Module 10: Evaluating Immune Correlates of Protection

Instructors: Peter Gilbert, Paul T. Edlefsen, Ying Huang

Supplemental: The Sieve Conditions

Summer Institute in Statistics and Modeling in Infectious Diseases

University of Washington, Department of Biostatistics

July 18-20, 2012

Statistical Center for HIV/AIDS Research & Prevention
SCHARP

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

The Sieve Conditions

- *per-contact* RR **is** *retrospective* RR **if**

  (during the trial follow-up period)

  1. Infection is possible from at most one strain
  2. The relative prevalence of strains is constant
  3. Exposure distributions are the same in both treatment groups, and homogeneous across subjects*

- Proof in Gilbert, Self, Ashby (1998)
  - Holds for all of the aforementioned models
  - * the homogeneity aspect of this assumption can be relaxed. See Gilbert, Statistics in Medicine 2000.
  - See Gilbert, et al (2001) for more discussion
- Allows for the interpretation of strain-specific VE as prospective, per-contact-by-*s* VE

ASSUMPTIONS and A1-A3

- ASSUMPTIONS
  1. Infection is possible from at most one strain
  2. (A3) The relative prevalence of strains is constant
  3. (A2) Exposure distributions are the same in both treatment groups, and homogeneous across subjects

A1: For each strain $s \in \{1, \cdots, K\}$,
the probability of infection with strain s
resulting from a specified amount of exposure
is homogeneous and constant over time among
vaccinated and placebo subjects, so that
vaccination reduces the transmission probability
by the same fraction $\exp\{\gamma_s\}$
for all vaccinees (i.e., "leaky" protection against
each strain; Halloran, Haber, and Longini, 1992)

Retrospective vs prospective

- Retrospective vs Prospective Category Probabilities

$$P^r_{Vs} \equiv \Pr\left(\text{ infected by strain } s \,\middle|\, \text{infected in } [0,\tau],\ \text{vaccine treatment assignment is } V \right)$$
$$\pi_V \equiv \Pr\left(\text{ infected in } [0,\tau] \,\middle|\, V \right)$$
$$P^p_{Vs} \equiv \Pr\left(\text{ infected by strain } s \text{ in } [0,\tau] \,\middle|\, V \right)$$
$$= \pi_V \times P^r_{Vs}$$

- Retrospective vs Prospective Relative Risks

$$\mathrm{RR}_r(s) \equiv \frac{\Pr\left(\text{ infected by strain } s \,\middle|\, \text{infected in } [0,\tau],\ \text{vaccine recipient} \right)}{\Pr\left(\text{ infected by strain } s \,\middle|\, \text{infected in } [0,\tau],\ \text{placebo recipient} \right)} = \frac{P^r_{vs}}{P^r_{us}}$$

$$\mathrm{RR}_\pi \equiv \frac{\Pr\left(\text{ infected in } [0,\tau] \,\middle|\, \text{vaccine recipient} \right)}{\Pr\left(\text{ infected in } [0,\tau] \,\middle|\, \text{placebo recipient} \right)} = \frac{\pi_v}{\pi_u}$$

$$\mathrm{RR}_p(s) \equiv \frac{\Pr\left(\text{ infected by strain } s \text{ in } [0,\tau] \,\middle|\, \text{vaccine recipient} \right)}{\Pr\left(\text{ infected by strain } s \text{ in } [0,\tau] \,\middle|\, \text{placebo recipient} \right)} = \frac{P^p_{vs}}{P^p_{us}}$$
$$= \frac{\pi_v P^r_{vs}}{\pi_u P^r_{us}}$$
$$= \mathrm{RR}_\pi \times \mathrm{RR}_r(s)$$

