

# Modelling Healthcare Associated Infections

## Identifying Routes of Transmission

Theo Kypraios

<http://www.maths.nott.ac.uk/~tk>



Loughborough University

## Joint work with:

- **Prof Phil O'Neill** @ University of Nottingham
- **Rosanna Cassidy** @ Maths, University of Nottingham
- **Dr Colin Worby** @ Harvard School of Public Health  
(previously at UoN)
- **Dr Eleni Verykouki** @ University of Thessaly, Greece  
(previously at UoN)
- **Dr Yinghui Wei** @ Plymouth University (previously at UoN)
- **Dr Ben Cooper** @ Mahidol University, Bangkok, Thailand  
(previously at the Public Health England, London)

# Plan

Motivation

Modelling

Statistical Inference and Model Choice

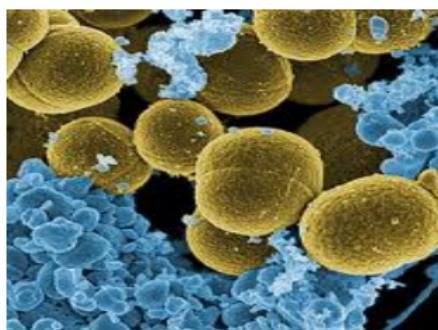
Illustration via Case Studies

Incorporation of High-Resolution Genetic Data

Conclusions

# Healthcare–Associated Infections (HCAIs)

- Caused by a wide variety of common and unusual bacteria, fungi, and viruses during the course of receiving medical care.
- Can be devastating and even deadly.
- Economic burden: cost between \$96 billion and \$147 billion annually [Marchetti and Rossiter, 2013, Journal of Economics]
- Great interest to investigate transmission dynamics, in order to improve infection control strategies.
- Despite recent successes in HCAI elimination, there is much more remains to be done.



# Healthcare-Associated Infections (HCAIs)



1.7 million

people per year get an infection during a hospital stay

98,987

people in the U.S. die annually from HAI

System  
\$35 Billion/yr



9.4% of total inpatient costs are HAI-related

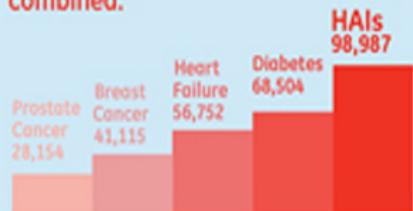


More than ½ of HAI cases affect people with Medicare or Medicaid

Patient  
\$1,100 per admission



HAI kill more people each year than Breast Cancer and Prostate Cancer combined.



# HCAIs: Statistical Modelling

- Statistical modelling framework enables a quantitative assessment of the main obstructions that undermine current efforts to control the spread of HCAIs
- Despite recent successes in HCAI elimination, there is much more remains to be done.
- Collection of high-resolution genetic data is becoming easier and cheaper.
- High-resolution genetic data potentially offers new insights into the dynamics of a hospital disease outbreak.



# Aims

Address a range of **scientific questions** via analyses of **detailed data sets** taken from observational studies on hospital wards.

For instance, we are interested in **answering important questions such as:**

- Do specific **control measures** work?
- What effects do **antimicrobial agents** have on transmission?
- Why do some **strains spread more rapidly** than others?
- How much extra information do **whole genome sequence** data provide?

# Aims

Address a range of **scientific questions** via analyses of **detailed data sets** taken from observational studies on hospital wards.

For instance, we are interested in **answering important questions such as:**

- Do specific **control measures** work?
- What effects do **antimicrobial agents** have on transmission?
- Why do some **strains spread more rapidly** than others?
- How much extra information do **whole genome sequence** data provide?

# Aims

Address a range of **scientific questions** via analyses of **detailed data sets** taken from observational studies on hospital wards.

For instance, we are interested in **answering important questions** such as:

- Do specific **control measures** work?
- What effects do **antimicrobial agents** have on transmission?
- Why do some **strains spread more rapidly** than others?
- How much extra information do **whole genome sequence** data provide?

