Lecture 6: Stochastic models for small groups such as households
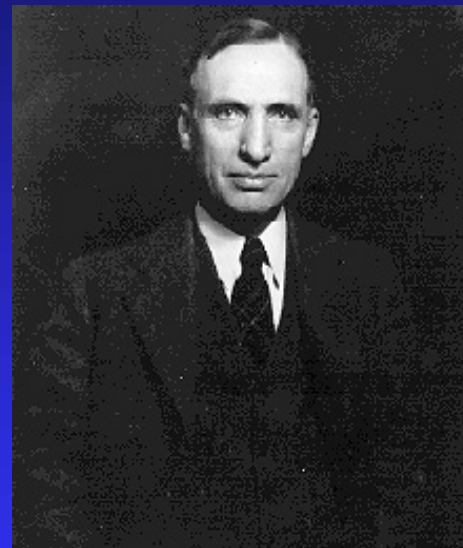
A Brief History of the Reed-Frost Model

- PD En'ko (1889) Deterministic difference equations
- L Reed and WH Frost (1930) Marbles and shoots
- M Greenwood (1931) Alternative formulation
- H Abbey (1952) 1st analysis as a stochastic process
- L Elveback, JP Fox, E Ackerman (1960) 1st computer program and lots of theory

Reed-Frost Model

Lowell Reed

1886 - 1966

Wade Hampton Frost

1880–1938

Both Former Deans: Johns Hopkins School of Public Health

Helen Abbey
1915 - 2001

Eugene Ackerman
1920 - 2014

John P. Fox
Died around 1989

Reed-Frost Model

Stochastic process: discrete state space and time t_0, t_1, t_2 ….

- **Infectious agent natural history**
  - ◆ Infectious for one time unit
- **Social contact structure**
  - ◆ Random mixing
  - ◆ *p = 1 – q,* probability two people make contact sufficient to transmit
- $R_0 = (n\text{-}1)p$

Reed-Frost Model

$$P(I_{t+1} | S_t, I_t) = \binom{S_t}{I_{t+1}} (1 - q^{I_t})^{I_{t+1}} q^{I_t (S_t - I_{t+1})}, S_t \geq I_{t+1},$$

$$S_{t+1} = S_t - I_{t+1},$$

$$R_{t+1} = R_t + I_t,$$

$$S_t + I_t + R_t = n, \forall t,$$

$$P[S(0) = n - 1] = 1, P[I(0) = 1] = 1, P[R(0) = 0] = 1$$

$$\{S_t, I_t\}_{t=0,1,\ldots} \; is \; a \; Markov \; chain$$

See chain binomial chapter in the *Encyclopedia Biostat., Vol 1, 593-7*

Greenwood Model

$$P(I_{t+1}|S_t, I_t) = \binom{S_t}{I_{t+1}} (1 - q^{I_t})^{I_{t+1}} q^{I_t(S_t - I_{t+1})}, S_t \geq I_{t+1},$$

$$S_{t+1} = S_t - I_{t+1},$$

$$R_{t+1} = R_t + I_t,$$

$$S_t + I_t + R_t = n, \forall t,$$

$$P[S(0) = n - 1] = 1, P[I(0) = 1] = 1, P[R(0) = 0] = 1$$

$$\{S_t, I_t\}_{t=0,1,\dots} \; is \; a \; Markov \; chain$$

A Chain

$$I_0 \to I_1 \to I_2 \to \cdots \to I_r$$

$$P(I_0, I_1, I_2, \cdots, I_r) =$$
$$P(I_1 \mid S_0, I_0) P(I_2 \mid S_1, I_1) \cdots$$
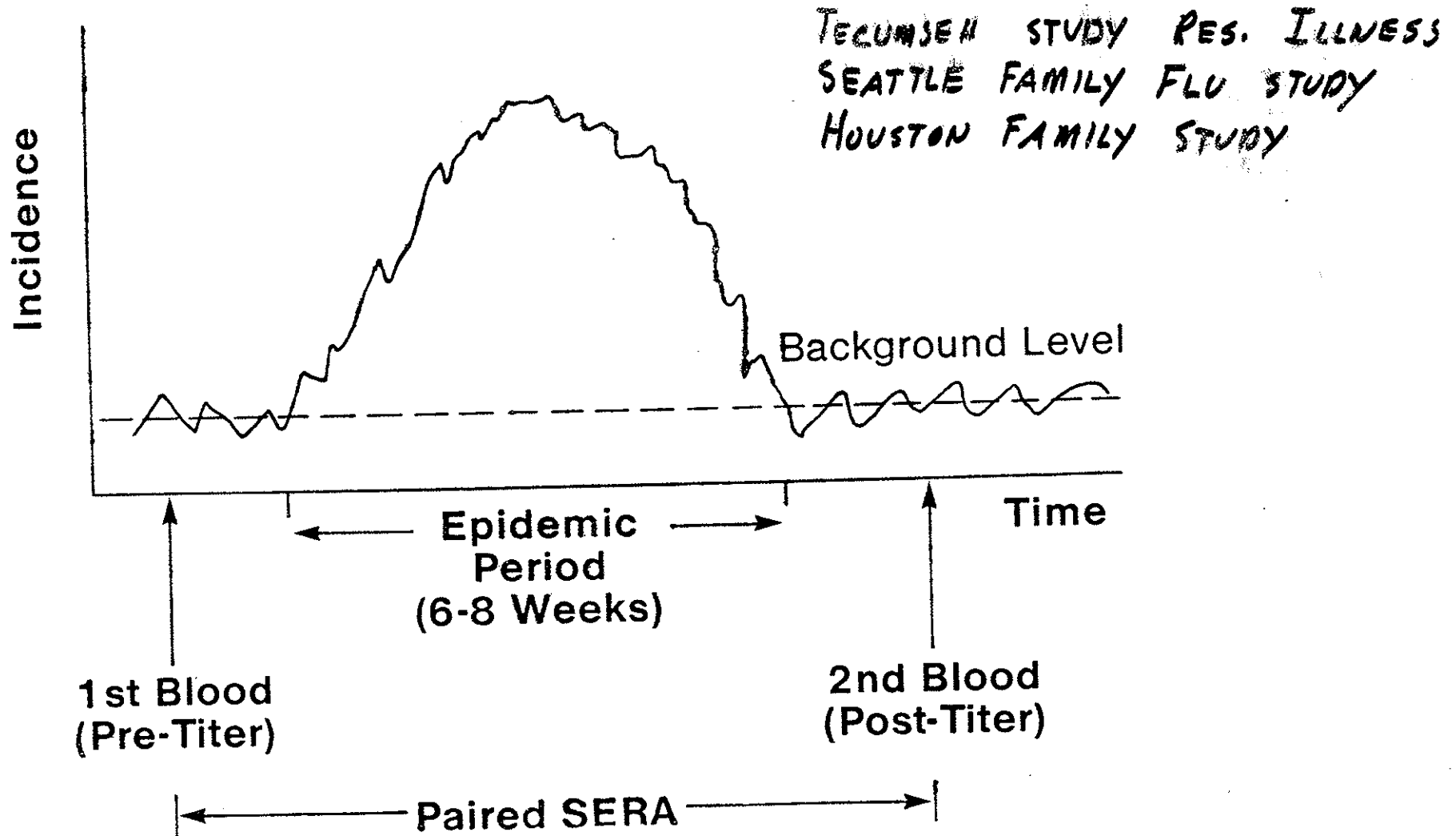$$P(I_r \mid S_{r-1}, I_{r-1})$$

$$T = \inf_{t \geq 0} \{t : S_t I_t = 0\}.$$

Example

Possible individual chains when $S_0 = 3,\ I_0 = 1$

Chain	Probability	Final Size
$\{i_0, i_1, i_2, ..., i_T\}$		R_T
$\{1\}$	q^3	1
$\{1, 1\}$	$3pq^4$	2
$\{1, 1, 1\}$	$6p^2q^4$	3
$\{1, 2\}$	$3p^2q^3$	3
$\{1, 1, 1, 1\}$	$6p^3q^3$	4
$\{1, 1, 2\}$	$3p^3q^2$	4
$\{1, 2, 1\}$	$3p^3q(1+q)$	4
$\{1, 3\}$	p^3	4

Epidemic in the Community

Tecumseh Study Res. Illness
Seattle Family Flu Study
Houston Family Study

Incidence

Background Level

Time

Epidemic ——→
Period
(6-8 Weeks)

1st Blood
(Pre-Titer)

2nd Blood
(Post-Titer)

