Statistical Methods for Infectious Diseases
Household Based Studies I
Lecture 7C

M. Elizabeth Halloran

Fred Hutchinson Cancer Research Center and
University of Washington
Seattle, WA, USA

January 27, 2009

VE from SAR: Index case identified

- The secondary attack rate (SAR), the proportion of susceptibles exposed to an infectious person who become infected, has been used to estimate protective effects of vaccination since the 1930's (Kendrick and Eldering 1939).

$$VE_{S,SAR} = 1 - \frac{SAR_1}{SAR_0}$$

- where 0,1 denote the unvaccinated and vaccinated susceptibles exposed to an infective within a transmission unit, such as household.

Time-of-onset data

- Collection of transmission units, such as households.
- Time-of-onset of infection or disease for each susceptible in the household
- Relevant covariates, such as vaccine status or age
- Choice of analyses

Assumptions of the conventional SAR approach

- The transmission units are independent.
- The incubation and latent periods are fixed.
- The infectious period is fixed.
- The co-primaries are irrelevant.
- Asymmetric assumption that the index case and co-primaries get infected from outside the unit, while the susceptibles are exposed only within the unit.

Case definition

- Definition of case can influence analysis
- In pertussis analysis, had 5 different clinical and 8 different biologic criteria, for 40 different case definitions (Préziosi and Halloran 2003)

Setting up SAR Analysis

- Choose transmission unit
- First case in transmission unit called the index case or primary case.
- A potentially infectious contact, or exposure is a susceptible living in the same transmission unit during the infectious period of the index case.
- Individuals in the transmission unit can be considered a *minicohort*.

Setting up SAR Analysis, con't

- Need to make an assumption about the relation of the latent period to incubation period, if only disease observed.
- Often assume that symptom onset is onset of infectiousness
- Co-primaries are those cases with onset of symptoms too soon after the index case to have been infected by the index case.
- Generally, co-primaries are simply thrown out of the analysis in simple SAR analyses, entering neither as susceptibles in the denominator or infectives.

Setting up SAR Analysis, con't

- Choose the time interval in which an exposed susceptible can be considered an secondary case
- Minimum incubation period between index case and possible secondary case
- Minimum and maximum duration of infectiousness of the index case
- Co-primaries, secondary cases, and others presumably not due to index case then determined.

VE based on nonparametric secondary attack rates (SAR)

- The three main unstratified vaccine effects are

$$\begin{aligned} VE_{S.1/.0} &= 1 - \frac{SAR_{.1}}{SAR_{.0}}, \\ VE_{I1./0.} &= 1 - \frac{SAR_{1.}}{SAR_{0.}}, \\ VE_{T} &= 1 - \frac{SAR_{11}}{SAR_{00}}. \end{aligned}$$

- The stratified measures of VE_S and VE_I are

$$\begin{aligned} VE_{S01/00} &= 1 - \frac{SAR_{01}}{SAR_{00}}, \quad VE_{S11/10} = 1 - \frac{SAR_{11}}{SAR_{10}}, \\ VE_{I10/00} &= 1 - \frac{SAR_{10}}{SAR_{00}}, \quad VE_{I11/01} = 1 - \frac{SAR_{11}}{SAR_{01}}. \end{aligned}$$

Estimating Vaccine Efficacy from Secondary Attack Rates

- Preźiosi and Halloran (2003), Halloran, Préziosi, and Chu (2003)
- In this analysis, use SARs to estimate VE_I and VE_T for pertussis vaccine.
- The pertussis study in rural Niakhar, Senegal (Simondon, et al. 1997; Préziosi, et al. 2002)
- This analysis focuses on the calendar year 1993, an epidemic year that produced a large number of cases and extensive exposure to pertussis.

Setting up secondary attack rate analysis

- Choose unit of transmission: compound with extended family
- Index case or primary: first case in transmission unit
- Potentially infectious contact: someone living in same compound during infectious period of index case: exposed susceptible children with no history of pertussis.
- Onset of pertussis symptoms assumed onset of infectiousness.