# Aims

Address a range of **scientific questions** via analyses of **detailed data sets** taken from observational studies on hospital wards.

For instance, we are interested in **answering important questions** such as:

- Do specific **control measures** work?
- What effects do **antimicrobial agents** have on transmission?
- Why do some **strains spread more rapidly** than others?
- How much extra information do **whole genome sequence** data provide?

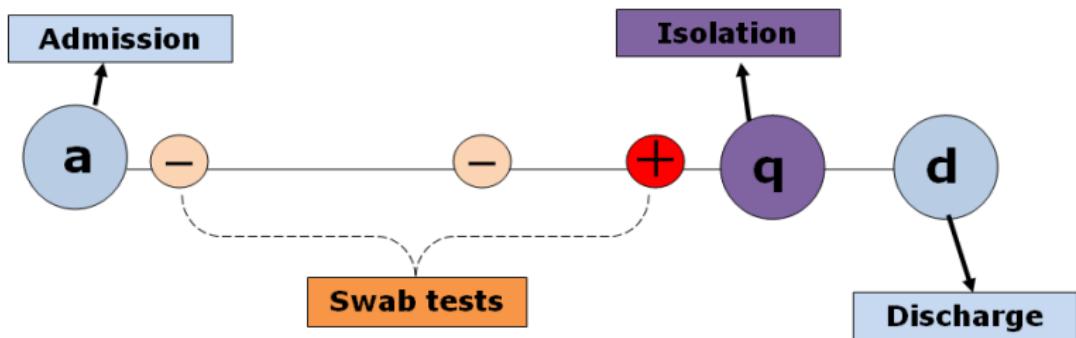
# Aims

Address a range of **scientific questions** via analyses of **detailed data sets** taken from observational studies on hospital wards.

For instance, we are interested in **answering important questions** such as:

- Do specific **control measures** work?
- What effects do **antimicrobial agents** have on transmission?
- Why do some **strains spread more rapidly** than others?
- How much extra information do **whole genome sequence** data provide?

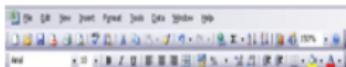
# Timeline: What happens to a patient?



# Typical Data sets

Typical data sets contain **anonymised ward - level** information on:

- Dates of patient **admission** and **discharge**.
- Dates/Outcomes of **swab tests**.
- Patient **location** (e.g. in isolation).
- Details of **antibiotics** administered.
- Whole genome sequence data.**



A	B	C	D	E	F	G	
1	Study_ID	SEX	age	med_surg	relative_precaution_date	relative_ICU_DT_IN	relative_ICU_DT_OUT
2	BM100	M	>=85	med		16350	16356
3	BM111	M	>=85	sur		16002	16007
4	BM112	M	65-<75	med		16004	16006
5	BM112	M	65-<75	sur		16006	16007
6	BM113	M	>=85	sur		16003	16007
7	BM113	M	>=85	sur		16007	16017
8	BM113	M	>=85	sur		16017	16026
9	BM113	M	>=85	med		16026	16042
10	BM114	F	75-<85	med		16081	16082
11	BM115	M	45-<55	med		16143	16144
12	BM116	F	35-<45	med		16148	16148
13	BM117	M	>=85	med		16346	16346
14	BM119	M	75-<85	med		16497	16498
15	BM119	M	75-<85	sur		16498	16498
16	BM12	F	45-<55	sur	15911	15940	16082
17	BM12	F	45-<55	sur	15911	16082	16180
18	BM120	F	65-<75	med		16003	16006
19	BM121	M	75-<85	med		16008	16009
20	BM124	M	65-<75	med	15717	16157	16158
21	BM127	M	35-<45	sur		15996	16006
22	BM128	M	65-<75	sur		16036	16036
23	BM129	F	55-<65	med		16388	16389
24	BM13	M	65-<75	sur	15992	15973	16016
25	BM139	F	75-<85	med		16000	16027
26	BM140	F	18-<35	med		16045	16046

# Principles

We use **stochastic models** to describe the (indirect) transmission of the pathogen between individuals.

