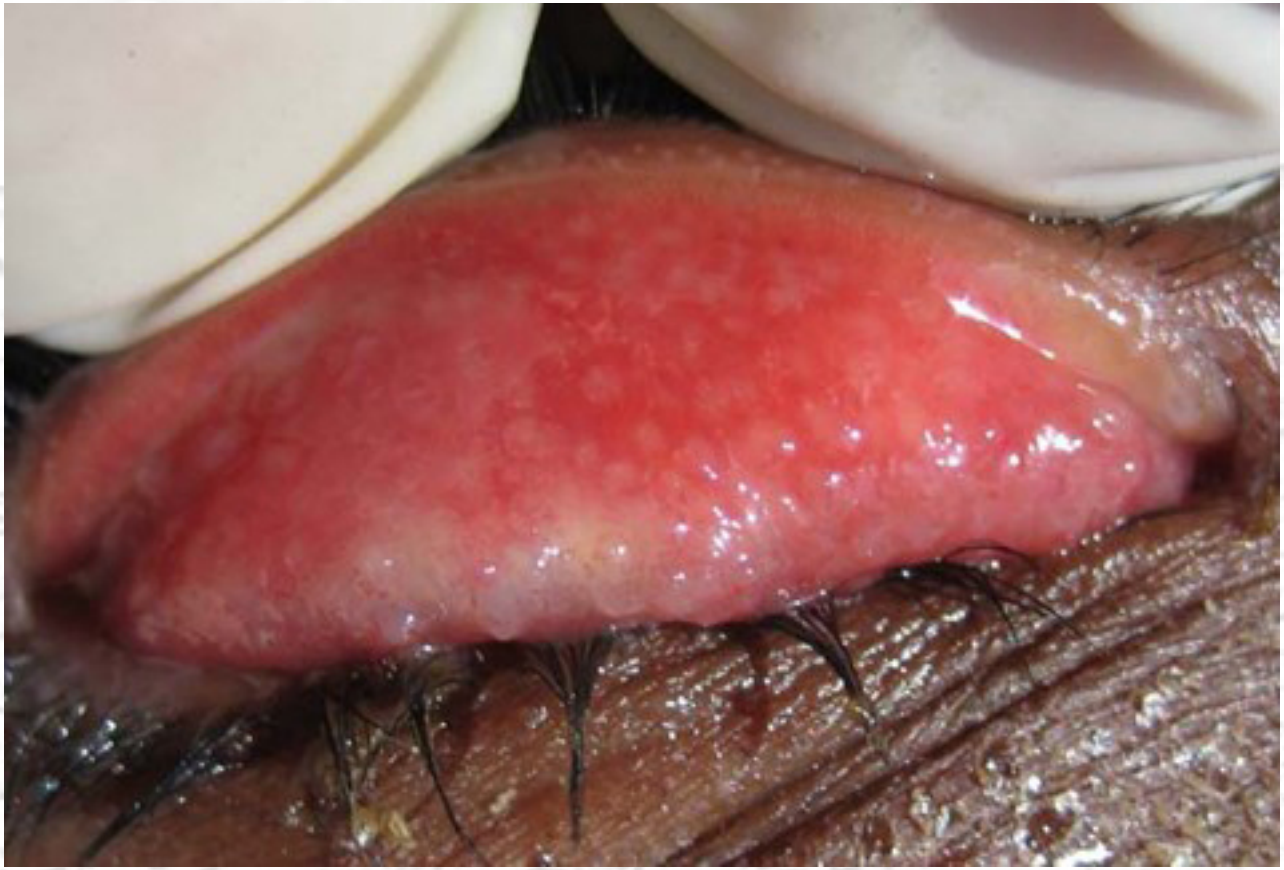
Forecasting trachoma: control, elimination, or eradication?

Thomas M. Lietman
Travis C. Porco
FI Proctor Foundation, UCSF
April 2014

Trachoma

- Causative agent *Chlamydia trachomatis*
- Infection in children leads to blindness later in life
- Slated for elimination according to the London Declaration
- No nonhuman reservoir

Severe TF/TI

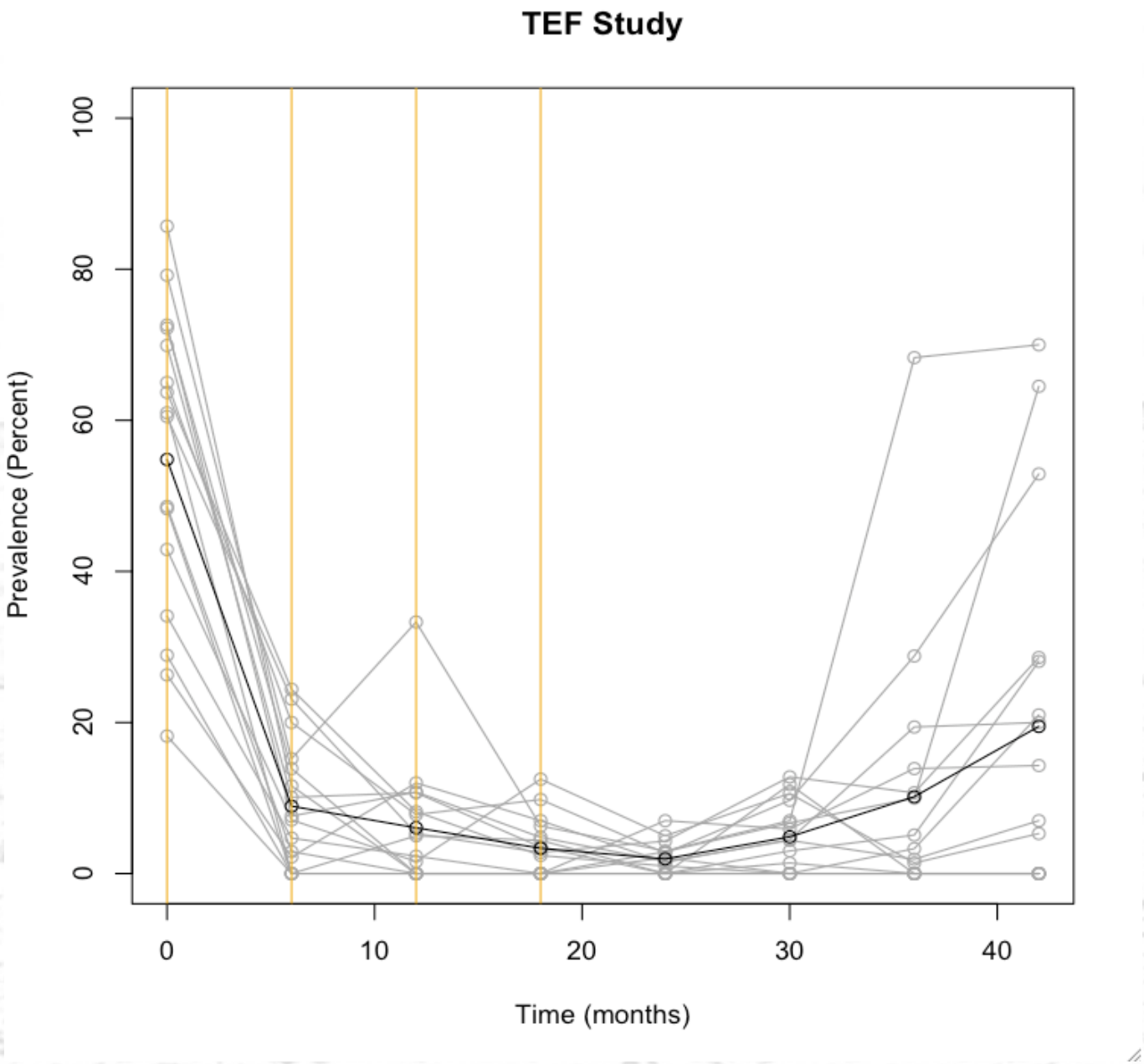


Important facts

- Ocular infection by *C. trachomatis* is easily cured with single-dose azithromycin (80-90% efficacy).
- No vaccine is currently available.
- Clinical signs are unreliable in detecting infection.