←——————— Paired SERA ———————→

Community Structure

Community
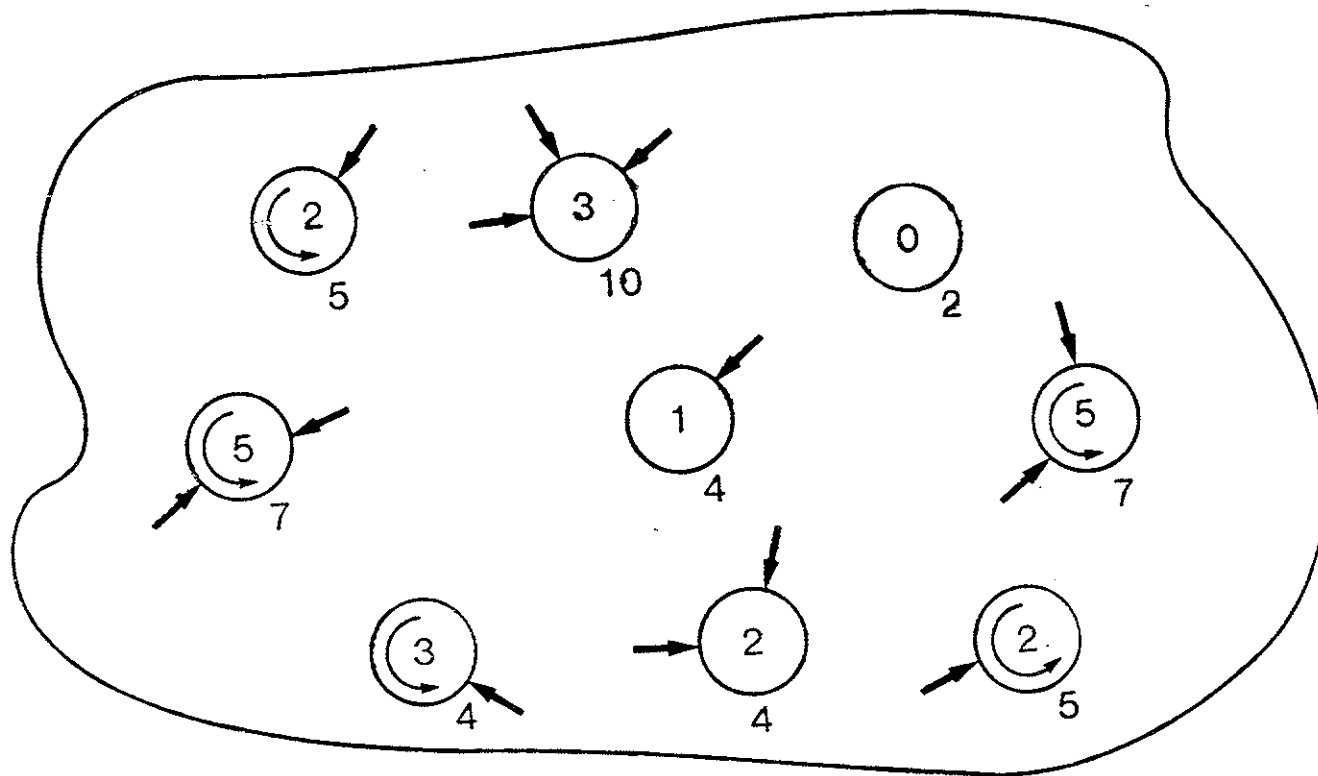
n – Households (Semiclosed Groups)
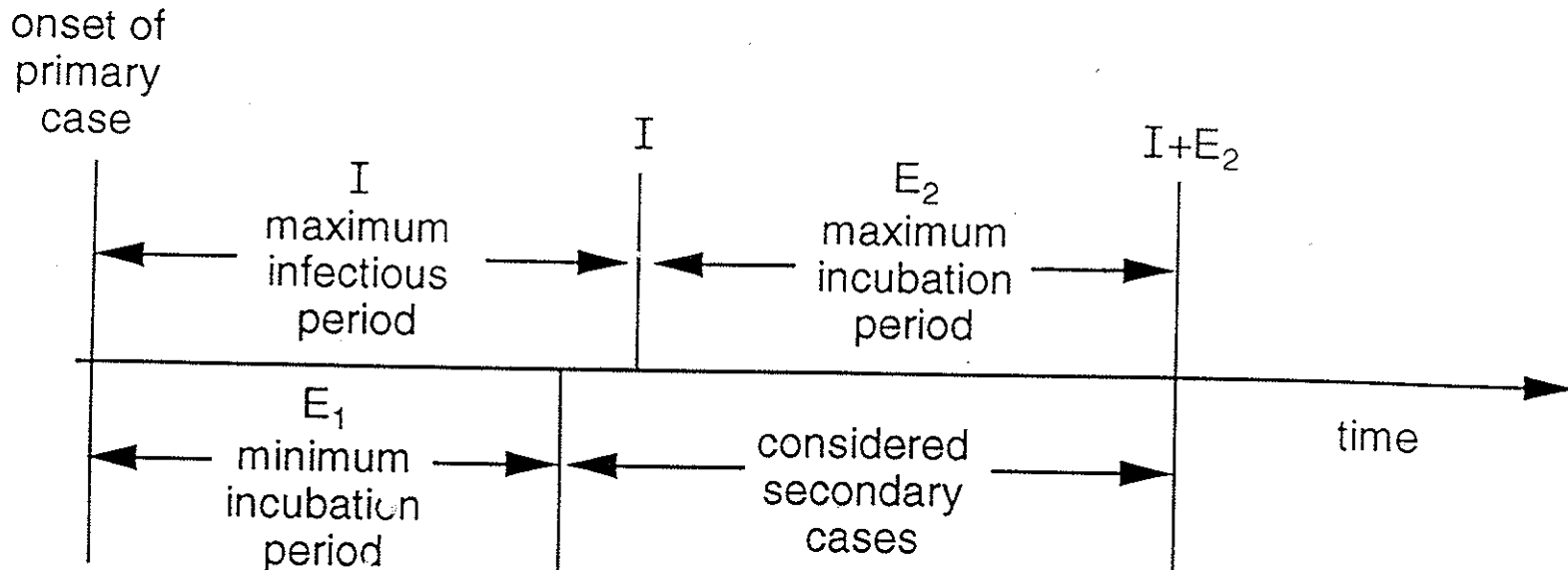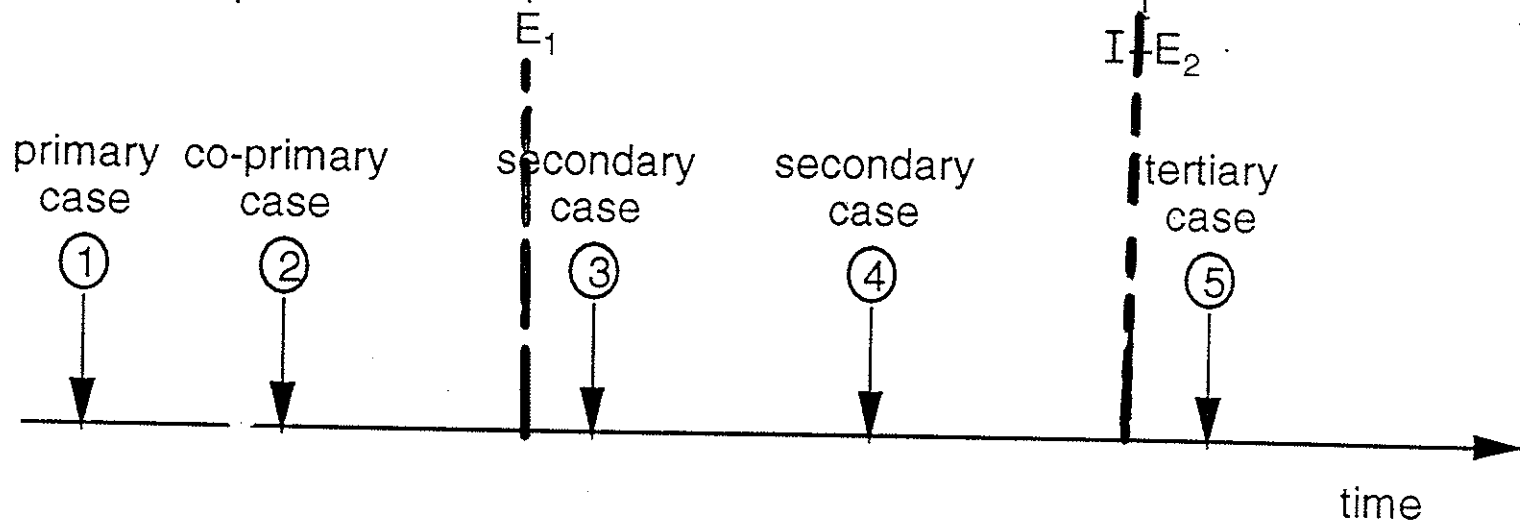
"IN ACCEPTING THE FAMILY
AS A SEMICLOSED GROUP, ONE
MUST BELIEVE THAT INFECTIONS
IN SUCCESSIVE FAMILY MEMBERS
WITHIN A SHORT PERIOD OF
TIME ARE MORE LIKELY TO BE
RELATED TO EACH OTHER THAN
TO HAVE BEEN SEPERATELY AC-
QUIRED FROM OUTSIDE SOURCES.
THIS IS ESSENTUALLY AN ACT
OF FAITH AND NEARLY IMPOSSIBLE
TO ESTABLISH IN TERMS OF EXACT
PROBABILITY."

BUCK, C.   ACUTE UPPER RESPIRATORY INFECTIONS IN FAMILIES
AM. J. HYG. 63, 1-12, 1956.

