USING LARGE PUBLIC HEALTH DATABASES: VACCINE EFFECTIVENESS

Ron Brookmeyer

Fielding School of Public Health

University of California, Los Angeles

BACKGROUND

- REAL WORLD VACCINE EFFECTIVENSS STUDIES
- USE PUBLIC HEALTH DATABASES
  - DATABASE (REGISTRY) OF VACCINATIONS
  - DATABASE (REGISTRY) OF CASES (DISEASE)
- LINK INDIVIDUALS IN DATABASES
  - (e.g., names, date of birth, address)
- COVID-19 EXAMPLES
  - Scobie et al (2021)
  - Rosenberg et al (2021)

OBJECTIVE
EVALUATE SOME SOURCES OF ERROR

- ERRORS IN REPORTING CASES
- ERRORS IN REPORTING VACCINCATIONS
- ERRORS IN LINKING INDIVIDUALS ACROSS DATABASES

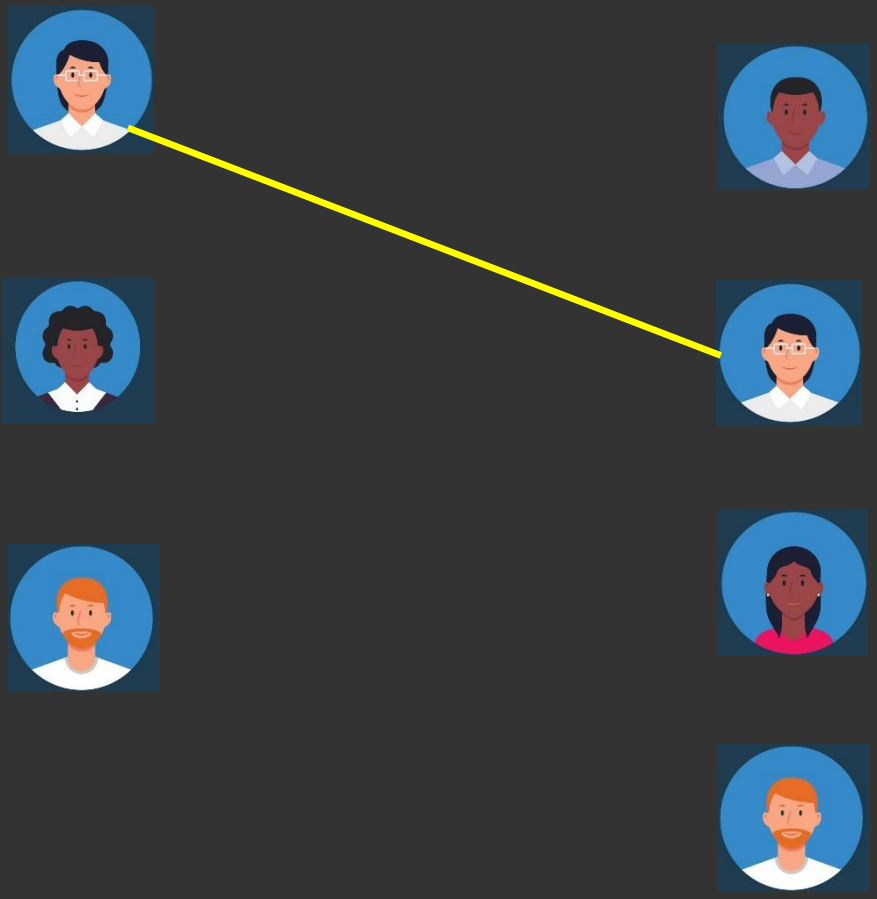
R. Brookmeyer & DE Morrison, Estimating Vaccine Effectiveness By Linking Population-based Health Registries: Some Sources of Bias, *American Journal Of Epidemiology*, 2022.