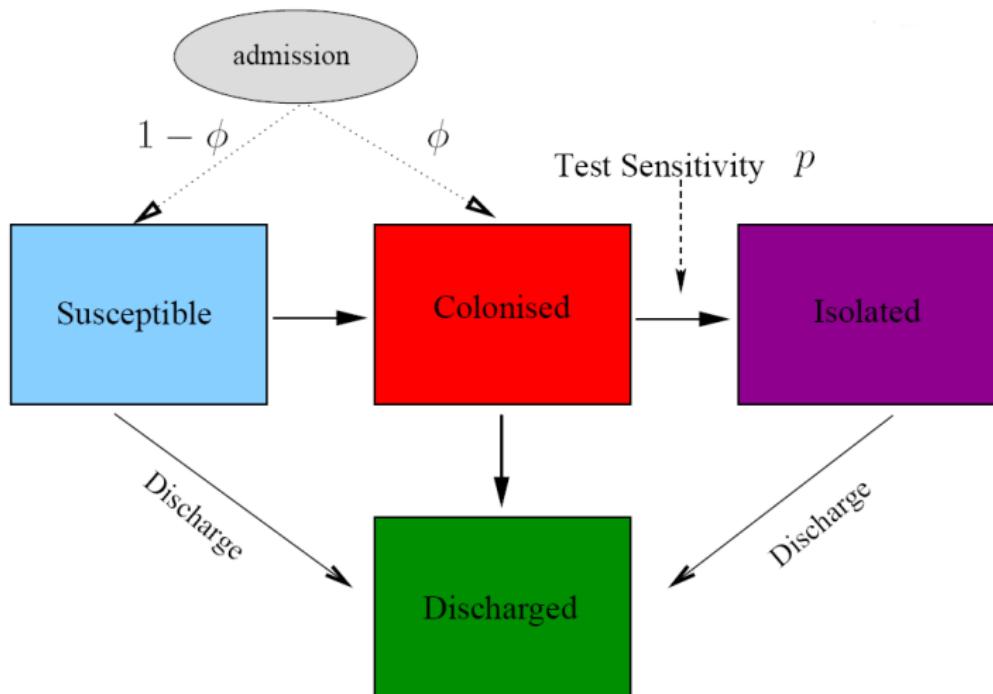
Such models are **useful tools** because they

- are **biologically meaningful**;
- are appropriate for the **questions of interest**;
- can **avoid unrealistic assumptions** of “standard” medical statistics methods.

**Stochasticity is vital** in this context because the number of individuals in a hospital ward or intensive care unit is **usually fairly small**.

# Modelling

# Schematic Representation of a “Baseline Model”



## Model Dynamics: Stochastic

- While susceptible an individual receives indirect colonisation pressure from each **colonised and non-isolated** (**colonised and isolated**) according to a **homogeneous Poisson process** with intensity  $\beta_1$  ( $\beta_2$ ).
- We also allow for **background transmission**, i.e. an individual receives colonisation pressure from outside the ward according to **homogeneous Poisson process** with intensity  $\beta_0$ .

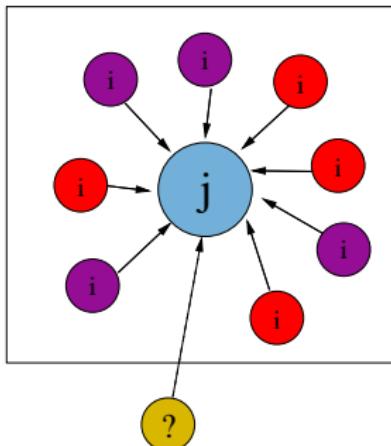
**Hint:** If  $\beta_1 > \beta_2$  may indicate that isolation is somewhat effective

# Assessing Effectiveness of Control Measures

The **total pressure** that susceptible individual  $j$  is subject to **just prior to their colonisation** is:

$$\lambda_j(t) = \beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t)$$

where  $n_C$  is number of **colonised** individuals on ward,  $n_I$  is number of **isolated** individuals on ward.