Per-contact probabilities and RRs

By assumption A1, the per-contact infection probability

$$\Pr\left(\text{ infected by strain } s \text{ at time } t \;\middle|\; \text{one exposure to strain } s \text{ at time } t,\ \text{vaccine treatment assignment is } V\right)$$

is independent of time t. Thus we may define

$$P_{Vs}^{pc} \equiv \Pr\left(\text{ infected by strain } s \;\middle|\; \text{one exposure to strain } s,\ \text{vaccine treatment assignment is } V\right), \text{ and}$$

$$\mathrm{RR}^{pc}(s) \equiv \frac{\Pr\left(\text{ infected by strain } s \;\middle|\; \text{one exposure to strain } s,\ \text{vaccine recipient}\right)}{\Pr\left(\text{ infected by strain } s \;\middle|\; \text{one exposure to strain } s,\ \text{placebo recipient}\right)} = \frac{P_{vs}^{pc}}{P_{us}^{pc}}.$$

Hazards and the exposure process

Define the hazard function of infection by strain s at time $t \in [0, \tau]$ for an individual with vaccination status V as

$$\lambda(t, s|V) \equiv \lim_{\triangle t \searrow 0} \frac{\Pr\big(t \leq T < t + \triangle t, Y = s \;\big|\; T \geq t, V\big)}{\triangle t},$$

and note that

$$\begin{aligned}
&\Pr\big(t \leq T < t + \triangle t, Y = s \;\big|\; T \geq t, V, \text{ subject-specific covariates }\big) = \\
&\quad \Pr\big(\text{ exposed to strain } s \text{ in } [t, t + \triangle t) \;\big|\; V, \text{ subject-specific covariates }\big) \\
&\quad \times \Pr\big(t \leq T < t + \triangle t, Y = s \;\big|\; T \geq t, V, \text{ exposed to } s \text{ in } [t, t + \triangle t), \text{ subject-specific covariates }\big).
\end{aligned}$$

If you assume ASSUMPTION 3 (A2): that the exposure distributions depend only on time and are independent of all subject-specific covariates, including treatment assignment, then the first term on the right-hand-side is the density of a Markov process in t, indexed by s but not by V. Then dividing both sides by $\triangle t$ and taking the limit as $\triangle t \searrow 0$ gives

$$\lambda(t, s|V) = f^N_{Es}(t) \times P^{pc}_{Vs},$$

where N_{Es} is the counting process counting exposures to strain s, with Markov intensity f^N_{Es} defined by

$$f^N_{Es}(t) \equiv \lim_{\triangle t \searrow 0} \frac{\Pr\big(\text{exposed to strain } s \text{ in } [t, t + \triangle t)\big)}{\triangle t},$$

the integral of which will be denoted by $F^N_{Es}(\tau) \equiv \int_0^\tau f^N_{Es}(t)dt$.

Unbiased odds ratios (pt 1 of 3)

Using the equation $Pr(T \geq t|V) = e^{-\Lambda(t|V)}$ relating a survivor function to a cumulative hazard function, it follows that

$$Pr(T \geq t|V) = \exp\left(-\int_0^t \lambda(u|V)du\right) = \exp\left(-\int_0^t \sum_l \lambda(u,l|V)du\right) = \exp\left(-\int_0^t \sum_l f_{El}^N(u)P_{Vl}^{pc}du\right).$$

$$\begin{aligned} P_{Vs}^p &= \int_0^\tau \lim_{\triangle t \searrow 0} \frac{\Pr(t \leq T < t+\triangle t, Y = s|T \geq t, V)}{\triangle t} dt = \int_0^\tau \lambda(t,s|V) \times \Pr(T \geq t|V)dt \\ &= \int_0^\tau f_{Es}^N(t)P_{Vs}^{pc}\Pr(T \geq t|V)dt \\ &= \int_0^\tau f_{Es}^N(t)P_{Vs}^{pc} \exp\left(-\int_0^t \sum_l f_{El}^N(u)P_{Vl}^{pc}du\right)dt. \end{aligned}$$

Unbiased odds ratios (pt 2 of 3)