Setting up secondary attack rate analysis

- Incubation period assumed at least 7 days long
- Co-primaries: symptoms within 7 days of index case; compounds with co-primaries excluded from this analysis because of interest in VE_I.
- Uncertainty in duration of infectiousness: varied cutoff from 28, 42, 56 and no cutoff after onset of index case.

Data

- 518 of 1,800 compounds were detected as having potential pertussis cases in 1993.
- Pertussis confirmed in 189 of those compounds
- Some more exclusionary criteria, partial vaccination, no susceptibles, households with co-primaries.
- 109 compounds with 109 primary cases and 790 susceptibles, 638 unvaccinated or completely vaccinated and 152 partially vaccinated.
- Biological confirmation available in 97% of suspected cases meeting clinical definition.

Data structure

- Let n be the number of compounds with a unique index case
- Let m_i be the number of susceptibles in the ith compound
- Let $\boldsymbol{x}_{ij} = (x_{ij1}, \cdots, x_{ijp})'$ denote a $p \times 1$ vector of explanatory variables associated with y_{ij}.
- In particular, let $x_{i.1}$ denote the vaccine status of the index case in compound i, and x_{ij2} the vaccine status of the jth exposed susceptible individual in compound i.

Correlated data structure

- Generally, confidence interval for VE based on SAR is simply based on log relative risk
- Does not take the correlated data structure into account within households
- Maybe not important in small households, but in pertussis study, compounds were large
- Considered marginal model (generalized estimating equations (PROC GENMOD))
- and random effects model (Bayesian (WinBUGS) and nonlinear mixed model (PROC NLMIXED))

Marginal Models

- In marginal models, inference about population averages is the focus.
- If there is heterogeneity across compounds in the baseline transmission, the estimated baseline coefficients represent an average over the heterogeneities.
- The correlation structure is some function of the marginal mean and possibly additional parameters.

Random Effects Models

- In the random effects model, a slightly different baseline transmission is estimated for each compound, with the degree of heterogeneity estimated in the variance of the random effect.
- The vaccine effects in each compound are interpreted in relation to that compound's baseline transmission.
- In this application, our primary scientific question is about the population average, or marginal, vaccine efficacy measures.
- So the marginal model is our model of choice.

The Marginal Model

- The marginal model for the logit of the SAR_{ij} of the jth person in the ith household is

$$\text{logit}(SAR_{ij}) \quad = \quad \beta_0 + \beta_1 x_{i\cdot 1} + \beta_2 x_{ij2} \ , \qquad (1)$$

- where $x_{i\cdot 1}$ denotes the vaccine status of the index case in compound i and x_{ij2} is the vaccine status of the jth exposed susceptible in compound i.
- The vaccine status of the index case, $x_{i\cdot 1}$, enters the analysis as a compound-level, environmental variable.

Transformation to SAR and VE scale

- Since we are interested in VE estimates on the SAR scale, we transform the parameters from the logistic model to the probability scale. The stratified SARs from model (1) are

$$\begin{aligned} SAR_{00} &= \frac{\exp\beta_0}{1+\exp\beta_0}, \quad SAR_{03} = \frac{\exp(\beta_0+\beta_2)}{1+\exp(\beta_0+\beta_2)}, \\ SAR_{30} &= \frac{\exp(\beta_0+\beta_1)}{1+\exp(\beta_0+\beta_1)}, \quad SAR_{33} = \frac{\exp(\beta_0+\beta_1+\beta_2)}{1+\exp(\beta_0+\beta_1+\beta_2)}. \end{aligned} \tag{2}$$

- Parameter estimates from the above model provide estimates for the stratified $VE_{S00/03}$ and $VE_{S30/33}$, the stratified $VE_{I00/30}$ and $VE_{I03/33}$, as well as VE_T.