1. NATURAL HISTORY OF S-I-R INFECTION PROCESS

LATENT   INFECTIOUS   REMOVED

$$\longleftarrow T_1 \longrightarrow \qquad \longleftarrow \longrightarrow T_2 \longrightarrow$$
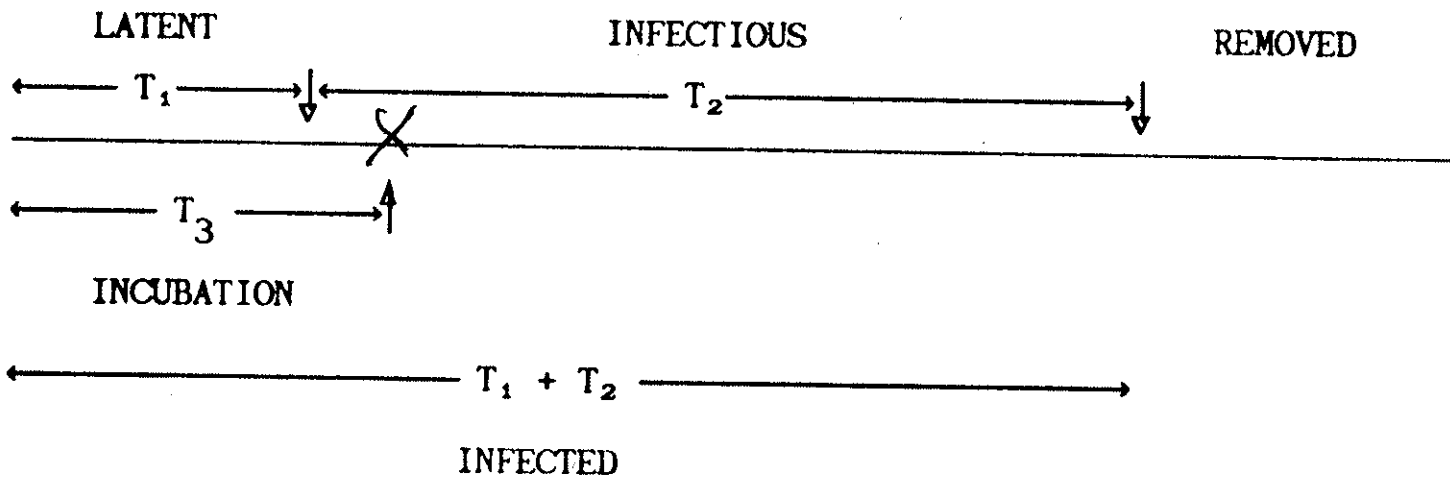
$$\longleftarrow T_3 \longrightarrow$$

INCUBATION

$$\longleftarrow T_1 + T_2 \longrightarrow$$

INFECTED

$$FLU: \quad \mathcal{E}(T_1) \cong \mathcal{E}(T_3) \cong 2 \ DAYS$$

$$\mathcal{E}(T_2) \cong 4 \ DAYS$$

$$MEASLES: \quad \mathcal{E}(T_1) \cong 8 \ DAYS$$

$$\mathcal{E}(T_3) \cong 9 \ DAYS$$

$$\mathcal{E}(T_2) \cong 4.5$$

PROBABILITY MODEL FOR ACUTE INFECTIOUS DISEASES

Π_{ij} PROB. THAT i OUT OF j PERSONS ARE INFECTED, $i \leq j$

ASSUMPTIONS:

1. T AND T_2 ARE FIXED ($B = b^t$ and $Q = q^{t_2}$)

2. RANDOM MIXING WITHIN HOUSEHOLDS

3. PROB. FAMILY MEMBER INFECTED FROM COMM.

    IS INDEPENDENT OF # INFECTED IN HOUSEHOLD

3.   TRANSMISSION PROBABILITIES

WITHIN FAMILIES:

p = PROB. INFECTED GIVEN CONTACT PER UNIT OF TIME

$q = 1 - p$

$Q = E(q^{T_2})$.

$SAR = (1 - Q) \times 100$ SECONDARY ATTACK RATE

FROM COMMUNITY SOURCES:

T = DURATION OF THE EPIDEMIC IN THE COMMUNITY

a = PROB. INFECTED FROM COMMUNITY SOURCES OVER

   THE COURSE OF THE EPIDEMIC, $b = 1 - a$

$B = E(b^T)$

$CPI = 1 - B$ COMMUNITY PROBABILITY OF INFECTION

PARTITION SAMPLE SPACE INTO i INFECTIVES AND j - i ESCAPES,

WHERE THERE ARE $\binom{j}{i}$ SUCH PARTITIONS

FOR ONE PARTITION OF THE SAMPLE SPACE:

E - EVENT i OUT OF i INFECTED

$P(E) = \Pi_{ii}$

F - EVENT j - i ESCAPE

$P(F|E) = B^{j-i} Q^{i(j-i)}$, i < j

THEN WE HAVE

$\Pi_{ij} = P(E \cap F) = P(E) \, P(F|E)$

$\Pi_{ij} = \Pi_{ii} \, B^{j-i} \, Q^{i(j-i)}$, i < j.

$\Pi_{jj} = 1 - \sum_{i=0}^{j-1} \Pi_{ij}$

SPECIAL CASES

BINOMIAL: WHEN $Q = 1$, THEN

$$\Pi_{1j} = \binom{j}{1} (1 - B)^1 B^{j-1}, \quad 1 \leq j$$

REED-FROST-EN'KO MODEL: WHEN AND $B = 1$, AND THERE ARE $1_o \geq 1$ INITIAL INFECTIVES, THEN

$$\Pi_{1j} = \binom{j}{1} \Pi_{11} Q^{(1_o+1)(j-1)}, \quad 1 < j.$$

$$\Pi_{jj} = 1 - \sum_{1=0}^{j-1} \Pi_{1j}$$

BAILEY (1975), LUDWIG (1975)

"DISCORDANT COUPLES MODEL": WHEN $j = 1$, $1_o = 1$, AND $B \leq 1$, THEN $\Pi_{01} = BQ$ AND $\Pi_{11} = 1 - BQ$.

$$\Pi_{ij} = \binom{j}{i} \Pi_{11} B^{j-i} Q^{i(j-i)}, \quad i < j.$$

$$\Pi_{jj} = 1 - \sum_{i=0}^{j-1} \Pi_{ij}$$

LONGINI & KOOPMAN (1982), LONGINI, ET AL. (1982)

$$\Pi_{01} = B$$

$$\Pi_{11} = 1 - B$$

PROBABILITY MASS FUNCTION FOR A HOUSEHOLD WITH FIVE SUSCEPTIBLES
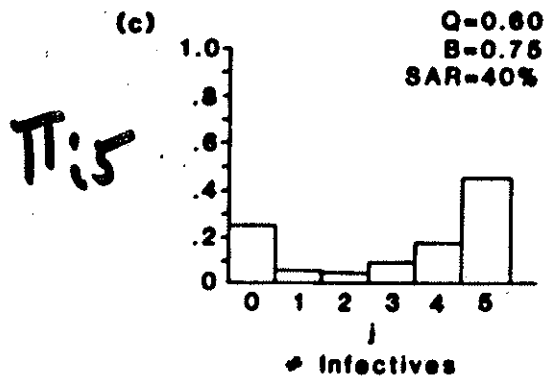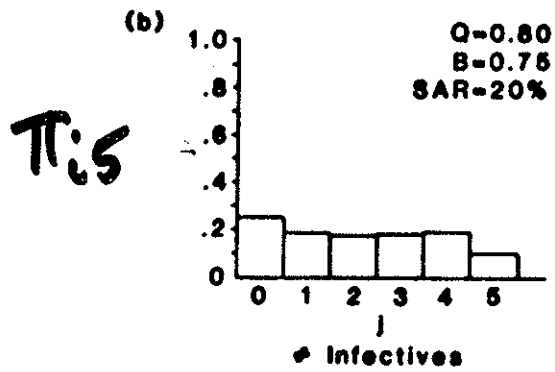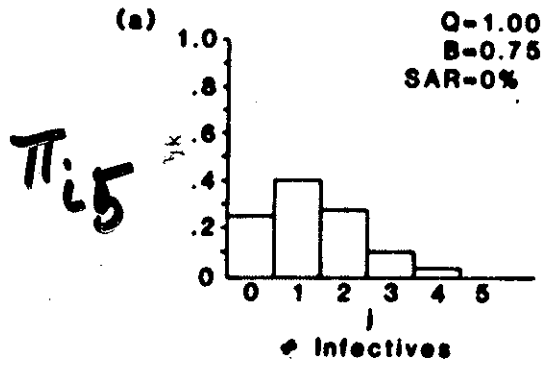
DISTRIBUTION OF RESPIRATORY SYMPTOMS
DURING H3N2 EPIDEMIC PERIOD OF 1977-78 IN TECUMSEH

NUMBER INFECTED	NUMBER OF INDIVIDUALS/HOUSEHOLD				
	1	2	3	4	5
0	38	48 (48)	13 (13)	6 (6)	4 (4)
1	19	29 (44)	19 (27)	13 (20)	3 (11)
2		25 (10)	20 (18)	17 (24)	9 (12)
3			10 (4)	23 (13)	8 (6)
4				7 (3)	5 (2)
5					6 (0)
TOTAL	57	102	62	66	35