Now if you assume ASSUMPTION 2 (A3): that the strain-specific exposure intensities are proportional, i.e.,

$$f^N_{Es}(t) = \theta_s f^N_{E0}(t),$$

then

$$\begin{aligned}
P^p_{Vs} &= \int_0^\tau f^N_{Es}(t) P^{pc}_{Vs} \exp\Big(-\int_0^t \sum_l f^N_{El}(u) P^{pc}_{Vl}\, du\Big) dt \\
&= \theta_s P^{pc}_{Vs} \int_0^\tau f^N_{E0}(t) \exp\Big(-\int_0^t f^N_{E0}(u) du \sum_l \theta_l P^{pc}_{Vl}\Big) dt \\
&= \theta_s P^{pc}_{Vs} \int_0^\tau f^N_{E0}(t) \exp\Big(-F^N_{E0}(t) \sum_l \theta_l P^{pc}_{Vl}\Big) dt \\
&= \left| \frac{-1}{\sum_l \theta_l P^{pc}_{Vl}} \exp\Big(-u \sum_l \theta_l P^{pc}_{Vl}\Big) \right|_0^{F^N_{E0}(\tau)} \times \theta_s P^{pc}_{Vs} \\
&= \Big(1 - \exp\big(-F^N_{E0}(\tau) \sum_l \theta_l P^{pc}_{Vl}\big)\Big) \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}} \\
&= \big(1 - Pr(T \geq \tau | V)\big) \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}} \\
&= \pi_V \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}},
\end{aligned}$$

which, since $P^p_{Vs} = \pi_V \times P^r_{Vs}$, implies that

$$P^r_{Vs} = \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}}.$$

Unbiased odds ratios (pt 3 of 3)

$$P^r_{Vs} = \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}}.$$

Finally, this result guarantees equivalence of retrospective, prospective, and per-contact odds ratios, since

$$\begin{aligned}
\mathrm{OR}_r(s) &\equiv \frac{P^r_{vs}/P^r_{us}}{P^r_{v1}/P^r_{u1}} \\
&= \frac{P^r_{vs}/P^r_{v1}}{P^r_{us}/P^r_{u1}} \\
&= \frac{\theta_s P^{pc}_{vs}/\theta_1 P^{pc}_{v1}}{\theta_s P^{pc}_{us}/\theta_1 P^{pc}_{u1}} \\
&= \frac{P^{pc}_{vs}/P^{pc}_{v1}}{P^{pc}_{us}/P^{pc}_{u1}} \\
&= \mathrm{OR}^{pc}(s).
\end{aligned}$$

Biased relative risks

The result that $P^r_{Vs} = \frac{\theta_s P^{pc}_{Vs}}{\sum_l \theta_l P^{pc}_{Vl}}$ implies that the retrospective relative risk is biased as an estimator of the per-contact relative risk, since

$$\begin{aligned} \mathrm{RR}_r(s) &= \frac{P^r_{vs}}{P^r_{us}} \\ &= \frac{P^{pc}_{vs} / \sum_l \theta_l P^{pc}_{vl}}{P^{pc}_{us} / \sum_l \theta_l P^{pc}_{ul}} \\ &= \mathrm{RR}^{pc}(s) \frac{\sum_l \theta_l P^{pc}_{ul}}{\sum_l \theta_l P^{pc}_{vl}} \end{aligned}$$

Note that the extra term (the multiplicative bias) is not a function of the strain category, so it cancels when taking a ratio as in $\mathrm{OR}_r(s) = \mathrm{RR}_r(s)/\mathrm{RR}_r(1)$.

Estimation of strain-specific VE and differential VE

Strain-specific VE is defined as $\mathrm{VE}_s = 1 - \mathrm{RR}^{pc}(s)$.

Together, the α_s^{mlr} and β_s^{mlr} parameters of the MLR model satisfy $\alpha_s^{mlr} + \beta_s^{mlr} = \log\left(\mathrm{RR}^{r}(s)\right)$. $\mathrm{RR}^{r}(s)$ is a biased estimate of $\mathrm{RR}^{pc}(s)$.