Trachoma now

- WHO plan to stamp it out as a public health problem
- Surgery, antibiotics, face-washing, environment
- The SAFE program
- Mass distribution of azithromycin the cornerstone *Schachter J, West SK, Mabey D, et al Lancet. 1999 Aug 21;354(9179):630-5*


TEF Study
Prevalence (Percent)
0
20
40
60
80
100
0
10
20
30
40
Time (months)

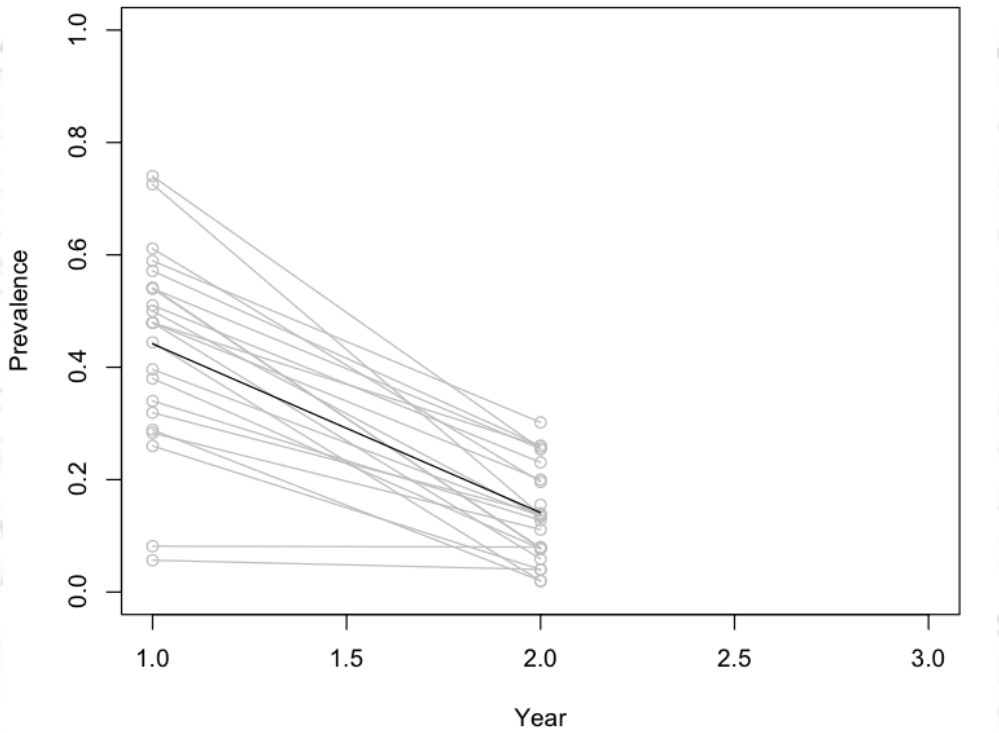
TANA Trial

- TANA data
- Annual treatment, biannual PCR
- Prevalence estimate from 50 children
- Use month 6 to simulate to month 12
- (A mass treatment occurs at month 12)

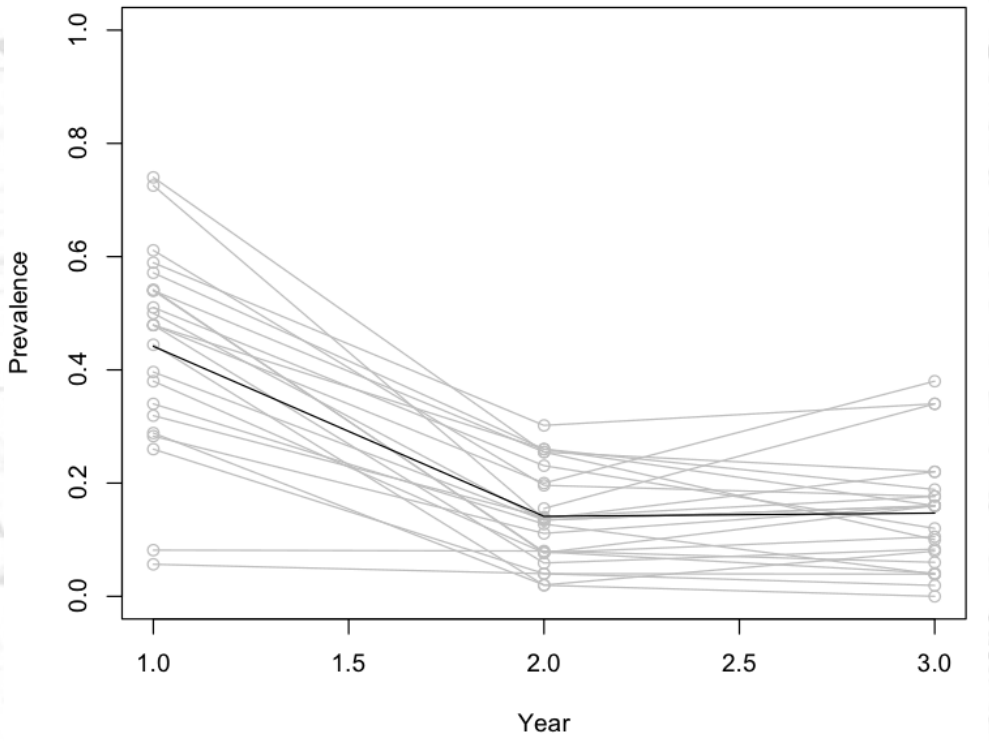
TANA Trial

- Multi-armed clinical trial
- Look first at two arms (24 villages):
  - Baseline MDA at month 0
  - Monitored at 12 and 24 months

