**Appendix**

Table 1 presents hypothetical data on infection after a first exposure. In these data 70% of people are low-risk (*U*=0); 10% of low-risk people are exposed to Chlamydia during the study, and have a 10% risk of infection upon exposure in the absence of vaccination. The other 30% of the population are sex workers (*U*=1); 60% of sex workers are exposed during the study, and also have a 10% risk of infection in the absence of vaccination. The per-exposure vaccine effects differ between strata of *U* as shown (e.g., the vaccine could be less effective among sex workers because of an increased prevalence of genital inflammation).

The observed intention-to-treat HR is a biased estimate of the counterfactual HR as a result of confounding for *Et*.

**Table 1:** Hypothetical data illustrating confounding of exposure to infection

|  |  |  |  |
| --- | --- | --- | --- |
| Counterfactual data | | | |
|  | =1  (Infected) | (Uninfected) | N |
| *A*=1 | 37 | 963 | 1000 |
| *A*=0 | 100 | 900 | 1000 |
| HRCtr=0.37 | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Observed data | | | |
|  | =1| Et =1, Ēt-1=  (Infected) | =0| Et =1, Ēt-1=  (Uninfected) | N |
| *A*=1 | 12 | 238 | 250 |
| *A*=0 | 25 | 225 | 250 |
| HRObs=0.48 | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Observed data, *U*=0 | | | |
|  | =1| Et =1, Ēt-1=  (Infected) | =0| Et =1, Ēt-1=  (Uninfected) | N |
| *A*=1 | 2 | 68 | 70 |
| *A*=0 | 7 | 63 | 70 |
| HRU=0=0.29 | | | |

|  |  |  |  |
| --- | --- | --- | --- |
| Observed data, *U*=1 | | | |
|  | =1| Et =1, Ēt-1=  (Infected) | =0| Et =1, Ēt-1=  (Uninfected) | N |
| *A*=1 | 10 | 170 | 180 |
| *A*=0 | 18 | 162 | 180 |
| HRU=1=0.56 | | | |

*A*: vaccine status (1= vaccinated, 0=unvaccinated)

: counterfactual Chlamydia infection status for a defined intervention and exposure status (1=infected, 0=uninfected)

*et*: counterfactual exposure status for Chlamydia at time *t* (1=exposed once, 0=unexposed)

*ēt*-1: counterfactual Chlamydia exposure history between time zero and *t*-1 (=unexposed)

: observed Chlamydia infection status at time *t*+1 (1=infected, 0=uninfected)

*Et*: observed exposure to Chlamydia at time *t* (1=exposed once, 0=unexposed)

*Ēt-1*: observed Chlamydia exposure history between time zero and *t*-1 (=unexposed)

*U*: sex worker status (1=sex worker, 0=general population)

N: number of participants

HR: discrete-time hazard ratio

Ctr: counterfactual

Obs: observed