Differential VE (between strains s and l) is defined as when $\mathrm{VE}_s \neq \mathrm{VE}_l$, and can be tested by comparing the strain-specific per-contact relative risks $\mathrm{RR}^{pc}(s)$ and $\mathrm{RR}^{pc}(l)$, for instance by testing the hypothesis that $\mathrm{RR}^{pc}(s)/\mathrm{RR}^{pc}(l) = 1$.

The β_s^{mlr} parameters of the MLR model satisfy $\beta_s^{mlr} = \log\left(\mathrm{OR}^{r}(s)\right)$, and therefore (under A1-A3), $\beta_s^{mlr} = \log\left(\mathrm{OR}^{pc}(s)\right)$.

Failure time models

Note that the "crude" hazard ratio $\lambda(t,s|v)/\lambda(t,s|u)$ is an unbiased estimator of the per-contact relative risk, even without ASSUMPTION 2 (A3), since

$$\frac{\lambda(t,s|v)}{\lambda(t,s|u)} = \frac{f_{Es}^{N}(t) \times P_{vs}^{pc}}{f_{Es}^{N}(t) \times P_{us}^{pc}} = \mathrm{RR}^{pc}(s).$$

Strain-specific VE is defined as $\mathrm{VE}_s = 1 - \mathrm{RR}^{pc}(s)$.

The β_s^{ph} parameters of the proportional hazards model satisfy $\beta_s^{ph} = \log\big(\lambda(t,s|v)/\lambda(t,s|u)\big)$, and therefore (under A1-A2), $\beta_s^{ph} = \mathrm{RR}^{pc}(s)$.

Differential VE (between strains s and l) is defined as when $\mathrm{VE}_s \neq \mathrm{VE}_l$, and can be tested by comparing the strain-specific per-contact relative risks $\mathrm{RR}^{pc}(s)$ and $\mathrm{RR}^{pc}(l)$, for instance by testing the hypothesis that $\mathrm{RR}^{pc}(s)/\mathrm{RR}^{pc}(l) = 1$.

The β_s^{lm} parameters of the Lunn & McNeil (1995) recoded proportional hazards model satisfy $\beta_s^{lm} = \log\big(\frac{\lambda(t,s|v)}{\lambda(t,s|u)}/\frac{\lambda(t,1|v)}{\lambda(t,1|u)}\big)$, and therefore (under A1-A2), $\beta_s^{lm} = \mathrm{OR}^{pc}(s)$.

Equivalence of failure time and count models under A3

ASSUMPTION 2 (A3) has two equivalent mathematical expressions. From the binary-endpoint perspective, the assumption says that the exposure intensities are proportional:

$$f_{Es}^{N}(t) \equiv \theta_s \times f_{E0}^{N}(t),$$

while from the time-to-event perspective, the assumption says that the strain-specific baseline hazards are proportional:

$$\lambda(t, s \mid V = u) \equiv \theta_s \times \lambda(t, 1 \mid V = u).$$

The fact that ASSUMPTION 2 (A3) represents the extra assumption needed for the MLR model to possess the same biological interpretation for sieve analysis as the proportional hazards model is connected to a result of Prentice et al. (1978). They showed that under the cause-specific proportional hazards model with proportional baseline risks, given by

$$\lambda_s(t \mid z) = \exp\left(\alpha_s + \beta_s^T z\right)\lambda_1(t),$$

the marginal probability distribution of failure type satisfies the MLR model. Thus, when censoring is non-informative, the failure time and count models intersect under A3.

Kochar and Proschan (Statistica Sinica 1991) showed that the failure time and type are mutually independent if and only if this model holds.

Under assumption A3 and non-informative censoring, strain-specific proportional hazards models and the MLR model are equivalent tools for sieve analysis. On the other hand, if A3 is violated, the incorporation of failure times can potentially improve inference.