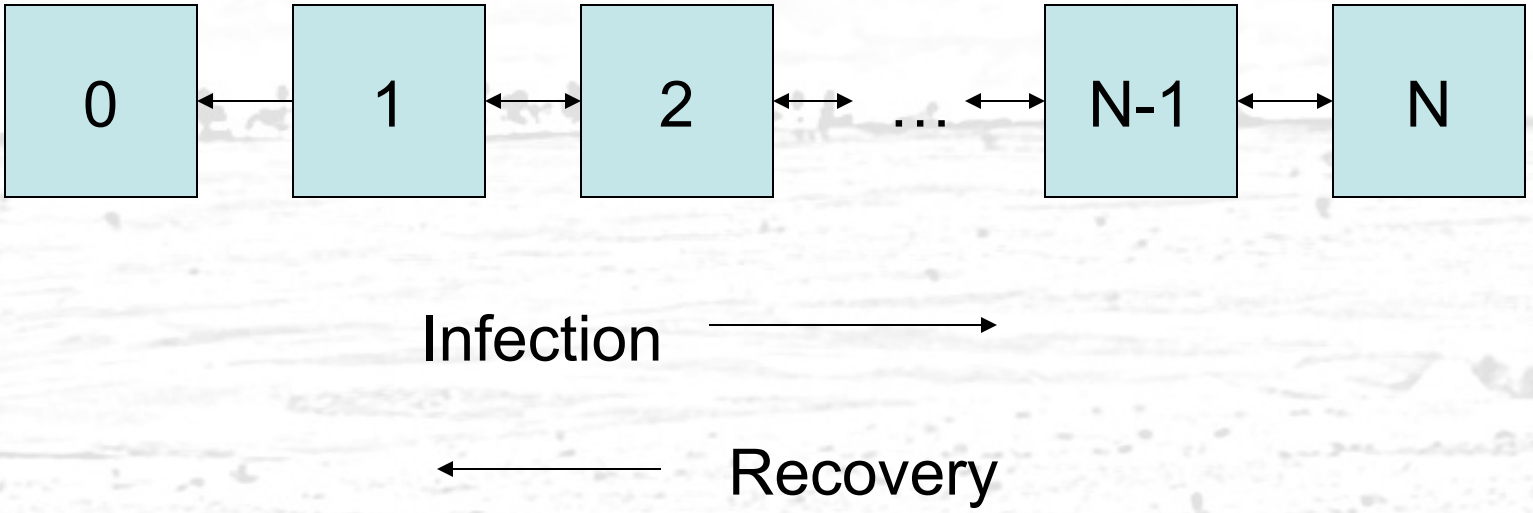
One round of MDA…



Then what?



State space (2)



Standard model

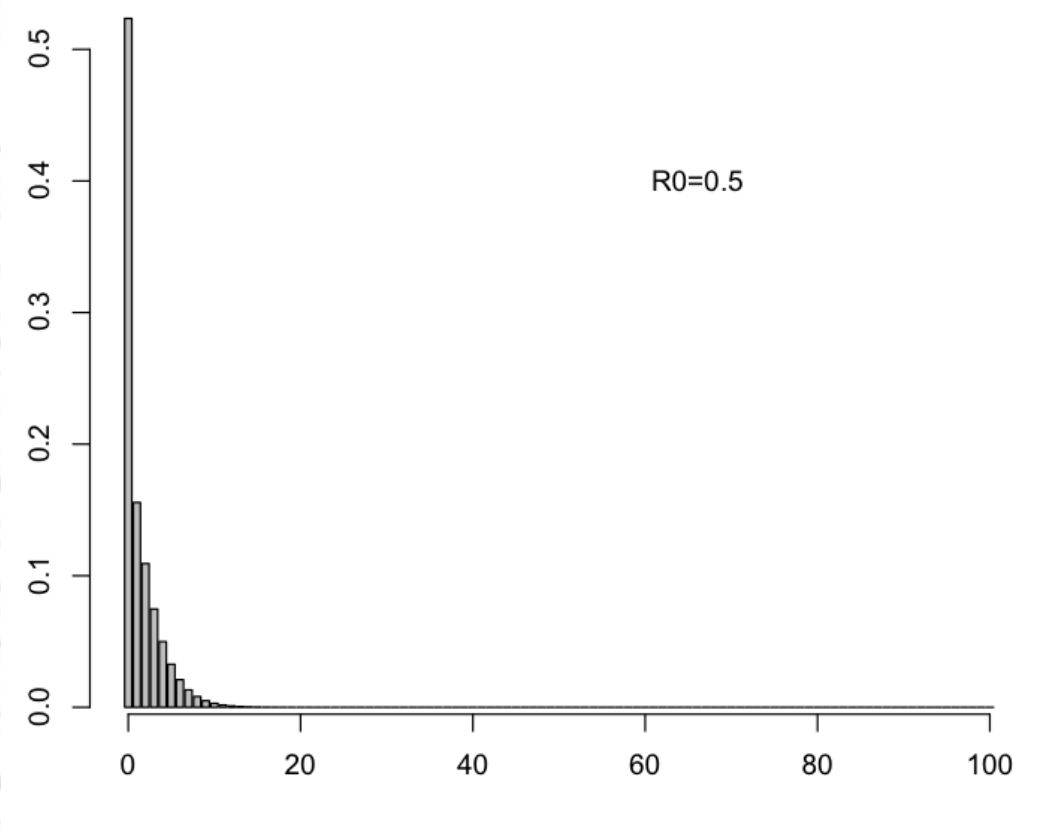
- Stochastic SIS model:

$$\frac{dp_i}{dt} = \left((N-i+1)\beta \left(\frac{i-1}{N-1} \right)^{\alpha} \right) p_{i-1}(t) -$$

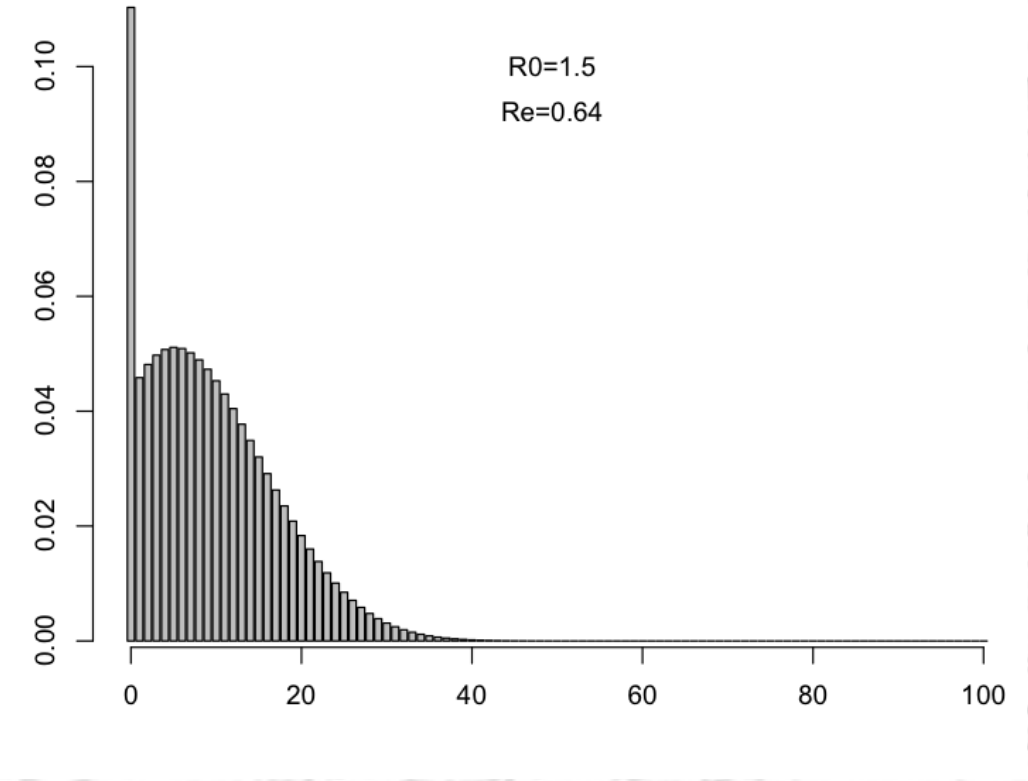
$$\left((N-i)\beta \left(\frac{i}{N-1} \right)^{\alpha} + i\gamma \right) p_i(t) +$$

