Real Time Modelling of Epidemics
(A Statistician's Perspective)

Dr. Rob Deardon

Department of Production Animal Health, Faculty of Veterinary Medicine
Department of Mathematics & Statistics, Faculty of Science
University of Calgary

Tuesday 22nd Nov, 2016
(BIRS One Health Disease Modelling Workshop)

Individual-level modelling

- The models I am interested in are infectious disease models that are
  - Mechanistic – account for transmission mechanisms
  - Individual level – account for heterogeneities in the population
    - spatial
    - networks
    - dynamic networks
    - population movement
    - treatment or vaccination status
    - genomics
    - epi-genetics
    - sub-population demographics
    - environmental
    - meterological

Individual-level modelling

- Let us build some example models.

Let the rate of infectivity of individual j at time point t be:

$$\lambda_{jt} = \alpha_0 \sum_{i \in I_t} 1$$

- where:
  - I_t is the set of infectious individuals at time t

Individual-level modelling

- Let us build some example models.

Let the rate of infectivity of individual j at time point t be:

$$\lambda_{jt} = \alpha_0 \sum_{i \in I_t} d_{ij}^{-\beta}$$

- where:
  - I_t is the set of infectious individuals at time t
  - d_{ij} is the spatial distance between i and j

Individual-level modelling

- Let us build some example models.

Let the rate of infectivity of individual j at time point t be:

$$\lambda_{jt} = \alpha_0 \sum_{i \in I_t} (d_{ij}^{-\beta} + \alpha_1 c_{ij})$$

- where:
  - I_t is the set of infectious individuals at time t
  - d_{ij} is the spatial distance between i and j
  - $c_{ij} = 1$ if i and j are connected in a network; 0 otherwise

Individual-level modelling

- Let us build some example models.

Let the rate of infectivity of individual j at time point t be:

$$\lambda_{jt} = (\alpha_0 + \alpha_2 X_j) \sum_{i \in I_t} (d_{ij}^{-\beta} + \alpha_1 c_{ij})$$

- where:
  - I_t is the set of infectious individuals at time t
  - d_{ij} is the spatial distance between i and j
  - $c_{ij} = 1$ if i and j are connected in a network; 0 otherwise
  - X_j is some individual level covariate

What do we want to use these models for?

1. Model choice: Identify significant risk factors for infection
2. Model assessment: Identify problems with models
3. Predict the course of an epidemic
4. Predict at risk individuals / sub-populations
5. Predict the course of an epidemic under various mitigation strategies

6. We want to be able to do all of the above:
    - during an epidemic in real time
    - and with realistic measures of **uncertainty/confidence** about model fit and prediction uncertainty

Solution: Statistical modelling

- Process of developing a mathematical model **directly from data** (e.g. observed epidemic)
- Model fitting methods designed to **extract information from data in optimal ways**
- Model will be as parsimonious as possible

- Aim is to produce a model which **best describes the system**...
- ...but also accounts for **underlying uncertainties** we have about the data and/or system being modelled.

Problems:

- 1. Data
  - Obtaining the data, especially for “one health”, is difficult
    - privacy, commercial sensitivity, etc.
    - multiple agencies and thus multiple bureaucracies

  - Data quality is generally problematic
    - uncertainty about infection times
    - uncertainty about recovery times (and infectious periods)
    - uncertainty about latent periods
    - uncertainty about spatial locations
    - uncertainty about movement patterns
    - uncertainty about networks
    - uncertainty about covariates
    - uncertainty about underreporting

    - Uncertainty here may mean:
      missing, censored or inaccurately measured data

Problems:

- 2. Classical techniques don't really work well (e.g. maximum likelihood and least squares estimation) so we tend to use **Bayesian methods**
    - allow us to naturally incorporate all sorts of uncertainty and information about parameters, data, observation models, disease models

Example results: Infection Time Uncertainty

- Consider simulation infection passing through a network

Let the rate of infectivity of individual j at time point t be:

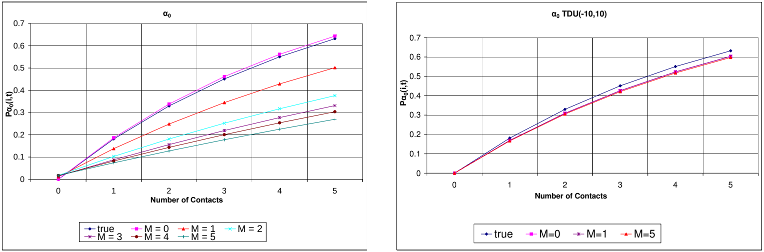
$$\lambda_{jt} = \epsilon + \left(\alpha_0 \sum_{i \in I_t} c_{ij} \right)$$

- ϵ represents infection coming in from outside network
- We simulate an epidemic through the network with $\epsilon = 0$

Example results: Infection Time Uncertainty

- Consider simulation infection passing through a network
- We then take this data and add noise (M) to the recorded infection times
- We then fit the original model back to this data:
    - Case 1: assuming infection times are correctly recorded
    - Case 2: incorporating infection time uncertainty

Daily Infection Probability vs. # of Infectious Contacts



(a) Case1: Fitting model to observed noisy data

(b) Case2: Fitting model accounting for noisy data

Problems:

- 3. Off the shelf software not readily available for fitting our models
- 4. Computation running time may be slow
- 5. Computation coding / implementation time may be slow

Problems:

- Why do we have these computational issues?
  - High dimensional problems (e.g. latent infection times)
  - Algorithms such as data augmented MCMC can be difficult to "tune" and get to run efficiently
  - Models are complex and can change a lot for different disease systems
  - Likelihood function is often cumbersome to compute
  - Need to code in low-level languages such as C/Fortran to get fast performance

Solutions:

- Approximate Bayesian computation (ABC)
- Emulation based inference
- Modelling aggregated data (or sub-sampled data)
- Mathematical tricks (e.g., through linear approximations of nonlinear functions)
- Composite likelihood
- Classification-based inference (machine learning tools)

Solutions:

- Development of stand-alone software
  (e.g. R packages that call Fortran code)
- Currently employing programmer/postdoc to specialize in this

Acknowledgements

- Work funded by:
  - University of Calgary Eyes High Postdoctoral Fellowship Scheme
  - Ontario Ministry of Agriculture, Food & Rural Affairs (OMAFRA)
  - Natural Sciences & Engineering Council of Canada (NSERC)
  - Canadian Foundation for Innovation (CFI)