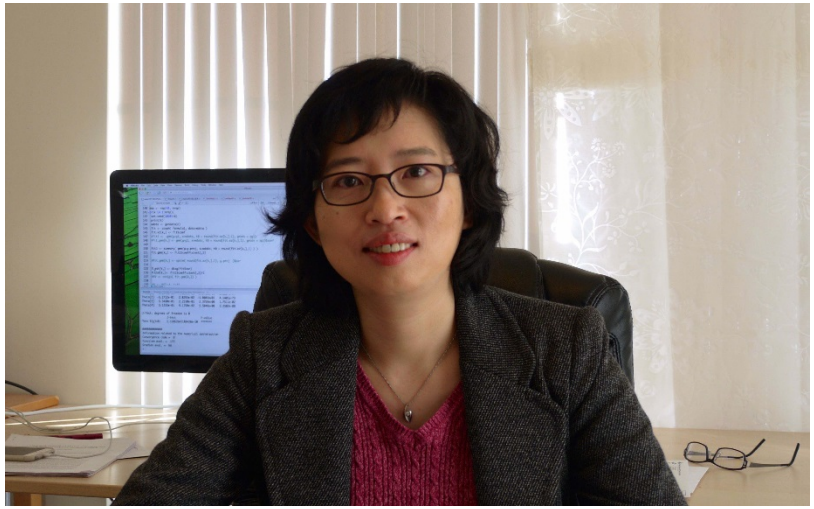
Global Polio Eradication

- Campaign started in 1988, WHO UNICEF & Rotary Foundation, now supported by BMGF & Hutch.
- Afghanistan & Pakistan two remaining countries with endemic polio
  - Challenge: vaccination is a western plot to sterilize
  - Challenge: sham Hep B vaccination campaign used to confirm Osama bin Laden's identity
- Oral polio vaccine (OPV) is highly effective but causes some polio making eradication difficult.
- Plan is to switch from OPV to killed (inactivated) IPV with last wild polio case

Acknowledgements

- Betz Halloran, Peter Gilbert, Michael Sachs, Erin Gabriel
- Chiung-Yu Huang

  (JHU)



INFECTION ENDPOINT WITH TIME CONSTANT VE_{If}
ONE DATA SET FOR EACH FEATURE

$$P(X_s > 0|Z = 1)/P(X_s > 0|Z = 0) = \exp(\beta_0 + \beta_1(s-1))$$

```
coxph(Surv(T,D) ~ W1+W2 + strata(Feature))
```

ID	T	Δ	Z	X_1	X_2	X_3
1	1.1	1	0	2	0	1
2	2.4	1	1	0	1	1
3	3.8	1	0	1	0	0
4	5.0	0	1	0	0	0

ID	T	D	Z	Feature	W_1	W_2
1	1.1	1	0	1	0	0
2	2.4	0	1	1	0	0
3	3.8	1	0	1	0	0
4	5.0	0	1	1	0	0
1	1.1	0	0	2	0	0
2	2.4	1	1	2	1	1
3	3.8	0	0	2	0	0
4	5.0	0	1	2	0	0
1	1.1	1	0	3	0	0
2	2.4	1	1	3	1	2
3	3.8	0	0	3	0	0
4	5.0	0	1	3	0	0

INFECTION ENDPOINT WITH ORCHESTRATED WANING VE_{If}
ONE DATA SET FOR EACH FEATURE

$$P(X_s > 0|Z = 1)/P(X_s > 0|Z = 0) = \exp(\beta_0 + \beta_1(s-1) + \beta_3 t + \beta_4 t(s-1))$$

```
coxph(Surv(T,D) ~ W1+W2 + W3 + W4  + strata(Feature))
```

ID	T	Δ	Z	X_1	X_2	X_3
1	1.1	1	0	2	0	1
2	2.4	1	1	0	1	1
3	3.8	1	0	1	0	0
4	5.0	0	1	0	0	0

ID	T	D	Z	Feature	W_1	W_2	W_3	W_4
1	1.1	1	0	1	0	0	0	0
2	2.4	0	1	1	0	0	0	0
3	3.8	1	0	1	0	0	0	0
4	5.0	0	1	1	0	0	0	0
1	1.1	0	0	2	0	0	0	0
2	2.4	1	1	2	1	1	2.4	2.4
3	3.8	0	0	2	0	0	0	0
4	5.0	0	1	2	0	0	0	0
1	1.1	1	0	3	0	0	0	0
2	2.4	1	1	3	1	2	2.4	4.8
3	3.8	0	0	3	0	0	0	0
4	5.0	0	1	3	0	0	0	0