$$\left((i+1)\gamma \right) p_{i+1}(t)$$

Elimination



Elimination under MDA



Scoring forecasts

- Probabilistic forecast—we produce the probability of the observed data at some time in the future
- Score the forecast by computing the quantity –log(L), where L is the probability of the data

Comparing predictions

- Suppose the previous is the true distribution
- Predictions made from a distribution with R_0=0.4 say are usually better (win 64% of the time, simulation N=10000), though have a lower expected score
- Simulation-based power studies for planning

How well can we do?

- TANA data
- Annual treatment
- Biannual PCR based survey
- Prevalence from 50 children at month 6
- Simulate to month 12
- (A mass treatment occurs at month 12)

Trachoma

- Calibrate on months 6 to 12 and 18 to 24
- Initialize with known results at month 30
- Project to month 36
- Compare with known data

Prediction

- Similarly for month 18 to month 24
- Assume conditional independence given unknown random true coefficient in each village
- Transmission model serves as a simple nonlinear clustered regression model
- Use it to forecast month 36 from month 30 using posterior mode for estimated transmission coefficient

Notes

- Note 1: true forecast score is computed from summing the probability in a sample given the true number infected times the probability of each true number infected
- Note 2: full analysis uses posterior density for village-specific transmission rate, instead of just posterior mode

Assessing forecasts

- Ignorance score (minus loglikelihood)
- -log(P(Y)), where Y is the observation, and P is the probability of the observation as predicted by the model
- Others (proper linear score)

Aside on loglikelihood

- What if the model is true, and I make forecasts from it. What is the expected ignorance score?

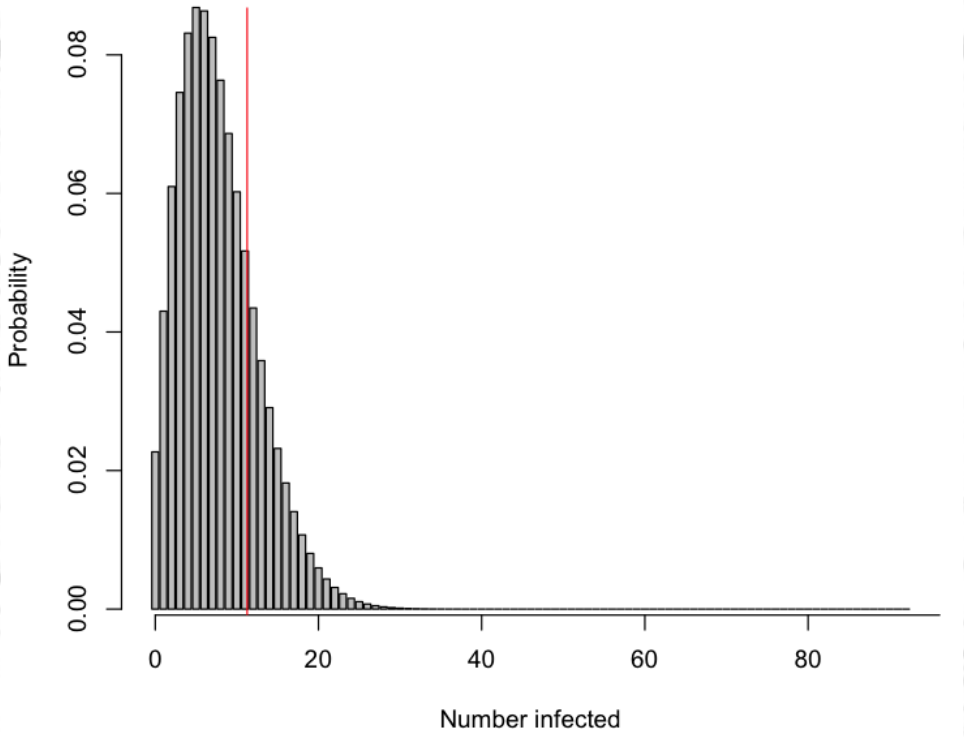
Aside on loglikelihood

- What if the model is true, and I make forecasts from it. What is the expected ignorance score?
- *Shannon entropy*