Note that this assumes **linear** colonisation pressure.

# Screening Tests

- Taken at **specific times** for every single patient
  - Protocol: If **positive** then the patient becomes **isolated**.
- This routine swabbing procedure may be subject to **imperfect sensitivity**, i.e. some false negative swabs are possible.
  - Therefore, we assume that the **sensitivity** of this swabbing procedure is denoted by  $p$ .
- **100% specificity** is assumed → there are no false positives. It can be easily relaxed.

# Inference

## Statistical Inference: Methodology

In order to estimate the model parameters, we need to construct a likelihood - i.e. a function (of the parameters) that tells us how likely the observed data are to occur.

$$L(\beta_0, \beta_1, \beta_2, \phi, p) = Pr(\text{ test results} | \beta_0, \beta_1, \beta_2, \phi, p)$$

However in our setting, such a likelihood is usually intractable due to the fact that key events are unobserved.

## Statistical Inference: Methodology (cont.)

For example, suppose we observe an individual who tests negative on admission (day 0), and after 7 days.

This could arise because:

1. Individual was not colonised and tests correct;
2. Individual was not colonised on admission, became colonised on the ward: first test correct, second incorrect;
3. Individual was colonised on admission and both tests wrong.

The probability of 1. happening is

$$(1 - \phi) \cdot Pr(\text{ Avoids colonisation for 7 days })$$

## Statistical Inference: Methodology (cont.)

$$\Pr(\text{individual colonised in } [t+, t + dt]) | n_C(t), n_Q(t)) = \\ (\beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t))dt + o(dt)$$

So, the avoidance probability is

$$\Pr(\text{Avoids colon. in 7 days}) = \\ \exp \left\{ - \int (\beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t))dt \right\}$$

**Problem** now is that  $n_C(t), n_I(t)$  are **unknown**, and their joint probability distribution is **hard to calculate**.

However, **if** the unobserved events (such as colonisation times, or true admission status of each individual) **were known** then the likelihood **becomes tractable**.

## Statistical Inference: Methodology (cont.)

- A standard way to proceed is then to **treat the “missing data” as extra model parameters** which can be estimated.
- This approach is **especially natural** in the context of **data augmentation** within a **Bayesian framework**.
- **Inference then is feasible** using state-of-the-art Markov Chain Monte Carlo algorithms {see, e.g., Forrester, Pettit & Gibson (2007), K, O'Neill, Huang, Rifas-Shiman and Cooper (2010)}

# Case Studies

# DataSet 1

Data on colonisation were collected from 8 adult intensive care units over a 17-month period.

- 10-bed ICUs in a tertiary academic medical center.
- Routine admission and weekly bilateral nares screening for MRSA (compliance 90%).
- Types of ICUs including:
  - medical,
  - cardiac,
  - general/cardiac/thoracic surgery,
  - burn trauma,
  - neurosurgery.
- Regular swabbing was carried out.

## Summary Statistics

---

Ward	Number of Admissions	Mean (SD) Length of Stay in Days	Percentage in contact precautions
M1	1293	3.4 (4.7)	11.4
M2	1018	4.4 (6.4)	19.1
GS1	1227	3.4 (5.2)	12.4
GS2	1030	4.0 (8.3)	10.7
SS1	706	5.8 (11.4)	12.5
SS2	888	4.8 (9.7)	7.5
SS3	1097	3.8 (6.4)	6.0
SS4	1263	3.6 (5.2)	5.1

# Dataset 1 (cont.)

## What is known?

- Newly-identified and previously known MRSA-positive patients were placed into contact precautions such as gown and glove use as well as use of single rooms.
- Dates of each ICU admission and discharge were obtained.
- Dates and outcomes of test results (if any).