VACCINE REGISTRY

CASE REGISTRY

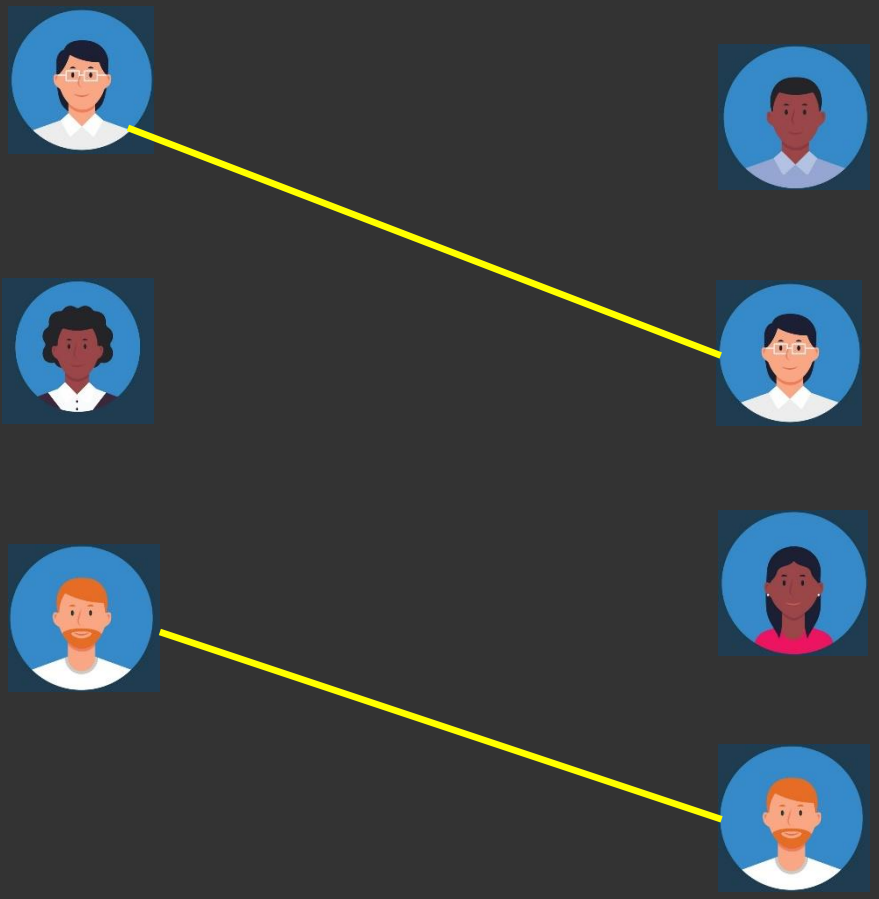
VACCINE REGISTRY

CASE REGISTRY



VACCINE REGISTRY

CASE REGISTRY



	Case	Non-Case	
Vaccinated	N_{VC}		N_V
Unvaccinated			
	N_C		N

	Case	Non-Case	
Vaccinated	N_{VC}		N_V
Unvaccinated			
	N_C		N

Population size of registry catchment area

	Case	Non-Case	
Vaccinated	N_{VC}	$N_{V\overline{C}} = N_V - N_{VC}$	N_V
Unvaccinated	$N_{\overline{V}C} = N_C - N_{VC}$	$N_{\overline{V}\overline{C}} = N - N_C - N_V + N_{VC}$	$N_{\overline{V}} = N - N_V$
	N_C	$N_{\overline{C}} = N - N_C$	N

	Case	Non-Case	
Vaccinated	N_{VC}	$N_{V\bar{C}} = N_V - N_{VC}$	N_V
Unvaccinated	$N_{\bar{V}C} = N_C - N_{VC}$	$N_{\bar{V}\bar{C}} = N - N_C - N_V + N_{VC}$	$N_{\bar{V}} = N - N_V$
	N_C	$N_{\bar{C}} = N - N_C$	N

$$\hat{R} = \frac{N_{VC} N_{\bar{V}}}{N_V N_{\bar{V}C}} \qquad \widehat{VE} = \left(1 - \hat{R}\right) \times 100\%$$

ERRORS EVALUATED

- UNDERREPORTING TO REGISTRIES
  NON-DIFFERENTIAL & INDEPENDENT:
  -reporting of C does not depend on V
  -reporting of V does not depend on C

- ERRORS IN LINKING
  -missed true matches

- ERRORS IN POPULATION SIZE (N)