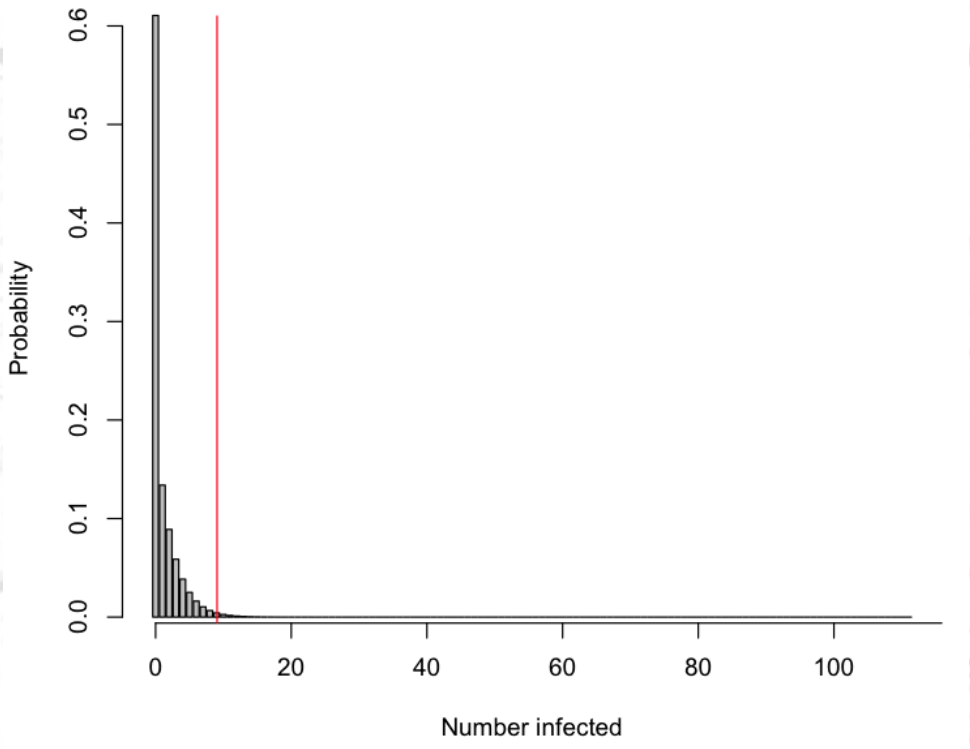
Next, look at 12 forecasts

- 12 villages
- Train on first 2 years
- Test on year 3

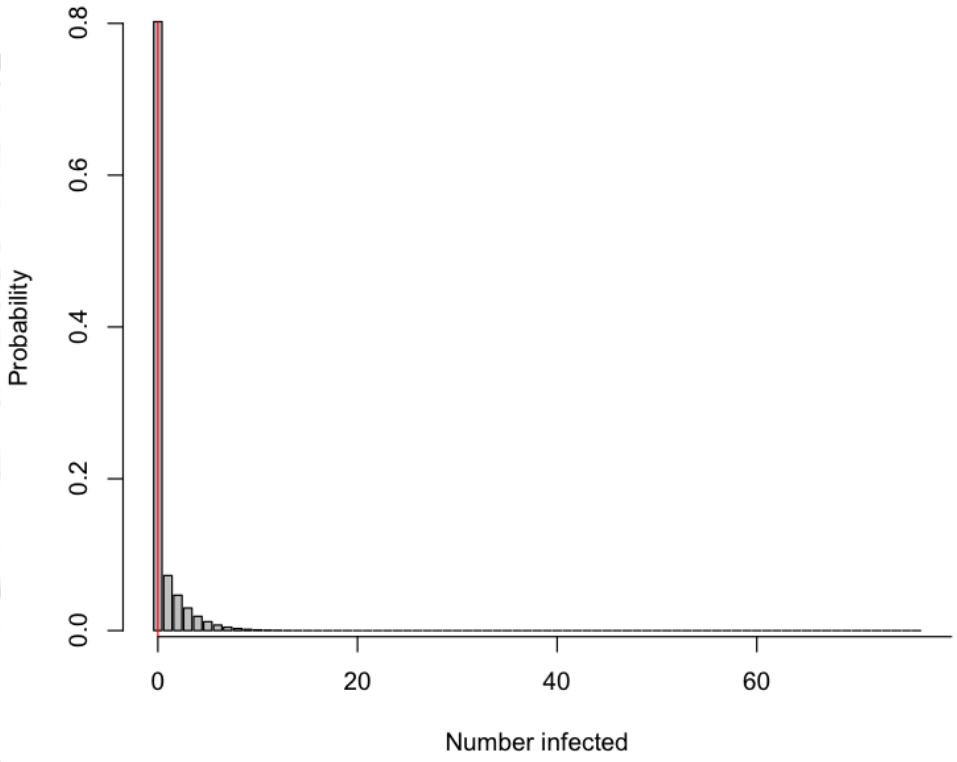
Village 1



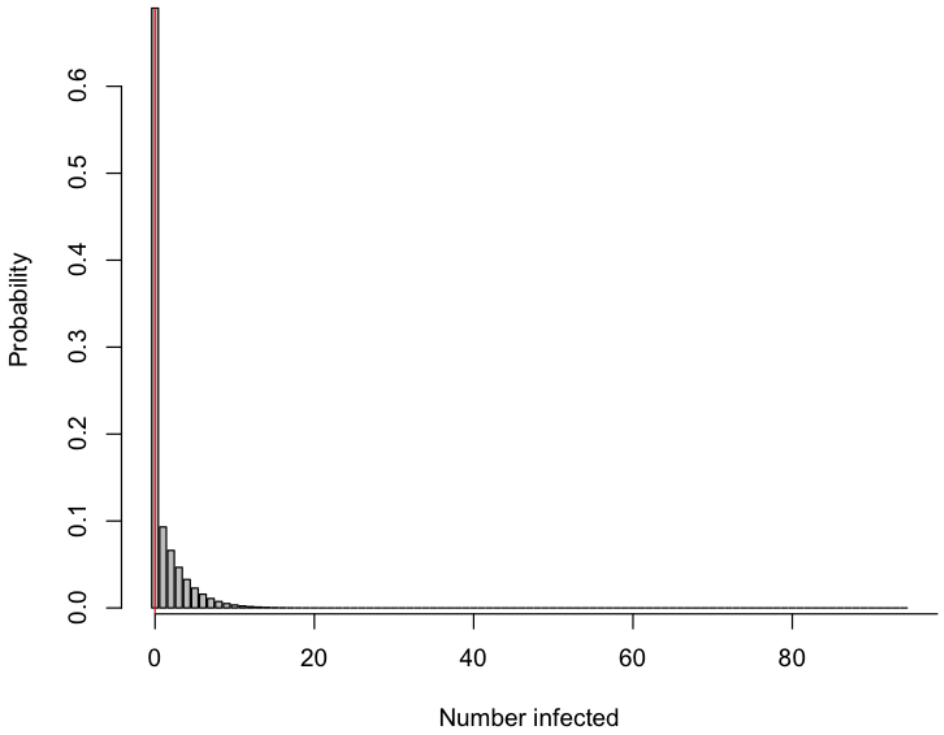
Village 2


Probability
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
20
40
60
80
100
Number infected