## What is (usually) not known?

- If the patient was colonised on admission.
- When the patient became colonised (if ever)?
- How sensitive the swab test was?
- Which apparently uncolonised patients were colonised?

# Q1: Are Contact Precautions Effective?

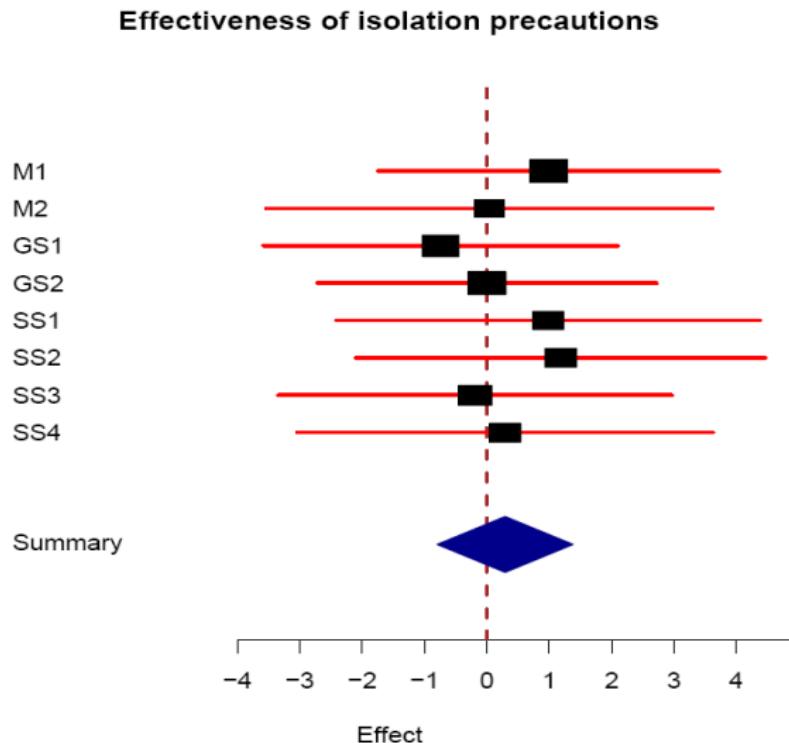
Some ways to measure this include calculating

- $\mathbb{P}(\beta_1 > \beta_2 | \text{data})$ , or
- considering the ratio  $\beta_1 / \beta_2$ .

Ward	$\mathbb{P}(\beta_1 > \beta_2   \mathbf{y})$	Median( $\beta_1 / \beta_2$ )
M1	0.82	2.7
M2	0.51	1.0
GS1	0.27	0.5
GS2	0.50	1.0
SS1	0.73	2.7
SS2	0.79	3.3
SS3	0.44	0.8
SS4	0.58	1.3

# Summarising the Results

By borrowing techniques from **Meta–Analysis** we can derive a *pooled estimate* for the  $\log(\beta_1/\beta_2)$ :



## Q1 (cont.) Are the findings model-dependent?

- Contact precautions appear to be effective - but what happens if we used a different model?
- We instead consider a simpler model in which
  - the colonisation pressure received by a susceptible individual does not increase with the number of colonised individuals.
- Specifically, the total pressure that susceptible individual  $j$  is subject to just prior to their colonisation is:

$$\lambda(t) = \beta_0 + \beta_1 \mathbb{1}_{\{n_C(t) \geq 1\}} + \beta_2 \mathbb{1}_{\{n_I(t) \geq 1\}},$$

- Another way of changing the model assumptions is via the choice of parameter prior distributions – informative prior distribution on  $\beta_0$ ?