$n_1 \qquad n_2 \qquad n_3 \qquad h_4$

LIKELIHOOD FUNCTION

n – Total number of households

$n_{\cdot j}$ – Number of households with j initial

   susceptibles

n_{ij} – Number of households with i of j

   individuals infected

$$\left(\sum_i n_{ij} = n_j \ , \ \sum_j \sum_i n_{ij} = \sum_j n_j = n \right)$$

$$L(Q,B) = \prod_j \ \prod_i \ \pi_{ij}^{n_{ij}}$$

LONGINI & KOOPMAN (1982)

$$\hat{B} = \frac{1}{n} \sum_j n_{\cdot j} \left(\frac{n_{0j}}{n_{\cdot j}} \right)^{1/j} \quad \sim 90\% \ \text{EFF.}$$

DISTRIBUTION OF INFECTIONS BY HOUSEHOLDS

Number of Susceptibles/Household

		j			
#Infected	1	2	3	. . . .	J
i					
0	n_{01}	n_{02}	n_{03}	. . . .	n_{0J}
1	n_{11}	n_{12}	n_{13}	. . . .	n_{1J}
2	–	n_{22}	n_{23}	. . . .	n_{2J}
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
.	.	.	.		.
J	–	–	–	. . . .	n_{JJ}
	$n_{.1}$	$n_{.2}$	$n_{.3}$	. . . .	$n_{.J}$ n

n_{ij}

DISTRIBUTION OF INFLUENZA A(H3N2) INFECTIONS

1980-81 TECUMSEH DATA

Number of Susceptibles/Household

	j				
Number Infected	1	2	3	4	
0	44	62	47	38	
1	10	13	8	11	
2	–	9	2	7	
3	–	–	3	5	
4	–	–	–	1	
	54	84	60	62	260
	$n_{.1}$	$n_{.2}$	$n_{.3}$	$n_{.4}$	

Maximum Likelihood Estimation

$$\ln L = c + \sum_j \sum_i n_{ij} \left\{ \ln \pi_{jj} + (j-i) \ln B + j(a-i) \ln Q \right\}$$

$$\{0 \div B \div 1, \; 0 \leq Q \leq 1\}$$

MLE's $\hat{Q}$ and $\hat{B}$ are solutions of

$$0 = \left. \frac{\partial \ln L}{\partial Q} \right|_{\hat{Q}, \hat{B}} = \sum_j \sum_i n_{ij} \left\{ \frac{1}{\pi_{ii}} \left(\frac{\partial \pi_{ii}}{\partial Q} \right) + \frac{i(j-i)}{Q} \right\},$$

$$0 = \left. \frac{\partial \ln L}{\partial B} \right|_{\hat{Q}, \hat{B}} = \sum_j \sum_i n_{ij} \left\{ \frac{1}{\pi_{ii}} \left(\frac{\partial \pi_{ii}}{\partial B} \right) + \frac{j-i}{B} \right\}.$$

Solve by method of scoring which also provides

$\hat{Var}(\hat{Q})$, $\hat{Var}(\hat{B})$ and $\hat{Cov}(\hat{Q}, \hat{B})$.

also $\hat{Q} \sim N$ and $\hat{B} \sim N$ large N.

Mongini & Koopman (1982)

INFORMATION MATRIX

$$-\mathcal{E}\left(\frac{\partial^2 \ln h}{\partial Q^2}\right) = \sum_k \sum_J n_k \, m_{Jk} \left\{ \frac{1}{m_{JJ}} \left[\frac{1}{m_{JJ}} \left(\frac{\partial m_{JJ}}{\partial Q}\right)^2 - \frac{\partial^2 m_{JJ}}{\partial Q^2} \right] \right.$$

$$\left. + \left[\frac{J(k-J)}{Q^2} \right] \right\}$$

$$-\mathcal{E}\left(\frac{\partial^2 \ln h}{\partial Q \partial B}\right) = \sum_k \sum_J n_k \left(\frac{m_{Jk}}{m_{JJ}}\right) \left[\frac{1}{m_{JJ}} \left(\frac{\partial m_{JJ}}{\partial Q}\right)\left(\frac{\partial m_{JJ}}{\partial B}\right) \right.$$

$$\left. - \left(\frac{\partial^2 m_{JJ}}{\partial Q \partial B} \right) \right]$$

$$-\mathcal{E}\left(\frac{\partial^2 \ln h}{\partial B^2}\right) = \sum_k \sum_J n_k \, m_{Jk} \left\{ \frac{1}{m_{JJ}} \left[\frac{1}{m_{JJ}} \left(\frac{\partial m_{JJ}}{\partial B}\right)^2 - \left(\frac{\partial^2 m_{JJ}}{\partial B^2}\right) \right] \right.$$

$$\left. + \left[\frac{k-J}{B^2} \right] \right\}$$

LOG-LINEAR MODEL FOR SMALL POPULATION GROUPS

$$\pi_{ij} = \binom{j}{i} \pi_{11} B^{(j-1)} Q^{j(j-1)} , \quad i < j$$

$$\text{LOG } \pi_{ij} - \text{LOG } \pi_{11} - \text{LOG} \binom{j}{i} = (j-1) \text{ LOG } B + j(j-1) \text{ LOG } Q.$$
$$i < j$$

$$\text{Let } a_{ij} = \text{LOG} \binom{j}{i} , \quad \beta = \text{LOG } B , \quad \gamma = \text{LOG } Q ,$$
$$(B = e^{\beta}) , \quad (Q = e^{\gamma})$$

and substituting we have

$$\text{LOG } \pi_{ij} - \text{LOG } \pi_{11} - a_{ij} = (j-1) \beta + j(j-1) \gamma$$

$$\underline{A} \text{ LOG } \underline{\pi} - \underline{a} = \underline{X} \underline{\beta}$$

Log-Linear Model

WEIGHTED LEAST - SQUARES

HABER, LONGINI, COTSONIS (1988)

PARAMETER ESTIMATION COMPARING PROBABILITY AND LOG-LINEAR MODELS – INFLUENZA A(H3N2) 1980–81 TECUMSEH DATA

I. Probability Model Using Maximum Likelihood

Number of Susceptibles/Household

Number Infected	1 Obs	1 Exp	2 Obs	2 Exp	3 Obs	3 Exp	4 Obs	4 Exp
0	44	47.6	62	65.3	47	41.1	38	37.4
1	10	6.4	13	14.0	8	10.6	11	10.3
2	–	–	9	4.7	2	5.7	7	6.6
3	–	–	–	–	3	2.6	5	4.9
4	–	–	–	–	–	–	1	2.8
Total	54	54.0	84	84.0	60	60.0	62	62.0

$$\chi^2_8 = 11.6 \quad p = .17 \, ,$$

$$\hat{SAR} = 20.1 \pm 3.5 \, , \quad \hat{CPI} = .118 \pm .013$$

II. Log-Linear Model Using Weighted Least Squares

$$\chi^2_8 = 9.9 \quad p = .27 \, ,$$

$$\hat{SAR} = 22.8 \pm 2.0 \, , \quad \hat{CPI} = .110 \pm .007$$

OBSERVED AND EXPECTED DISTRIBUTION OF INFLUENZA B
1977 - TECUMSEH DATA
Number of Susceptibles/Household

Number Infected	1 Obs.	1 Exp.	2 Obs.	2 Exp.	3 Obs.	3 Exp.	4 Obs.	4 Exp.	5 Obs.	5 Exp.
0	42	44.1	79	71.4	23	26.4	16	17.8	6	5.5
1	14	11.9	24	36.9	24	19.6	19	16.9	7	6.2
2	-	-	12	6.7	7	6.9	8	8.5	4	4.0
3	-	-	-	-	0	1.1	2	2.5	1	1.7
4	-	-	-	-	-	-	1	.4	0	.5
5	-	-	-	-	-	-	-	-	0	.1
Total	56	56.0	115	115.0	54	54.0	46	46.1	18	18.0

$\hat{Q} = .960$ $\qquad \hat{B} = .788$ $\quad CPI = .212$

$Var(\hat{Q}) = .0007$ $\qquad Var(\hat{B}) = .0003$ $\qquad Cov(\hat{Q},\hat{B}) = -.0002$

$X^2(13) = 15.1$ $\qquad P = .30$

Actual SAR = 4.0% $\qquad$ Apparent SAR = 22.6%

OBSERVED AND EXPECTED DISTRIBUTION OF RESPIRATORY SYMPTOMS DURING H3N2 EPIDEMIC PERIOD OF 1977-78 IN TECUMSEH

NUMBER ILLNESSES	1		2		3		4		5	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
0	38	35.4	48	39.5	13	14.9	6	9.9	4	3.3
1	19	21.6	29	40.0	19	18.9	13	14.9	3	4.8
2			25	22.5	20	17.8	17	16.4	9	6.3
3					10	10.4	23	15.9	8	7.6
4							7	9.8	5	7.9
5									6	5.2
TOTAL	57	57.0	102	102.0	62	62.0	66	66.0	35	35.1

$\hat{Q} = 0.835 \pm 0.024$, ACTUAL SAR $= 16.5 \pm 2.4$, APPARENT SAR $= 46.7$.

$\hat{B} = 0.622 \pm 0.020$; $\hat{CPI} = 0.378 \pm 0.020$, $Cov(\hat{Q},\hat{B}) = -0.0002$.

$x^2(13) = 14.95$, P VALUE $= 0.31$, N $= 322$.