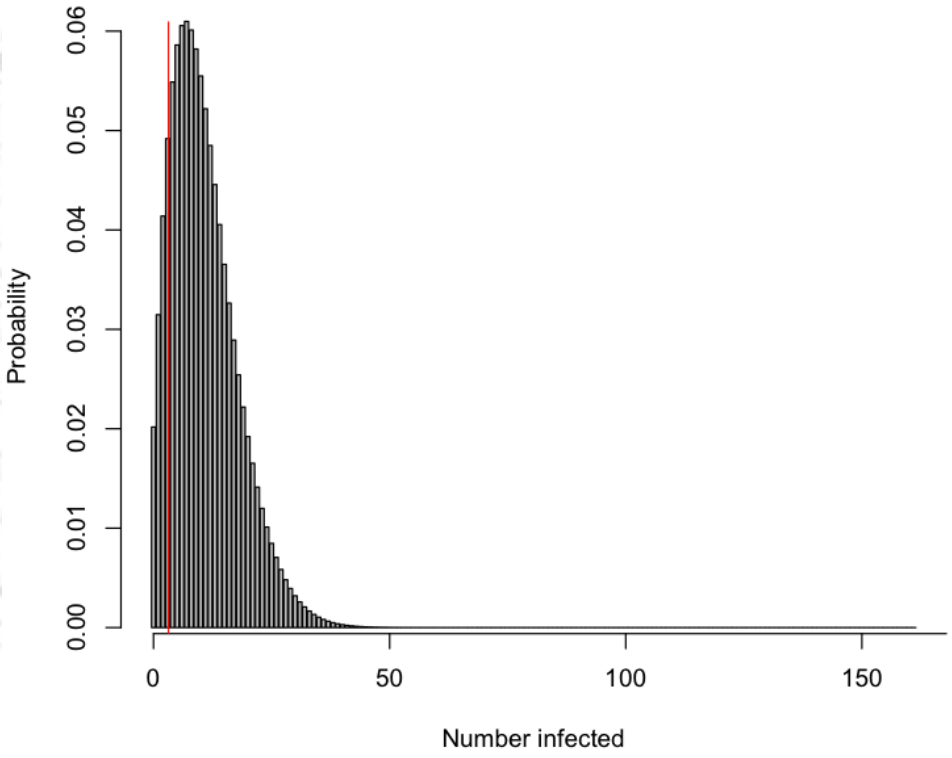
Village 3



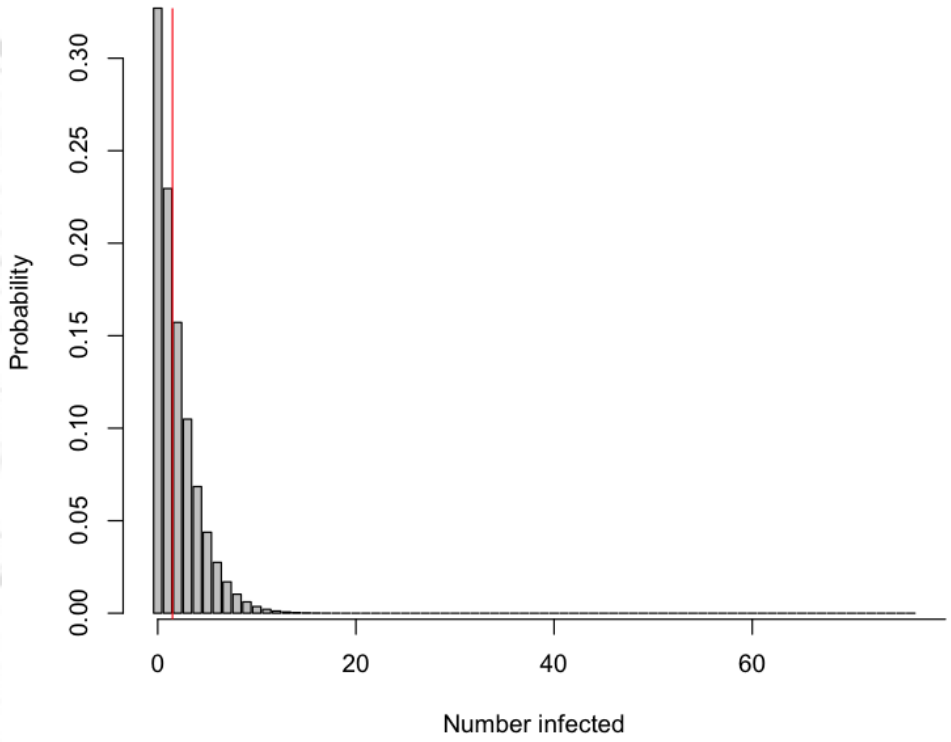
Village 4



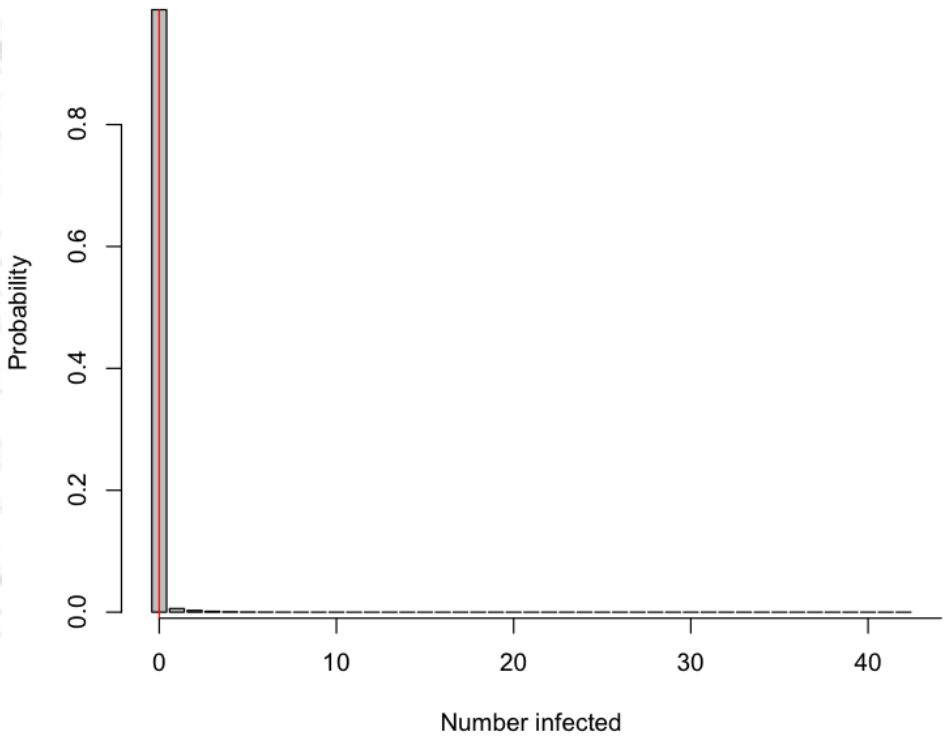
Village 5


Probability
0.00
0.01
0.02
0.03
0.04
0.05
0.06
0
50
100
150
Number infected

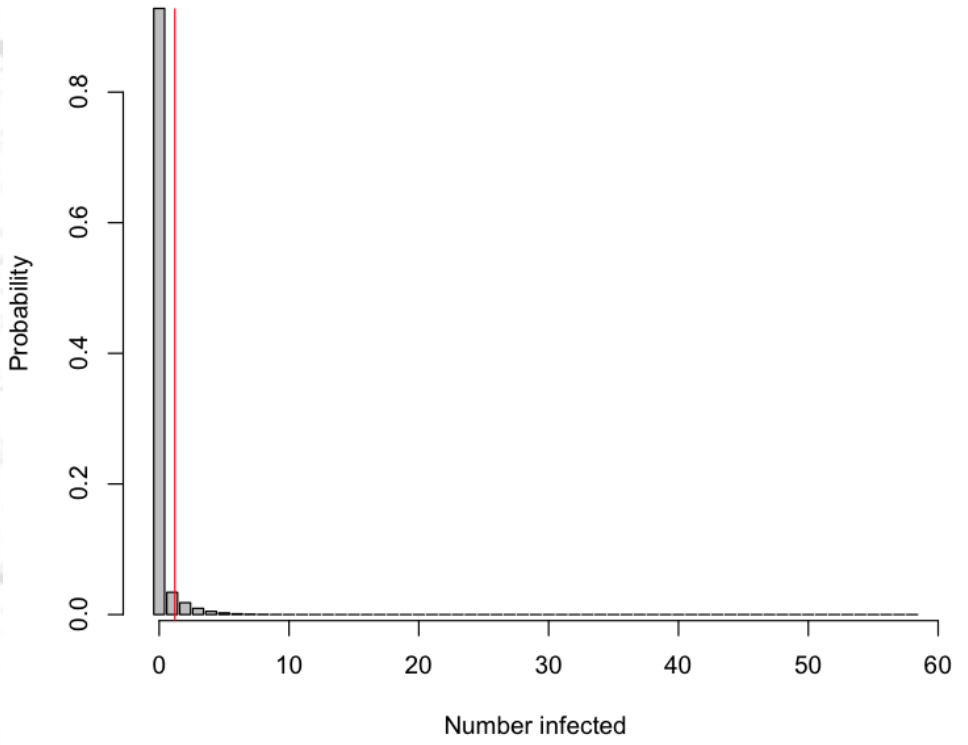
Village 6


Probability
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0
20
40
60
Number infected