Transformation to SAR and VE scale, cont'd

- Plugging the expressions for the SARs into equations the VE equations based on the SAR's, the expressions for the VE measures are

$$VE_{S03/00} = \frac{1-\exp(\beta_2)}{1+\exp(\beta_0+\beta_2)}, \quad VE_{S33/30} = \frac{1-\exp(\beta_2)}{1+\exp(\beta_0+\beta_1+\beta_2)},$$

$$VE_{I30/00} = \frac{1-\exp(\beta_1)}{1+\exp(\beta_0+\beta_1)}, \quad VE_{I33/03} = \frac{1-\exp(\beta_1)}{1+\exp(\beta_0+\beta_1+\beta_2)},$$

$$VE_T = \frac{1-\exp(\beta_1+\beta_2)}{1+\exp(\beta_0+\beta_1+\beta_2)}.$$

Transformation to SAR and VE scale, cont'd

- To obtain estimates of the unstratified $VE_{I3./0.}$ and $VE_{S.3/.0}$, we fit additional submodels, such as $\text{logit}(SAR_{ij}) = \beta_0' + \beta_1' x_{i\cdot 1}$ and $\text{logit}(SAR_{ij}) = \beta_0'' + \beta_2'' x_{ij2}$ and transform back to get

$$VE_{I3./0.} = \frac{1 - \exp(\beta_1')}{1 + \exp(\beta_0' + \beta_1')}, \quad VE_{S.3/.0} = \frac{1 - \exp(\beta_2'')}{1 + \exp(\beta_0'' + \beta_2'')}. \quad (3)$$

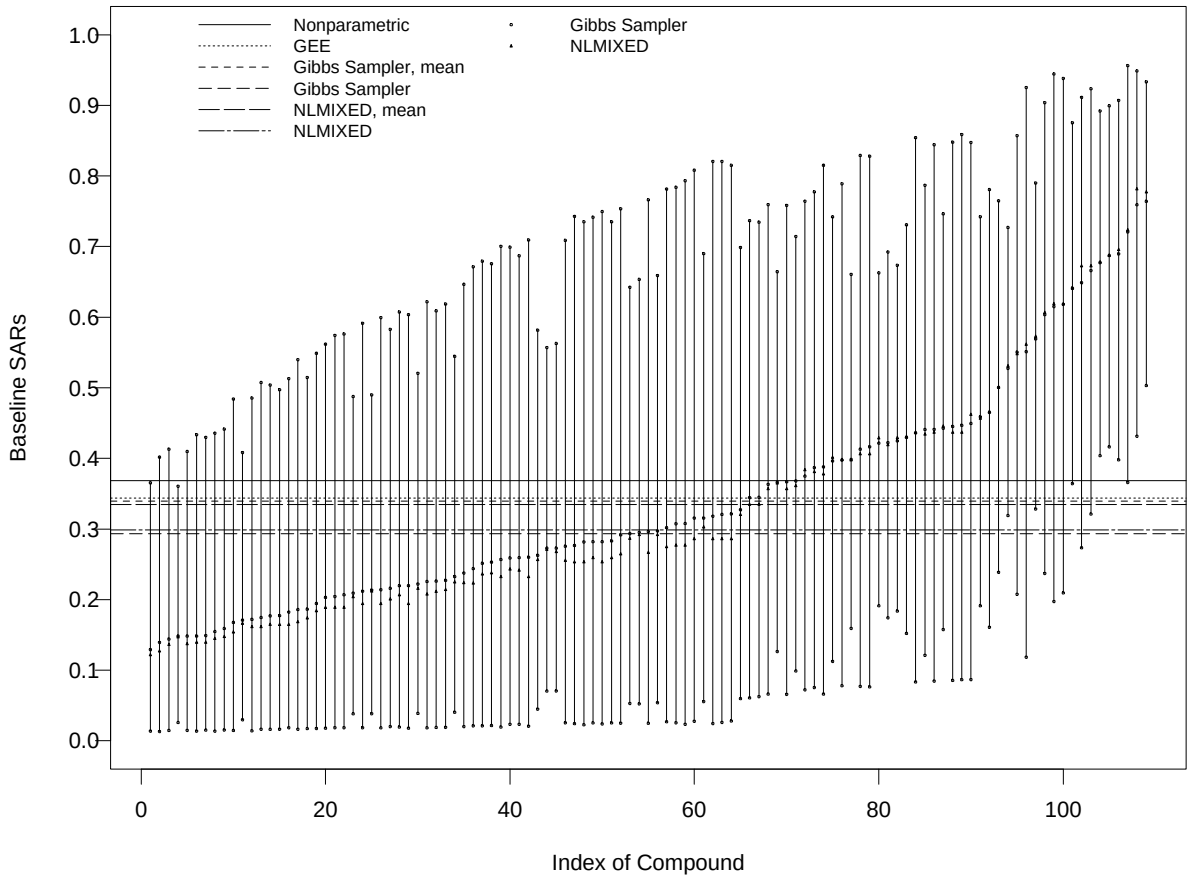
- Alternatively, we could have used the parameter estimates from the full model (1) and substitute the respective means of $x_{i\cdot 1}$ and x_{ij2}.