OBSERVED AND EXPECTED DISTRIBUTION OF INFLUENZA A(H1N1) 1978-1979
TECUMSEH DATA

Number Infected	1		2		3	
	Obs.	Exp.	Obs.	Exp.	Obs.	Exp.
0	28	34.1	22	18.7	4	2.3
1	33	26.9	18	22.9	3	3.2
2			20	18.4	5	4.0
3					1	3.5
Total	61	61.0	60	60.0	13	13.0

$\widehat{SAR} = 22.5 \pm 8.6$, Apparent SAR $= 48.2$, $\widehat{CPI} = 0.441 \pm 0.038$.

$x^2(4) = 7.62$, $p = .106$

PARAMETERS

r – level of susceptibility $(r = 1, 2, \ldots, R)$

h – level of infectiousness $(h = 1, 2, \ldots, H)$

B_r = probability of escape from the community

$CPI_r = 1 - B_r$ community probability of infection

Q_{rh} = probability of escape from infected household member

$SAR_{rh} = (1 - Q_{rh}) \times 100$ secondary attack rate

Household with s initial susceptibles

Susceptibility $\quad\quad \mathbf{r} = (r_1, \ldots, r_s)$

Infectiousness $\quad\quad \mathbf{h} = (h_1, \ldots, h_s)$

Infection $\quad\quad\quad \mathbf{x} = (x_1, \ldots, x_s)$

where $\quad\quad x_j = \begin{cases} 0 \text{ if } j\text{-th person not infected.} \\ 1 \text{ if } j\text{-th person infected.} \end{cases}$

k_h – number persons infected at h-th level

$$k = \sum_h k_h$$

Find $P[\mathbf{x} | \mathbf{r}, \mathbf{h}] \quad : \quad 2^s \times R^s \times H^s$

e.g., $R = 2$, $H = 2$, $s = 5$

$2^5 \times 2^5 \times 2^5 = 32{,}768$

PROBABILITY MODEL

$$p(x|r,h) = \begin{cases} \prod\limits_{i=1}^{s} B_{r_i} , & \text{for } k = 0 . \\[2ex] p(1_k|r_j,h_j) \prod\limits_{i:x_i=0} B_{r_i} \prod\limits_{h=1}^{H} [Q_{r_i h}]^{k_h} , & \text{for } k=1,\ldots,s-1 . \\[2ex] 1 - \sum\limits_{x:\Sigma x_i < k} p(x|r,h) , & \text{for } k = s . \end{cases}$$

where $j = (j_1,\ldots,j_k)$ are the indices of the k infected individuals,

$r_j = (r_{j_1},\ldots,r_{j_k})$ and $h_j = (h_{j_1},\ldots,h_{j_k})$. Also, 1_k denotes the array

$(1,\ldots,1)$ of order k.

Constructed recursively -- e.g., R = 2, H = 2, s = 5

There are 37,448 probabilities.

$$\underline{EXAMPLE} \qquad R = 2 \;,\quad S = 1, 2 \;,\; H = 1$$

WHEN $\quad S = 1 :\qquad 2^1 \times 2^1 = 4$ PROBABILITIES

$$P(0|1) = B_1 \qquad P(1|1) = 1 - B_1$$
$$P(0|2) = B_2 \qquad P(1|2) = 1 - B_2$$

WHEN $\quad S = 2 :\qquad 2^2 \times 2^2 = 16$ PROBABILITIES

$$P(0,0|1,1) = B_1^2 \qquad P(0,0|1,2) = P(0,0|2,1) = B_1 B_2$$
$$P(0,0|2,2) = B_2^2$$

$$P(1,0|1,1) = B_1 Q_1 \, P(1|1) = B_1 (1-B_1) Q_1 = P(0,1|1,1)$$
$$P(1,0|1,2) = B_2 Q_2 \, P(1|1) = B_2 (1-B_1) Q_2 = P(0,1|2,1)$$
$$P(1,0|2,1) = B_1 Q_1 \, P(1|2) = B_1 (1-B_2) Q_1 = P(0,1|1,2)$$
$$P(1,0|2,2) = B_2 Q_2 \, P(1|2) = B_2 (1-B_2) Q_2 = P(0,1|2,2)$$

$$P(1,1|1,1) = 1 - B_1^2 - 2 B_1 (1-B_1) Q_1$$
$$P(1,1|1,2) = 1 - B_1 B_2 - 2 B_2 (1-B_1) Q_2 = P(1,1|2,1)$$
$$P(1,1|2,2) = 1 - B_2^2 - 2 B_2 (1-B_2) Q_2$$

Infection attack rates by pre-season hemagglutination inhibition (HI) titer level stratified by age group: Influenza A(H3N2) epidemic seasons (1977-1978) and (1980-1981) combined in Tecumseh, Michigan

Infection status

Pre-HI titer (1:x)	No. infected	No. not infected	Total	Attack Rate[†]	Risk Ratio[††]
Children (0-17)					
Low level (x<8)	100	200	300	0.333	3.330*
Higher level (8≤x≤64)	20	180	200	0.100	
Total	120	380	500**	0.240	
Adults (18+)					
Low level (x<8)	96	440	536	0.179	1.884*
Higher level (8≤x≤64)	42	402	444	0.095	
Total	138	842	980**	0.141	

[†] Attack rate = No. infected/No. at risk

[††] Risk ratio = Ratio of the attack rates

* $p < 0.001$

** The total of 1480 individuals does not include the 26 "immune" individuals

The risk ratios across levels of age or across levels of pre-season HI tite are different ($p < 0.0001$) using the chi-square test for lack of interaction.

STATISTICS NOT APPROPRIATE

Observed and expected frequencies for households by the number of susceptibles from the influenza A(H3N2) seasons (1977-1978) and (1980-1981) combined, in Tecumseh, Michigan, stratified by pre-season HI titer level[†]

No. susceptible/household Pre-season titer level		No. infected/household Pre-season titer level		No. of households	
low	higher	low	higher	observed	expected
1	0	0	0	63 { 45	52.7
		1	0	18	10.3
0	1	0	0	70 { 65	63.6
		0	1	5	6.4
2	0	0	0	71 { 52	49.6
		1	0	11	14.4
		2	0	8	7.0
1	1	0	0	66 { 52	50.1
		0	1	2	3.8
		1	0	8	9.6
		1	1	4	2.5
0	2	0	0	45	42.9
		0	1	6	(8.5(a
		0	2	1	0.6(a
3	0	0	0	17	16.9
		1	0	4	5.5
		2	0	3	3.9
		3	0	5	2.7
2	1	0	0	28	24.7
		0	1	1	1.3(b
		1	0	6	7.0
		1	1	0	1.4(b
		2	0	2	3.3
		2	1	2	1.2(b
1	2	0	0	16	17.2
		0	1	6	2.5(c
		0	2	0	0.1(c
		1	0	2	3.2
		1	1	1	1.7(c
		1	2	0	0.2(c

0	3	0	0	11	11..2
		0	1	4	3.3(d
		0	2	0	0.5(d
		0	3	0	0.0(d
4	0	0	0	16	13.7
		1	0	4	4.3
		2	0	6	3.5
		3	0	0	3.5
		4	0	2	3.0
3	1	0	0	13	13.8
		0	1	0	0.6(e
		1	0	6	4.3
		1	1	1	0.6(e
		2	0	1	3.0
		2	1	0	0.8(e
		3	0	5	2.0(e
		3	1	0	0.8(e
2	2	0	0	11	11.5
		0	1	0	1.3(f
		0	2	1	0.1(f
		1	0	1	3.2
		1	1	3	1.2(f
		1	2	1	0.1(f
		2	0	3	1.5(f
		2	1	0	1.0(f
		2	2	0	0.2(f
1	3	0	0	10	12.5
		0	1	5	2.7
		0	2	0	0.3(g
		0	3	0	0.0(g
		1	0	2	2.3(g
		1	1	1	1.7(g
		1	2	2	0.4(g
		1	3	0	0.0(g
0	4	0	0	10	8.2
		0	1	2	3.1(h
		0	2	0	0.7(h
		0	3	0	0.1(h
		0	4	0	0.0(h

5	0	0	0	3	2.4
		all other		3	3.6
4	1	0	0	2	2.7
		all other		4	3.3
3*	2*			4	–
2	3	0	0	4	5.2
		all other		6	4.8
1*	4*			2	–
0*	5*			3	–
Total				567	

Point estimates and standard errors:

Low pre-season antibody titer $\quad \hat{CPI}_1 = 0.164 \pm 0.015 \quad \hat{SAR}_1 = 26.0 \pm 3.0$

Higher pre-season antibody titer $\hat{CPI}_2 = 0.092 \pm 0.013 \quad \hat{SAR}_2 = 2.1 \pm 2.6$

Overall goodness-of-fit $\quad \chi^2(28df)^* = 33.01, \ p = 0.235$

Tests: $H_o : CPI_1/CPI_2 = 1 \quad \hat{RR} = 1.783 \quad p < 0.0005$

$\qquad H_o : SAR_1/SAR_2 = 1 \quad \hat{RR} = 12.4 \quad p < 0.0001$

† Only households with 5 or fewer susceptibles are included.

(a)-(h) Outcomes with the same letter were pooled for the goodness-of-fit test.

* These combinations were not included in the goodness-of-fit test.