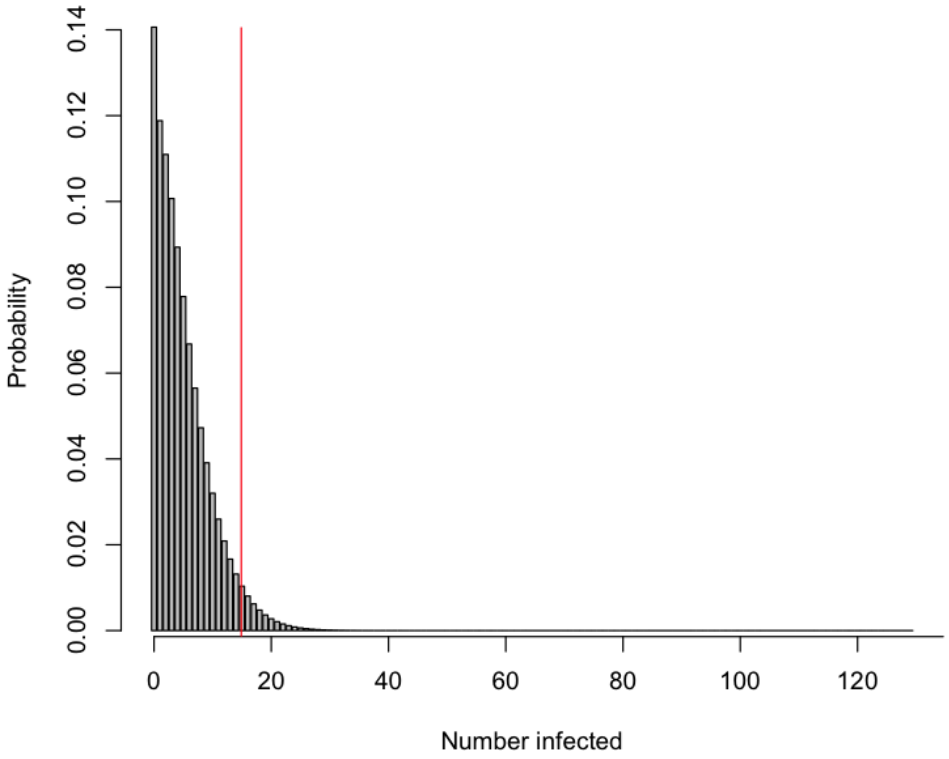
Village 7



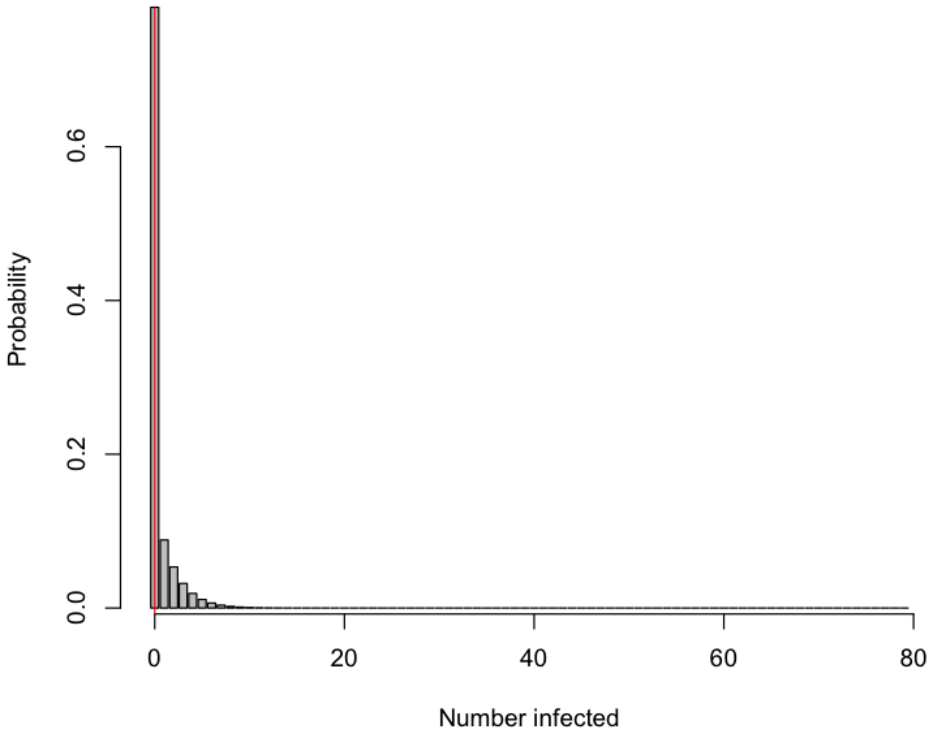
Village 8



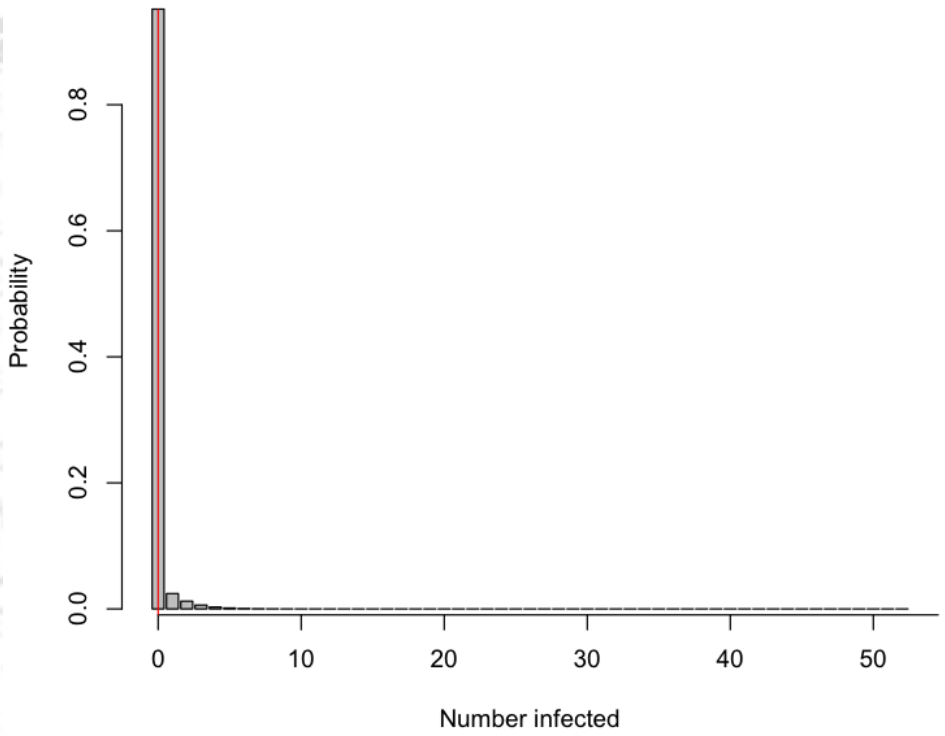
Village 9


Probability
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
0
20
40
60
80
100
120
Number infected

