$$OR = \frac{\pi_1/(1-\pi_1)}{\pi_2/(1-\pi_2)} \quad \begin{cases} \text{Odds of } y=1 \text{ for } x=1 \\ \\ \text{Odds of } y=1 \text{ for } x=2 \end{cases}$$

$$= \left(\frac{\pi_1}{\pi_2}\right)\left(\frac{1-\pi_2}{1-\pi_1}\right)$$

$$= RR \times \left(\frac{1-\pi_2}{1-\pi_1}\right)$$

When does $OR \approx RR$ $\longrightarrow$ $\pi_i \approx .5$

$\leadsto$ Both $\pi_1 \times \pi_2$ very small

If $\pi_1 \approx \pi_2$

Sampling options — $X \begin{array}{c} 1 \\ 2 \end{array} \left| \begin{array}{cc} 1 & 2 \\ \end{array} \right.$

① Both row totals × col. totals random
   (unknown prior to data collection)
   (e.g. multinomial or Poisson sampling)
   $\longrightarrow$ Cross-sectional or Cross-classification design.

   — Values of $X \to y$ unknown prior to data.

   $\Rightarrow$ Both RR and OR estimable.

② Row totals fixed by researchers

$X = 1 \longrightarrow n_1$
$X = 2 \longrightarrow n_2$

— Y-values unknown
→ clinical trials
    → randomly assign to x-value
— prospective

Estimable? $P(Y=1 \mid X=x)$

→ Both RR & OR

$P(X = x \cancel{\mid} Y=1)$

	Y	
	1	2
1	n_{11}	n_{12}
2	n_{21}	n_{22}

X

③ Col totals fixed by researchers

$Y = 1 \longrightarrow n_{+1}$
$Y = 2 \longrightarrow n_{+2}$

— X-values unknown
→ case-control
— retrospective

Estimable? $P(X=1 \mid Y=y)$

$P(Y=1 \cancel{\mid X=x})$

→ RR not estimable

But OR is!

Odds of $Y=1$ comparing $X=1$ to $X=2$:

$$OR_{Y=1} = \frac{Odds(Y=1 \mid X=1)}{Odds(Y=1 \mid X=2)}$$

$$= \frac{n_{11}/n_{12}}{n_{21}/n_{22}} = \boxed{\frac{n_{11} n_{22}}{n_{12} n_{21}}}$$

Compare to Odds of $X=1$ comparing $Y=1$ to $Y=2$

$$OR_{X=1} = \frac{Odds(X=1 \mid Y=1)}{Odds(X=1 \mid Y=2)} = \frac{n_{11}/n_{21}}{n_{12}/n_{22}} \quad \text{equal!}$$

$$= \boxed{\frac{n_{11} n_{22}}{n_{12} n_{21}}}$$

≠ Not true for RR.

Take-home message:

- RR usually preferred, but not estimable for case-control studies

- OR always estimable

Why case-control study?
- less expensive than longitudinal prospective study
- ethical issues
- rare diseases $* \longrightarrow \pi_1 \, \& \, \pi_2$ small

$$\Longrightarrow RR \approx OR$$

- OR more common to estimate in GLM

$\Longrightarrow$ Inference on OR $\longrightarrow$ interpret as if it's RR!