VARIABLE SUSCEPTIBILITY: PRE-SEASON ANTIBODY × AGE

$r = 1,2,3,4.$

1. LOW, 0 - 17

2. HIGHER, 0 - 17

3. LOW, 18^+

4. HIGHER, 18^+

Estimated CPIs and SARs

Age		Pre-season titer			RR
		X < 8		$8 \leq X \leq 64$	Low to high titer
Children	CPI_1	0.23 ± 0.03	CPI_2	0.09 ± 0.03	2.5**
	SAR_1	36.6 ± 6.2	SAR_2	3.4 ± 4.7	10.8**
Adults	CPI_3	0.13 ± 0.02	CPI_4	0.09 ± 0.01	1.5
	SAR_3	18.2 ± 4.4	SAR_4	1.6 ± 3.7	11.4**

Goodness-of-fit $\chi^2 = 24.8$ (df = 23), p = 0.36, **p < 0.05

Antibody Efficacy: Direct Measures

$$AE_{S,CI} = 1 - \frac{AR_H}{AR_L}$$

AR   Lev. IV

$$AE_{S,CI} = 1 - \frac{CPI_H}{CPI_L}$$

CPI   Lev. IV

$$AE_{S,P} = 1 - \frac{SAR_H}{SAR_L}$$

SAR   Lev. I

Estimated Antibody Efficacy

Index	Children	Adults	Crude
AR	0.700	0.469	0.588
CPI	0.503	0.321	0.439
SAR	0.907	0.912	0.919

Table 3

Maximum likelihood estimates and standard errors for parameters of the model of influenza A(H3N2) infections in 1977–1978 and 1980–1981 combined epidemics in Tecumseh, Michigan, with unrestricted contact parameters by age (0–17 vs 18+)

	Estimate	Transformation
Constant distribution: $T_1 \equiv 4.1$		
	[a]$\beta_{11} = .0805 \pm .0208$	$SAR_{11} = 28.1186 \pm 6.1227$
	$\beta_{12} = .0354 \pm .0291$	$SAR_{12} = 13.4996 \pm 10.314$
	$\beta_{21} = .0268 \pm .0135$	$SAR_{21} = 10.4080 \pm 4.9593$
	$\beta_{22} = .0401 \pm .0127$	$SAR_{22} = 15.1662 \pm 4.4096$
Child	$B_1 = .8184 \pm .0254$	$CPI_1 = .1816 \pm .0254$
Adult	$B_2 = .8897 \pm .0128$	$CPI_2 = .1103 \pm .0128$
Log likelihood	$= -522.333$	

* ADDY, ET AL: _BIOMETRICS_ 47, 961–974 (1991)

The Transmissibility and Control of Pandemic Influenza A (H1N1) Virus

Yang Yang,[1] Jonathan D. Sugimoto,[1,2] M. Elizabeth Halloran,[1,3] Nicole E. Basta,[1,2] Dennis L. Chao,[1] Laura Matrajt,[4] Gail Potter,[5] Eben Kenah,[1,3,6] Ira M. Longini Jr.[1,3*]

[1]Center for Statistics and Quantitative Infectious Diseases, Fred Hutchinson Cancer Research Center and the University of Washington, Seattle, WA, USA. [2]Department of Epidemiology, School of Public Health, University of Washington, Seattle, WA, USA. [3]Department of Biostatistics, School of Public Health, University of Washington, Seattle, WA, USA. [4]Department of Applied Mathematics, University of Washington, Seattle, WA, USA. [5]Department of Statistics, University of Washington, Seattle, WA, USA. [6]Department of Global Health, University of Washington, Seattle, WA, USA.

*To whom correspondence should be addressed. E-mail: longini@scharp.org

Pandemic influenza A (H1N1) 2009 (pandemic H1N1) is spreading throughout the planet. It has become the dominant strain in the southern hemisphere, where the influenza season is underway. Here, based on reported case clusters in the USA, we estimate the household secondary attack rate for pandemic H1N1 to be 27.3% [95% confidence interval (CI) 12.2%–50.5%]. From a school outbreak, we estimate a school child infects 2.4 (95% CI: 1.8–3.2) other children within the school. We estimate the basic reproductive number, R_0, to range from 1.3 to 1.7 and the generation interval to range from 2.6 to 3.2 days. We use a simulation model to evaluate the effectiveness of vaccination strategies in the USA for Fall 2009. If vaccine were available soon enough, vaccination of children, followed by adults, reaching 70% overall coverage, in addition to high risk and essential workforce groups, could mitigate a severe epidemic.

corresponding to the rainy season. The last influenza pandemic was the Hong Kong A (H3N2) 1968–1969 pandemic. At that time, the first large epidemic was in Hong Kong in July 1968, followed by epidemics in South East Asia in August–September 1968, in the upper northern hemisphere between September 1968 and April 1969 (peaking in late December 1968 and early January 1969) and in the lower southern hemisphere between June and September 1969 (9). In the USA and the upper northern hemisphere, shifted (i.e., pandemic) or drifted strains of influenza tend to have a relatively small Spring "herald wave" before returning in the Fall (10). In the upper northern hemisphere, the 1918–1919 A (H1N1) pandemic had a mild Spring 1918 herald wave, followed by a severe second wave in the Fall of 1918. Pandemic Asian influenza A (H2N2), 1957–1958, caused mid-Summer 1957 outbreaks in Louisiana schools that were open in the Summer because of the need for children helping

Estimating Transmissibility

- Natural history of disease

  - Incubation period: $\Pr(\delta = l) = \alpha_l,$ subject to $\sum_{l=\delta_{\min}}^{\delta_{\max}} \alpha_l = 1$

  - Infectious period: from ILI onset $\tilde{t}_i$ to $\tilde{t}_i + D$, and probability of being infectious is β_d for day $\tilde{t}_i + d - 1$.

- Transmission model

  - Daily transmission probabilities: b (C2P) and $p(t)$ (P2P)
  - Daily escape probability

  $$e_i(t) = (1-b) \prod_{j \in h(i)} (1 - p(t)\beta_{t-\tilde{t}_j+1})$$

  - Likelihood

  $$L_i(b, p \mid \tilde{\mathbf{t}}_h) = \begin{cases} \prod_{t=1}^{T} e_i(t), & \tilde{t}_i = \infty, \\ \sum_{t=\tilde{t}_i-\delta_{\max}}^{\tilde{t}_i-\delta_{\min}} \alpha_{\tilde{t}_i-t}(1-e_i(t)) \prod_{\tau=1}^{t} e_i(t), & \tilde{t}_i < T \end{cases}$$

Estimating Transmissibility (continue)

- Accounting for missing data

  - Household sizes and some ILI onset dates are missing

  - Multiple imputation (Schaffer, 1997).

- Calculating SAR and R_0 for US data

$$SAR = 1 - \prod_{d=1}^{D} (1 - p\beta_d)$$

$$R_0 = f(R_H + R_C + R_S) + (1 - f)(R_H + R_C)$$

$$R_H = N_H \times SAR_H, \quad R_S = N_S \times SAR_S, \quad R_C = \rho R_H$$

- Calculating R_0 for Mexican data
  - For large population, chain binomial becomes Poisson.
  - Assume each case corresponds to K-1 uninfected people.

Household SAR

	Individuals 11 and 12 are infected			Individuals 11 and 12 are not infected		
	Estimates	s.e.	95% CI	Estimates	s.e.	95% CI
Individual 10 is an index case						
p_H	0.0625	0.0268	(0.0265, 0.140)	0.0457	0.0213	(0.0181, 0.111)
SAR_H	0.273	0.101	(0.122, 0.505)	0.207	0.0870	(0.0849, 0.425)
R_H	0.82	0.303	(0.397, 1.693)	0.622	0.261	(0.274, 1.416)
Individual 10 is a secondary case						
p_H	0.0709	0.0288	(0.0314, 0.152)	0.0543	0.0232	(0.0231, 0.122)
SAR_H	0.304	0.105	(0.143, 0.535)	0.242	0.0915	(0.107, 0.459)
R_H	0.913	0.314	(0.466, 1.791)	0.726	0.274	(0.346, 1.523)

Household SAR: Sensitivity

|		Incubation Period			Infectious Period		
		Estimates	s.e.	95% CI	Estimates	s.e.	95% CI
Short[†]							
	p_H	0.0447	0.0208	(0.0177, 0.108)	0.0662	0.0305	(0.0262, 0.157)
	SAR_H	0.204	0.0854	(0.0834, 0.418)	0.203	0.0853	(0.0833, 0.417)
	R_H	0.611	0.256	(0.268, 1.390)	0.610	0.256	(0.268, 1.388)
Long[†]							
	p_H	0.0462	0.0216	(0.0183, 0.112)	0.0376	0.0176	(0.0149, 0.0919)
	SAR_H	0.210	0.0880	(0.0857, 0.429)	0.211	0.0884	(0.0862, 0.431)
	R_H	0.629	0.264	(0.276, 1.432)	0.632	0.265	(0.278, 1.438)

[†] Short incubation period: $\delta_{min} = 1$, $\delta_{max} = 3$ and $(\alpha_1, \alpha_2, \alpha_3) = (0.6, 0.3, 0.1)$

Long incubation period: $\delta_{min} = 1$, $\delta_{max} = 4$ and $(\alpha_1, \alpha_2, \alpha_3, \alpha_4) = (0.25, 0.25, 0.25, 0.25)$