Random effects model

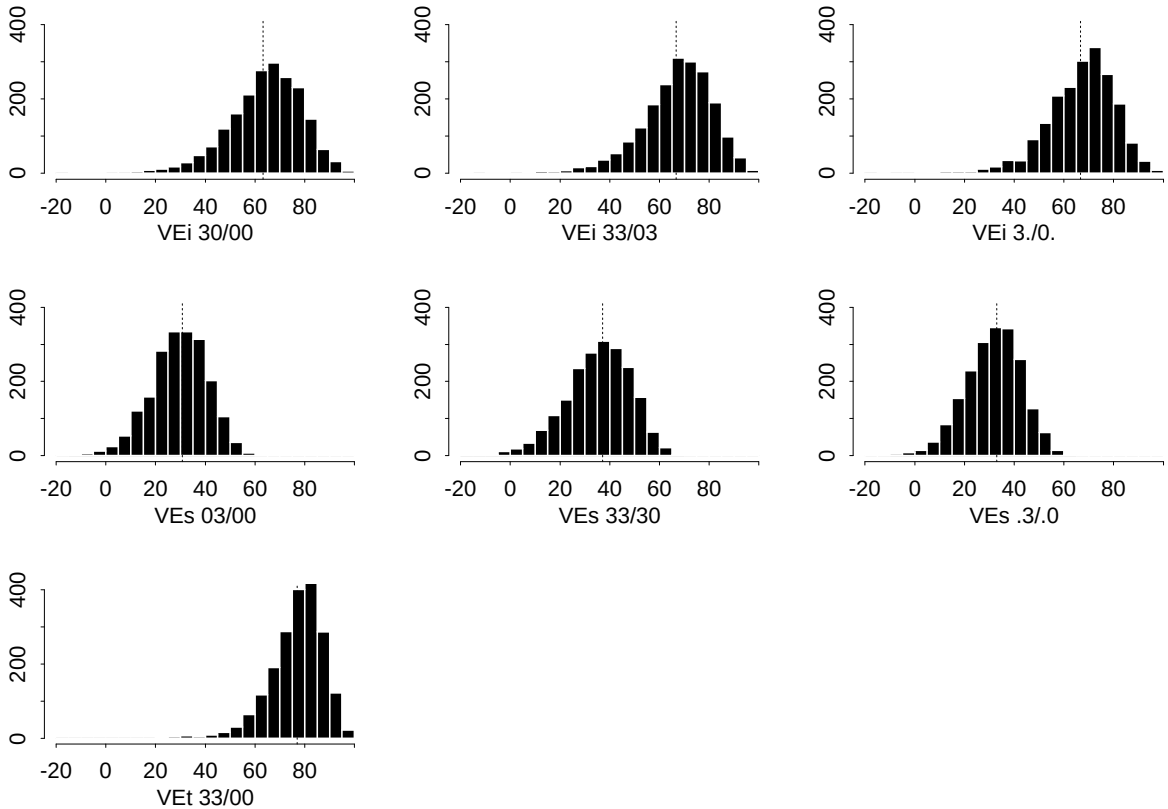
- Random effects model needs to compute the baseline SAR_{00} by integrating over the estimated random effects