METHODS

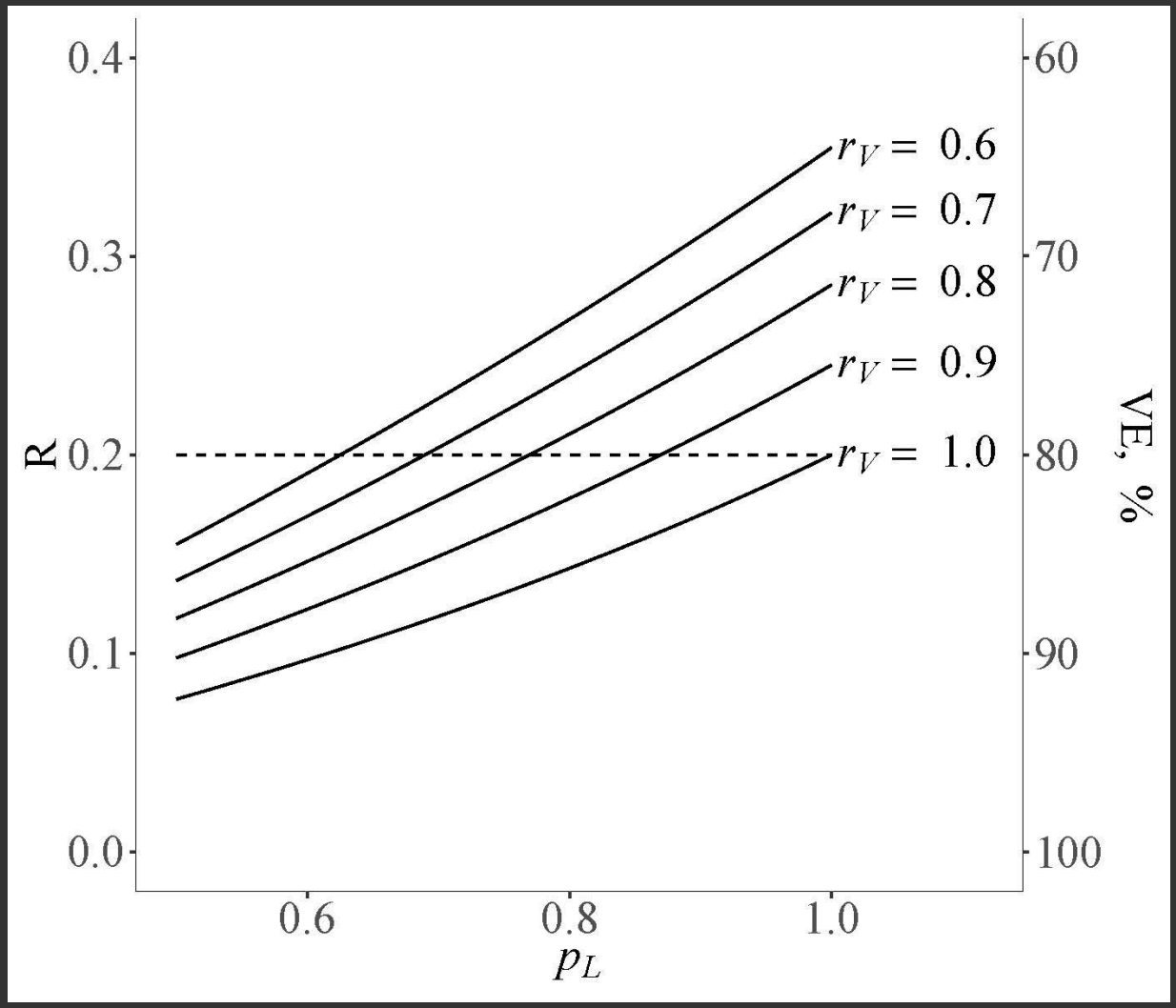
- ANALYSIS
- SIMULATION

RESULT WITH ONE SOURCE OF ERROR

ERROR	DIRECTION OF BIAS in VE
• UNDERREPORT VACCINATIONS	VE BIASED TOWARD 0
• UNDERREPORT CASES	NO BIAS IN VE
• LINKING ERROR: MISSED TRUE MATCHES	OVERESTIMATE VE
• ERROR IN *N*	
UNDERESTIMATE *N*	OVERESTIMATE VE
OVERESTIMATE *N*	UNDERESTIMATE VE

MULTIPLE ERRORS: LINK ERROR & VACCINE UNDERREPORT

True VE=80%



ADJUST VE TO CORRECT FOR ERRORS

- ADJUSTMENT FORMULA FOR VE
- STUDIES TO ESTIMATE MAGNITUDE OF ERRORS
- SENSITIVITY ANALYSES

SIMULATION RESULTS

True matches missed	Vaccine Under-report	Error in N	True VE	Unadjusted VE	Adjusted VE
10%	10%	0	0%	25%	0%
10%	10%	-10%	0%	48%	0%
10%	10%	-10%	80%	86%	80%
10%	10%	+10%	80%	72%	80%
10%	10%	-20%	80%	92%	80%

Average over 1000 simulations. $N=11x10^6$

CONCLUDING REMARKS

- **SUMMARY: SOURCES OF ERROR**
  - -One error: Direction of bias is predictable.
  - - Multiple errors: direction is either way.
    - Strong enough to make harmful vaccines appear effective
  - - Adjust VE for biases.
  - - Enables comparisons with clinical trials, countries

- **OTHER ERRORS**
  - -"differential" registry underreporting
  - - missed true matches & false matches

- **THE WAY FORWARD**
  - **-Improve public health databases**
    - - Improve completeness & accuracy of databases
    - -Improve data linking across databases
  - **-Studies to quantify errors to inform VE adjustment**