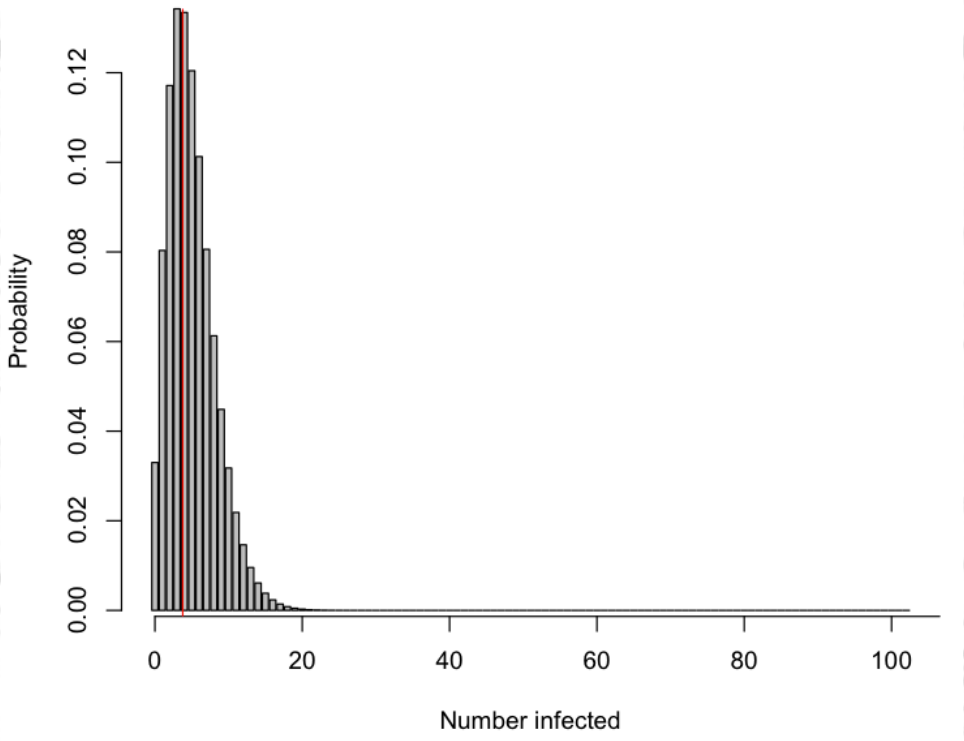
Village 10



Village 11



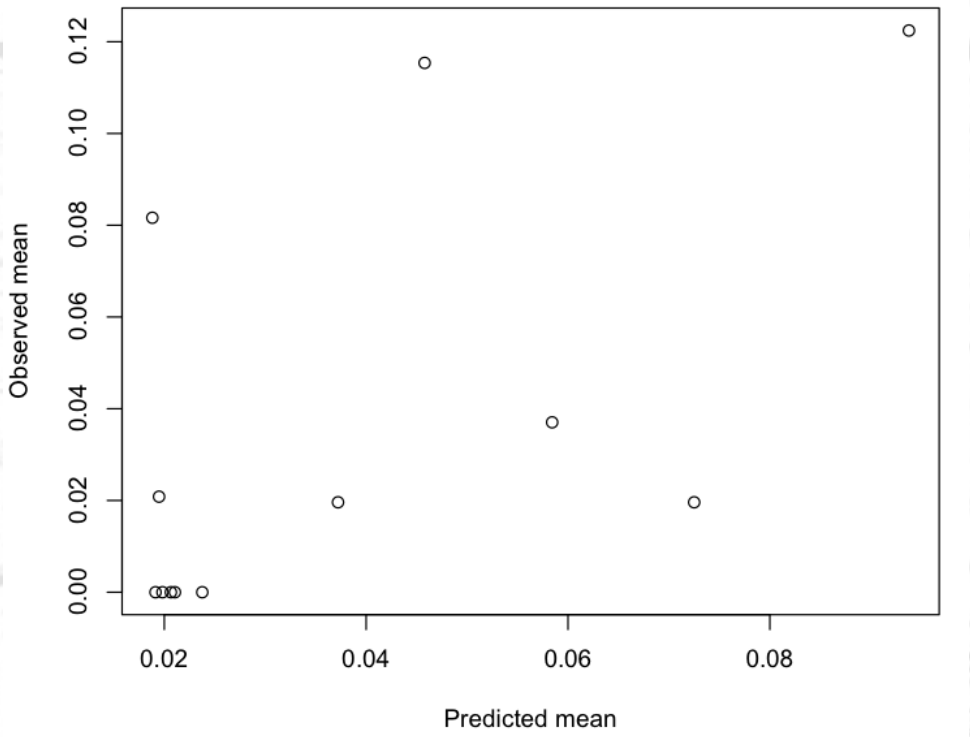
Village 12


Probability
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
20
40
60
80
100
Number infected

Comment

- Much of our apparent forecast skill seems to be keeping the zero-prevalence villages at zero

Summarizing



Trachoma

- If these models are correct, there should be substantial stochastic variability at the small community level, though greater predictability for a group of communities.
- Striking outliers are possible even among theoretically identical villages

What could we expect from this class of models?

- Simple models of this sort give a correlation of about 0.5 between one year and the next
- Observed correlation between baseline and 1 year for the 24 villages was 0.58

Chasing ghosts

- Unpredictability of trachoma at the village level
- Long tail of the distribution
- Expect transient local hot spots
- The presence of a local hot spot does NOT imply failure

Importance

- Expect substantial variability even among identical villages

Future directions

- Include clinical signs into model

$$\frac{dp_{i,j,k}}{dt} = -\left(\gamma j + \rho k + \lambda_j (i+k)\right) p_{i,j,k} +$$

$$\lambda_{j-1}\left(i+1\right) p_{i+1,j-1,k} + \gamma(j+1) p_{i,j+1,k-1} + \lambda_{j-1}(k+1) p_{i,j-1,k+1} + \rho(k+1) p_{i,j-1,k+1}$$

- Include between-village differences (random effects)

Funding

That Man May See
Bernard Osher Foundation
Bodri Foundation
Harper-Inglis Trust
Peierls Foundation
Jack and DeLoris Lange Foundation
Research to Prevent Blindness
International Trachoma Initiative/Pfizer
NIAID: RO1-AI48789, R21-AI55752
NIGMS MIDAS
NEI: U10-EY016214
Bill and Melinda Gates Foundation

With grateful acknowlegment