Short infectious period: $D = 7$ and $(\beta_1, \ldots, \beta_7) = (1.0, 1.0, 0.6, 0.4, 0.2, 0.1, 0.05)$

Long infectious period: $D = 8$ and $(\beta_1, \ldots, \beta_8) = (1.0, 1.0, 1.0, 1.0, 1.0, 0.7, 0.4, 0.1)$
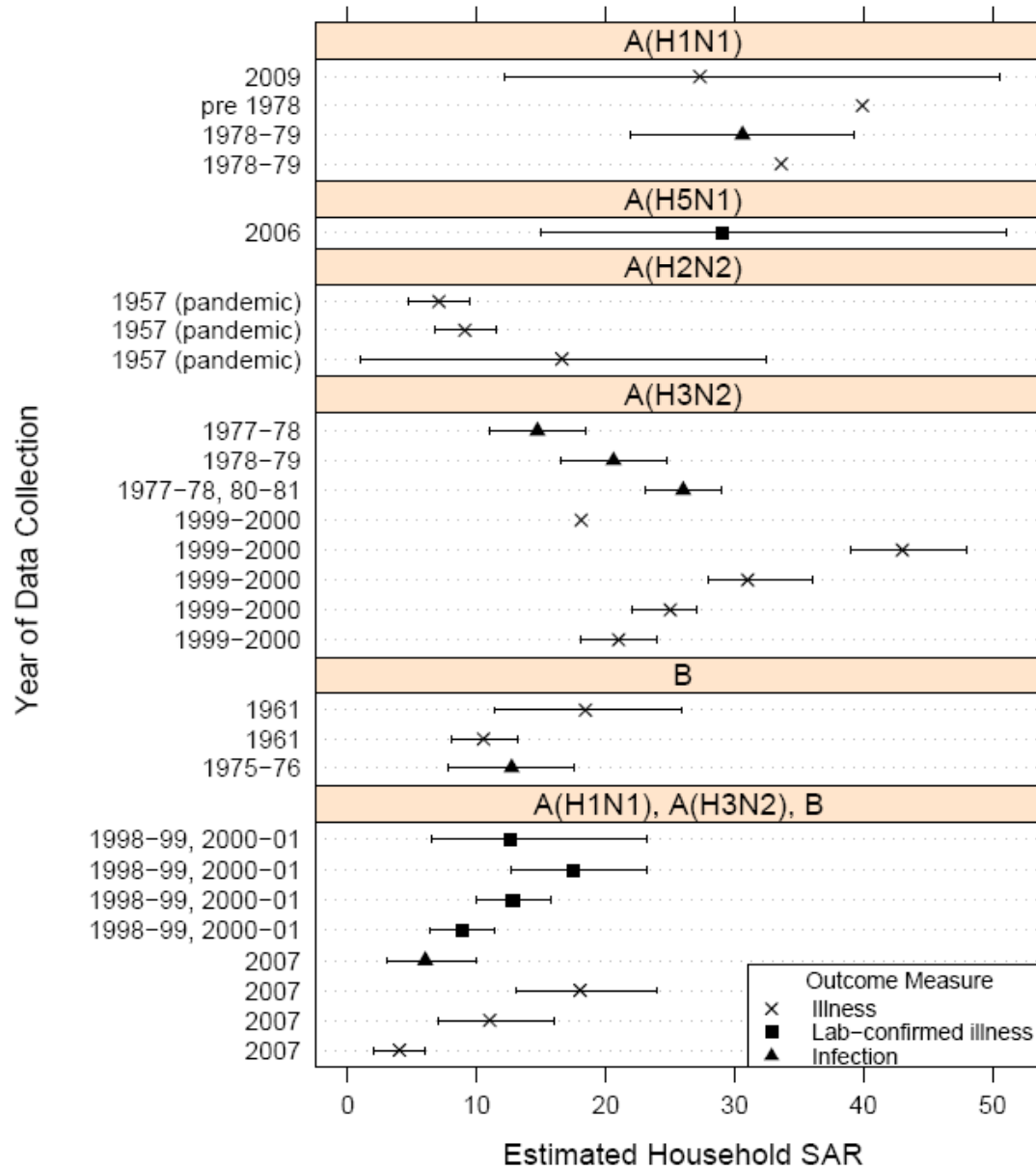
		Article	Cases SAR	Type of Confirmed	Data Source
A (H1N1)	2009	this paper	27.3 [12.2, 50.5]	ILL	U.S.
A (H5N1)	2006	(19) Yang 2007	29 [15, 51]	ILL-LAB	Sumatra
A (H2N2)	1957 (pandemic)	(20) Nishiura 2007	7.06 [4.73-9.44]	ILL	Tokyo
	1957 (pandemic)	(20) Nishiura 2007	9.07 [6.73-11.53]	ILL	Osaka
	1957 (pandemic)	(21) Longini 1982	16.6 [1, 32.5] (hh3)	ILL	Sugiyama
A (H1N1)	pre 1978	(21) Longini 1982	39.9 (hh4)	ILL	Hope,Simpson,Sutherland
	1978-79[a]	(22) Longini 1982	30.6 [21.9, 39.3]	LAB	Seattle
	1978-79	(21) Longini 1982	33.6 (hh5)	ILL	Hope,Simpson,Sutherland
A (H3N2)	1977-78	(22) Longini 1982	14.7 [11, 18.4]	LAB	Tecumseh
	1978-79	(22) Longini 1982	20.6 [16.5, 24.7]	LAB	Seattle
	1977-78, 80-81	(23) Longini 1988	26 [23, 29]	LAB	Tecumseh
	1999-2000	(24) Viboud 2004	18.1	ILL	France
	1999-2000	(25) Cauchemez 2004	43 [39, 48] (hh2)	ILL	France
	1999-2000	(25) Cauchemez 2004	31 [28, 36] (hh3)	ILL	France
	1999-2000	(25) Cauchemez 2004	25 [22, 27] (hh4)	ILL	France
	1999-2000	(25) Cauchemez 2004	21 [18, 24] (hh5)	ILL	France
B	1961[b]	(20) Nishiura 2007	18.41 [11.37, 25.95]	ILL	Osaka
	1961	(20) Nishiura 2007	10.51 [8.01, 13.15]	ILL	Osaka
	1975-76	(22) Longini 1982	12.7 [7.8, 17.6]	LAB	Seattle
Mixed	1998-1999 and 2000-2001[c]	(28) Yang 2009[d]	12.6 [6.5, 23.2]	ILL-LAB	
	1998-1999 and 2000-2001[e]	(31) Halloran 2007	17.5 [12.6, 23.2]	ILL-LAB	
	1998-1999 and 2000-2001[f]	(32) Yang 2007	12.8 [9.9,15.7]	ILL-LAB	
	1998-1999 and 2000-2001[g]	(11) Yang 2006	8.9 [6.4, 11.4]	ILL-LAB	
	2007[h]	(33) Cowling 2008	6 [3, 10]	LAB	Hong Kong trial
	2007	(33) Cowling 2008	18 [13, 24]	ILL	Hong Kong trial
	2007	(33) Cowling 2008	11 [7, 16]	ILL	Hong Kong trial
	2007	(33) Cowling 2008	4 [2, 6]	ILL	Hong Kong trial

Table: Parameters used for measuring various effects of vaccination∗

	Comparison groups and effect		
Level Parameter choice	Susceptibility	Infectiousness	Combined change in susceptibility and infectiousness
Conditional on exposure:			
I Transmission probability	$\text{VE}_{S,p}\dagger = 1 - \frac{p_{\cdot 1}}{p_{\cdot 0}}$	$\text{VE}_{I,p} = 1 - \frac{p_{1\cdot}}{p_{0\cdot}}$	$\text{VE}_{T,p} = 1 - \frac{p_{11}}{p_{00}}$

	Study design			
	I direct	IIA indirect	IIB total	III overall
Unconditional:				
II Incidence or hazard rate, IR, λ	$\text{VE}_{S,IR} = 1 - \frac{\text{IR}_{A1}}{\text{IR}_{A0}}$	$\text{VE}_{IIA,IR} = 1 - \frac{\text{IR}_{A0}}{\text{IR}_{B0}}$	$\text{VE}_{IIB,IR} = 1 - \frac{\text{IR}_{A1}}{\text{IR}_{B0}}$	$\text{VE}_{III,IR} = 1 - \frac{\text{IR}_{A\cdot}}{\text{IR}_{B\cdot}}$
	$\text{VE}_{S,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{A0}}$	$\text{VE}_{IIA,\lambda} = 1 - \frac{\lambda_{A0}}{\lambda_{B0}}$	$\text{VE}_{IIB,\lambda} = 1 - \frac{\lambda_{A1}}{\lambda_{B0}}$	$\text{VE}_{III,\lambda} = 1 - \frac{\lambda_{A\cdot}}{\lambda_{B\cdot}}$
III Proport. hazards, PH	$\text{VE}_{S,PH} = 1 - e^{\beta_1}$	NA	NA	NA
IV Cumulative incidence	$\text{VE}_{S,CI} = 1 - \frac{\text{CI}_{A1}}{\text{CI}_{A0}}$	$\text{VE}_{IIA,CI} = 1 - \frac{\text{CI}_{A0}}{\text{CI}_{B0}}$	$\text{VE}_{IIB,CI} = 1 - \frac{\text{CI}_{A1}}{\text{CI}_{B0}}$	$\text{VE}_{III,CI} = 1 - \frac{\text{CI}_{A\cdot}}{\text{CI}_{B\cdot}}$

∗ From Halloran, Struchiner, Longini, Am. J. Epidemiol 1997; 146;789–803.

