Evaluating the predictive power of $R_0$ in wildlife populations:
dueling timescales of host movement and disease dynamics

Philip Johnson, UC Berkeley
Paul Cross, USGS
James Lloyd-Smith, Pennsylvania State
Wayne Getz, UC Berkeley
Our question:

- Why does a given disease affect some species much more than other species?

Potential answers:
- Immunological differences
- Behavioral differences
Population structure
Outline

- Why structure matters
  - Phenomenological coupling
  - Mechanistic coupling
- Our model
- Threshold metrics for disease invasion
  - $R_0$
  - $R^*$
- Decision tree prediction of invasion
Why structure matters

Keeling & Rohani (2002): phenomenological coupling is OK when mixing is fast.
Discrete SIR Model

Parameters
- $\beta$ – force of infection
- $\gamma$ – $Pr$(recovery)
But what about the herds?

**Parameters**
- $\beta$ – force of infection
- $\gamma$ – $Pr(\text{recovery})$
- $\mu$ – $Pr(\text{movement})$
- $n$ – group size
Simulation

- Discrete-time, nonhomogenous Markov chain
- Initial state: one infected host in one group
- Transition matrices
  - Spatial: 11x11 torus = 1210 groups = 1210x1210 matrix
  - Disease: S,I,R states – 3x3 matrix where S→I transition depends on I and n
R₀: The basic reproductive #

- How the theoreticians see it:
  - **expected** number of new infections caused by a single index case in an **infinite** population where **everyone** is susceptible.

- How it is often interpreted:
  - extrapolation of R₀ > 1 threshold to real-life **stochastic** contacts, **finite** population, **depletion** of susceptible pool
Why structure really matters

Acute disease
\[ \gamma = 0.1 \]
\[ \beta = 0.5 \]
\[ R_0 = 5 \]

Chronic disease
\[ \gamma = 0.01 \]
\[ \beta = 0.05 \]
\[ R_0 = 5 \]

1 group of 1000
25 groups of 40
100 groups of 10
$R_*: \text{taking } R_0 \text{ to the next level}$

- expected number of new infected \textit{groups} caused by a single initially infected group in a population with infinite susceptible groups
- First coined by Ball \textit{et al.} 1997
- Calculating $R_*$
  - analytic
  - \textit{simulation}
Effect of movement ($\mu$)

$n = 10$

Mean proportion infected

Movement rate ($\mu$) / recovery rate ($\gamma$)

$R_0 = 20$

$R_0 = 10$

$R_0 = 5$

$R_0 = 2$
Effect of group size (n)

Rule of thumb for high $R_0$: $n \cdot \mu / \gamma > 1$
Effect of stochasticity

- $R_0 > 1$ and movement necessary but not sufficient

Proportion of population infected

Number of simulations

- $\mu/\gamma = 0.001$
  - $\beta/\gamma = 5$
  - $n = 10$

- $\mu/\gamma = 0.2$
  - $\beta/\gamma = 5$
  - $n = 10$

- $\mu/\gamma = 1$
  - $\beta/\gamma = 5$
  - $n = 10$
Empirical $R_0$ versus $R_*$

$R_*$ predicts better than $R_0$
A problem: how to predict?

- $R_*$ is great, but cannot be calculated beforehand
- $R_0$ is easy to calculate, but runs into trouble as a predictor

- What to do, what to do?
Classification trees to the rescue!

- Past applications:
  - Clinical risk assessments
  - Growing use in ecological literature
- Goal: minimize misclassification rate of response variable (disease invasion) via binary decision tree
- Heuristic penalty used to avoid huge trees
SIRS & SIR-bd

\[ R_0 \approx \frac{\beta}{\gamma} \]

\[ R_0 \approx \frac{\beta}{\gamma + \delta} \]
Hypothetical classification tree

Has feathers?
  yes
  no

Body weight > 5kg?
  yes
  no

Lives on ground
  yes
    Lives on ground
      yes
        Lives on ground
          10/5
    no
      Flies
        yes
          Flies
            3/9
        no
          Lives on ground
            10/1

Misclassification (error) rate: (5+3+1)/28 = 0.32
CART method

- Aggregate, measurable parameters:
  - $\beta/\gamma$ ($R_0$)
  - $\mu \cdot n/\gamma$ (previous rule of thumb)
  - $\rho \cdot n/\gamma$ or $\delta \cdot n/\gamma$ (influx of new susceptibles)

- Other combinations possible (including raw parameters)
CART results

\[ R_0 \geq 2.0 \]

- \[ \mu n/\gamma \geq 2.7 \]
  - Invasion 2084/557
  - Invasion 257/165
  - Extinction 34/709

- \[ \rho n/\gamma \geq 7.2 \]
  - Extinction 89/2105

Error rates
- \( R_0 > 1 \): 0.35
- CART: 0.14
- Empirical \( R_0 > 1 \): 0.097
Recap

- Model design
  - Population structure critical
  - Mechanistic host movement
- Single predictors
  - $R_0$ gives high false-positives
  - $R_*$ best, but no way to calculate
- CART’s measurable parameters
  - $R_0$, $\mu \cdot n/\gamma$, $\rho \cdot n/\gamma$
Acknowledgements

- Getz Lab

- For more info:
  - Cross, et al. (submitted) *J R Soc Interface*