Inference

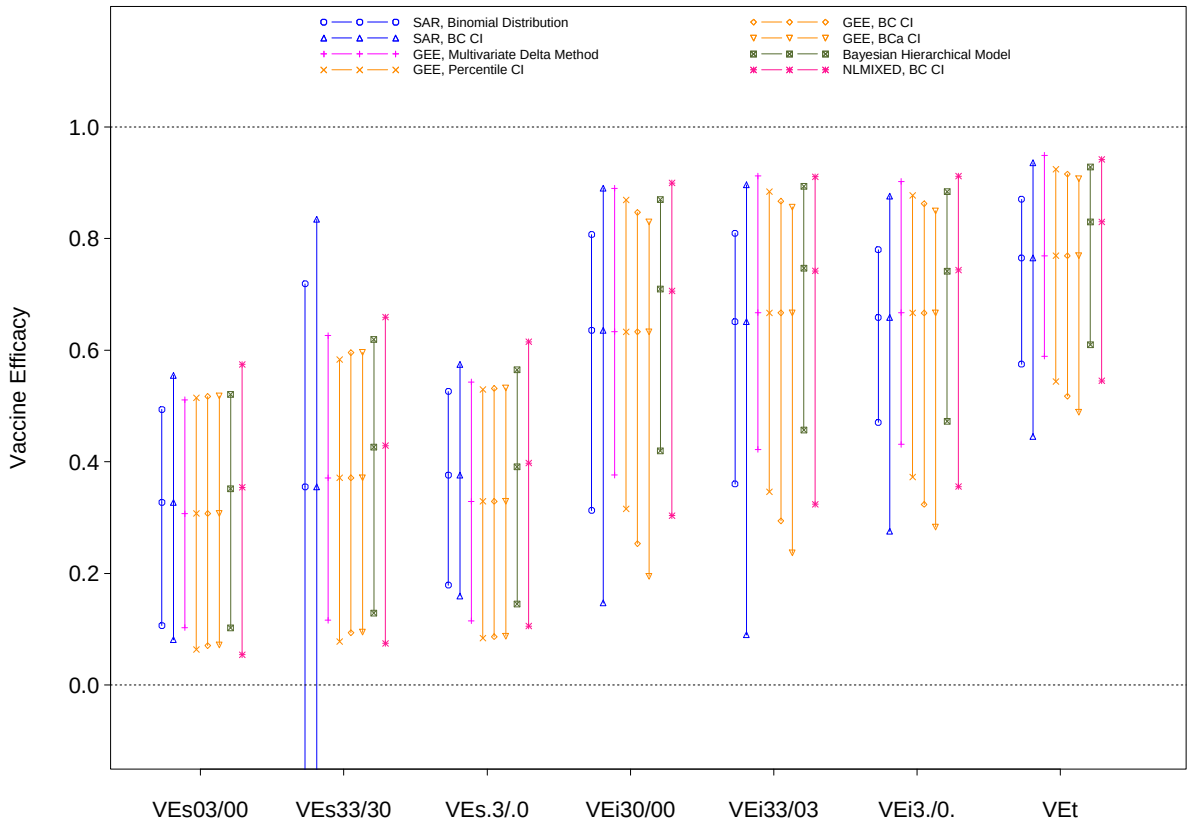
- Confidence intervals on the transformed SAR scale were obtained using the bootstrap
- Sampled by compound
- Estimated coefficients for each bootstrap sample then transformed back to SAR scale.



Halloran, Préziosi and Chu (2003)



Halloran, Préziosi and Chu (2003)



Halloran, Préziosi and Chu (2003)

Pertussis VE, Niakhar region, Senegal, 1993.

	Vaccine Efficacy (VE) x 100% (95% confidence interval)			
	VE for susceptibility			
Estimator	$VE_{S03/00}$	$VE_{S33/30}$	$VE_{S.3/.0}$	
SAR (BC*)	33 (8,55)	36 (−62,88)	38 (16,57)	
GEE (BC)	31 (7,52)	37 (9,60)	33 (9,53)	
	VE for infectiousness			Total VE
	$VE_{I30/00}$	$VE_{I33/03}$	$VE_{I3./0.}$	VE_T
SAR (BC*)	64 (15,89)	65 (9,90)	66 (28,88)	77 (45,94)
GEE (BC)	63 (25,85)	67 (29,87)	67 (32,86)	77 (52,92)

* BC = bias-corrected bootstrap confidence interval

Source: Préziosi and Halloran (2003); Halloran, Préziosi and Chu (2003)

Thank you!