Estimation of household SARs for SARS-CoV-2*

Figure 1. Household Secondary Attack Rates Over Time, by Study Midpoint, in 135 Studies of Unvaccinated Index Cases and Unvaccinated Contacts
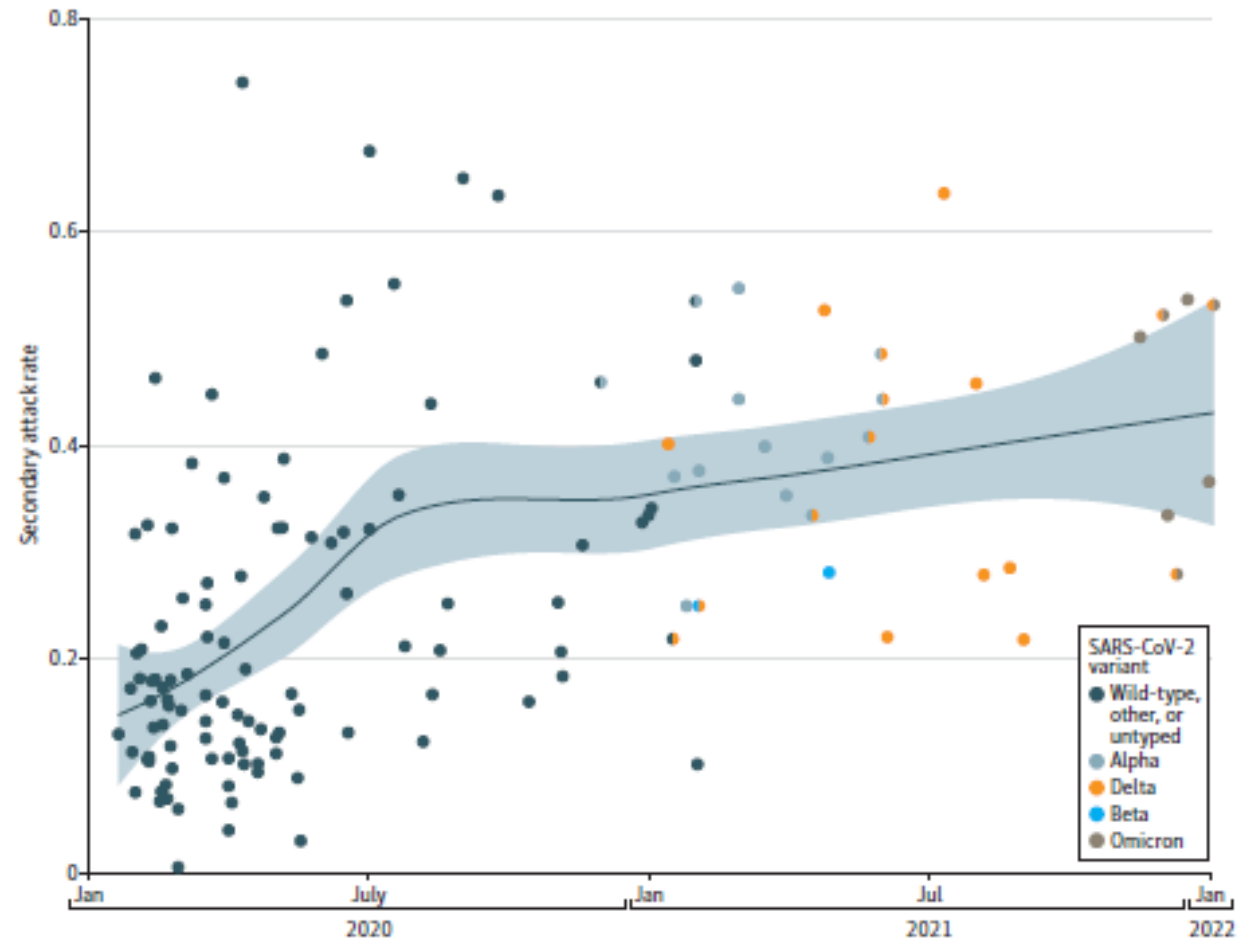
*Madewell ZJ, Yang Y, Longini IM, Halloran ME, Dean NE., Household SARS-CoV-2 secondary attack rates by variant and vaccination status: an updated systematic review and meta-analysis. *JAMA Network Open* **5**(4) (2022):e229317. doi:10.1001/jamanetworkopen.2022.9317.

Estimated VE's from household SAR's

Table. Estimated Vaccine Effectiveness From Household Secondary Attack Rates

Variant and Vaccination Type	Estimated vaccine effectiveness				
	VE$_{I,P}$, % (95% CI)	Studies	VE$_{S,P}$, % (95% CI)	Studies	VE$_{T,P}$, % (95% CI)
All					
Booster vaccination	31.8 (27.1 to 36.2)	Lyngse et al,[62] 2022; Baker et al,[63] 2022; Jalali et al,[64] 2022	49.5 (37.7 to 59.1)	Lyngse et al,[55] 2021; Lyngse et al,[62] 2022; Baker et al,[63] 2022; Jalali et al,[64] 2022	65.4 (55.7 to 74.9)
Full vaccination	44.2 (20.7 to 60.8)	de Gier et al,[19] 2021; Gazit et al,[21] 2021; Layan et al,[28] 2021; Meyer et al,[33] 2021; Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Águila-Mejía et al,[61] 2022; Baker et al,[63] 2022; Jalali et al,[64] 2022; Lyngse et al,[65] 2022; de Gier et al[70]2021	61.4 (45.6 to 72.6)	de Gier et al,[19] 2021; Gazit et al,[21] 2021; Layan et al,[28] 2021; Martínez-Baz et al,[32] 2021; Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Sachdev et al,[50] 2021; Yi et al,[57] 2022; Baker et al,[63] 2022; Jalali et al,[64] 2022; Lyngse et al,[65] 2022; de Gier et al[70]2021	78.5 (64.8 to 86.8)
Partial vaccination	23.6 (−6.0 to 44.9)	de Gier et al,[19] 2021; Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Harris et al,[48] 2021; Baker et al,[63] 2022; Jalali et al,[64] 2022; de Gier et al[70]2021	37.2 (16.4 to 53.0)	de Gier et al,[19] 2021; Martínez-Baz et al,[32] 2021; Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Sachdev et al,[50] 2021; Baker et al,[63] 2022; Jalali et al[64]2022	52.1 (27.7 to 68.8)
Alpha					
Full vaccination	75.3 (69.9 to 79.8)	de Gier et al,[19] 2021; Gazit et al,[21] 2021; Layan et al,[28] 2021; Meyer et al[33]2021	78.6 (76.0 to 80.9)	de Gier et al,[19] 2021; Gazit et al,[21] 2021; Layan et al[28]2021	94.7 (93.3 to 95.8)
Delta					
Booster vaccination	NA	NA	68.0 (62.3 to 72.8)	Lyngse et al[55] 2021; Jalali et al[64]2022	NA
Full vaccination	21.9 (11.0 to 31.5)	Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Águila-Mejía et al,[61]2022; Jalali et al,[64]2022; Lyngse et al,[65] 2022; de Gier et al[70]2021	56.4 (54.6 to 58.1)	Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Yi et al,[57] 2022; Jalali et al,[64] 2022; Lyngse et al,[65] 2022; de Gier et al[70]2021	64.4 (58.0 to 69.8)
Partial vaccination	16.0 (−46.9 to 51.9)	Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Jalali et al,[64] 2022; de Gier et al[70]2021	37.8 (12.0 to 56.0)	Ng et al,[37] 2021; Singanayagam et al,[42] 2022; Yi et al,[57] 2022; Jalali et al[64]2022	51.2 (6.1 to 74.6)

$$VE_T = 1- (1 - VE_S)(1-VE_I)$$

Omicron					
Booster vaccination	32.3 (25.6 to 38.3)	Lyngse et al,[62] 2022; Baker et al,[63] 2022; and Jalali et al[64] 2022	40.8 (35.9 to 45.3)	Lyngse et al,[55] 2021; Lyngse et al,[62] 2022; Baker et al,[63] 2022; and Jalali et al[64] 2022	59.8 (54.7 to 64.5)
Full vaccination	18.2 (0.6 to 32.6)	Águila-Mejia et al,[61] 2022; Baker et al,[63] 2022; Jalali et al[64] 2022	18.1 (−18.3 to 43.3)	Baker et al[63] 2022; Jalali et al[64] 2022	35.8 (13.0 to 52.6)
Partial vaccination	NA	NA	6.9 (−38.0 to 37.2)	Baker et al[63] 2022; Jalali et al[64] 2022	NA

Abbreviations: NA, not applicable because not reported in at least 2 studies; $VE_{I,p}$, vaccine effectiveness for infectiousness based on the transmission probability p; $VE_{S,p}$, vaccine effectiveness for susceptibility; $VE_{T,p}$, total vaccine effectiveness.