## Q1 (cont.) Are the findings model-dependent?

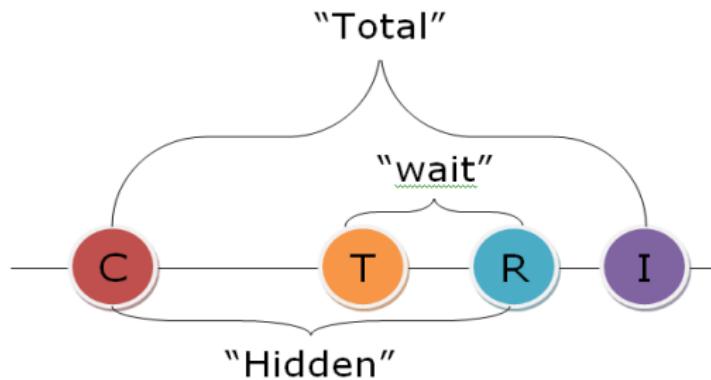
Ward	$\mathbb{P}(\beta_1 > \beta_2   \mathbf{y})$	Median( $\beta_1 / \beta_2$ )
M1	0.82 (0.75) [0.83]	2.7 (2.3) [2.4]
M2	0.51 (0.45) [0.54]	1.0 (0.8) [1.1]
GS1	0.27 (0.36) [0.15]	0.5 (0.6) [0.3]
GS2	0.50 (0.62) [0.58]	1.0 (1.5) [1.2]
SS1	0.73 (0.53) [0.57]	2.7 (1.1) [1.3]
SS2	0.79 (0.71) [0.79]	3.3 (2.0) [2.0]
SS3	0.44 (0.67) [0.60]	0.8 (2.0) [1.3]
SS4	0.58 (0.59) [0.70]	1.3 (1.4) [2.3]

where

- results in  $(\cdot, \cdot)$  refer to the model assuming **non-linear colonisation pressure**,
- results in  $[\cdot, \cdot]$  refer to the model assuming that the **background pressure if zero**.

## Q2: Undetected Cases and Test Delays?

- Our methodology enable us to assess:
  - how much transmission is due to patients who are colonised but not yet detected and
  - how much transmission is due to patients who are colonised and have been tested, but who are awaiting results.
- Define 1 CPD to be one Colonised-Patient-Day, i.e. each colonised patient contributes one unit of CPD for each day they remain colonised.



## Undetected cases and test delays (cont.)

Table:  $p_{hidden}$  and  $p_{wait}$  fitting the “standard” model

Ward	$p_{hidden}$	$p_{wait}$
M1	16.5 (14.9, 18.2)	11.5 (10.6, 12.4)
M2	10.5 (8.3, 13.2)	6.7 (5.6, 7.9)
GS1	13.8 (12.2, 15.5)	8.3 (7.6, 8.9)
GS2	17.3 (15.1, 19.8)	7.4 (6.6, 8.2)
SS1	9.6 (8.1, 11.2)	4.7 (4.4, 4.9)
SS2	10.7 (9.0, 12.7)	5.7 (5.4, 6.1)
SS3	15.6 (13.0, 18.4)	7.9 (6.9, 8.7)
SS4	19.8 (16.2, 23.3)	10.1 (8.7, 11.4)

So, roughly speaking:

- about 10% - 15% of patient-colonised days are undetected
- about 10% of patient-colonised days occur due to delays in obtaining test results.

**Important:** Robust results to the different choice of models.

## Q3: Effect of Colonisation Pressure

- Such a question can be addressed by comparing different models (e.g. one with colonisation rate linear, one with an alternative) and seeing which is most likely under the data.
  - Model 0:  $\lambda(t) = \beta_0$
  - Model 1:  $\lambda(t) = \beta_0 + \beta_1 \mathbb{1}_{\{n_C(t) \geq 1\}} + \beta_2 \mathbb{1}_{\{n_I(t) \geq 1\}}$
  - Model 2:  $\lambda(t) = \beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t)$
- Bayesian Model Choice.
  - Posterior Model Probabilities - Bayes Factors
  - Within-Model prior distributions and Lindley's paradox : Prior's Matching & Prior Sensitivity
  - Trans-dimensional MCMC algorithm

## Q3: Effect of Colonisation Pressure

- Such a question can be addressed by comparing different models (e.g. one with colonisation rate linear, one with an alternative) and seeing which is most likely under the data.
  - Model 0:  $\lambda(t) = \beta_0$
  - Model 1:  $\lambda(t) = \beta_0 + \beta_1 \mathbb{1}_{\{n_C(t) \geq 1\}} + \beta_2 \mathbb{1}_{\{n_I(t) \geq 1\}}$
  - Model 2:  $\lambda(t) = \beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t)$
- Bayesian Model Choice.
  - Posterior Model Probabilities - Bayes Factors
  - Within-Model prior distributions and Lindley's paradox : Prior's Matching & Prior Sensitivity
  - Trans-dimensional MCMC algorithm

## Q3: Effect of Colonisation Pressure

- Such a question can be addressed by comparing different models (e.g. one with colonisation rate linear, one with an alternative) and seeing which is most likely under the data.
  - Model 0:  $\lambda(t) = \beta_0$
  - Model 1:  $\lambda(t) = \beta_0 + \beta_1 \mathbb{1}_{\{n_C(t) \geq 1\}} + \beta_2 \mathbb{1}_{\{n_I(t) \geq 1\}}$
  - Model 2:  $\lambda(t) = \beta_0 + \beta_1 n_C(t) + \beta_2 n_I(t)$
- Bayesian Model Choice.
  - Posterior Model Probabilities - Bayes Factors
  - Within-Model prior distributions and Lindley's paradox : Prior's Matching & Prior Sensitivity
  - Trans-dimensional MCMC algorithm

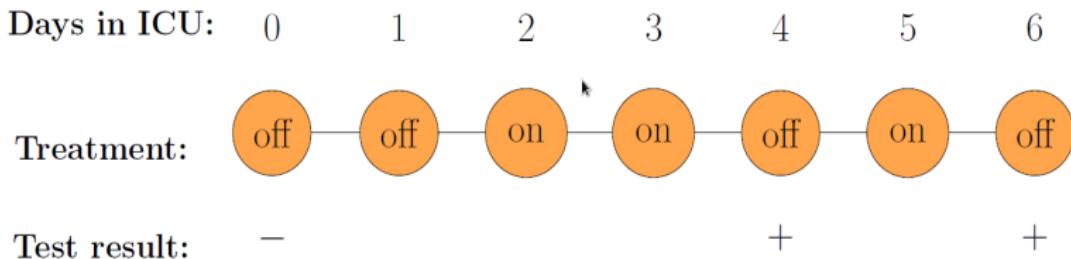
### Q3 (cont): Bayesian Model Choice

Ward	$\lambda = 10^{-3}$			$\lambda = 10^{-2}$			$\lambda = 10^{-1}$		
	$M_0$	$M_1$	$M_2$	$M_0$	$M_1$	$M_2$	$M_0$	$M_1$	$M_2$
M1	0.25	0.32	0.43	0.43	0.23	0.34	0.33	0.31	0.36
M2	0.34	0.39	0.27	0.57	0.22	0.21	0.57	0.22	0.21
GS1	0.16	0.41	0.43	0.23	0.33	0.44	0.23	0.33	0.44
GS2	0.62	0.18	0.20	0.82	0.06	0.12	0.82	0.06	0.12
SS1	0.84	0.12	0.04	0.99	0.01	0.00	0.98	0.02	0.00
SS2	0.79	0.18	0.03	0.97	0.02	0.01	0.97	0.02	0.01
SS3	0.41	0.32	0.27	0.62	0.21	0.17	0.52	0.28	0.21
SS4	0.76	0.19	0.05	0.73	0.19	0.08	0.80	0.15	0.05

- Results do not suggest much support for the full model.
- However, closer scrutiny reveals that, typically,  $n_C(t)$  and  $n_I(t)$  are on average 0, 1, or 2.

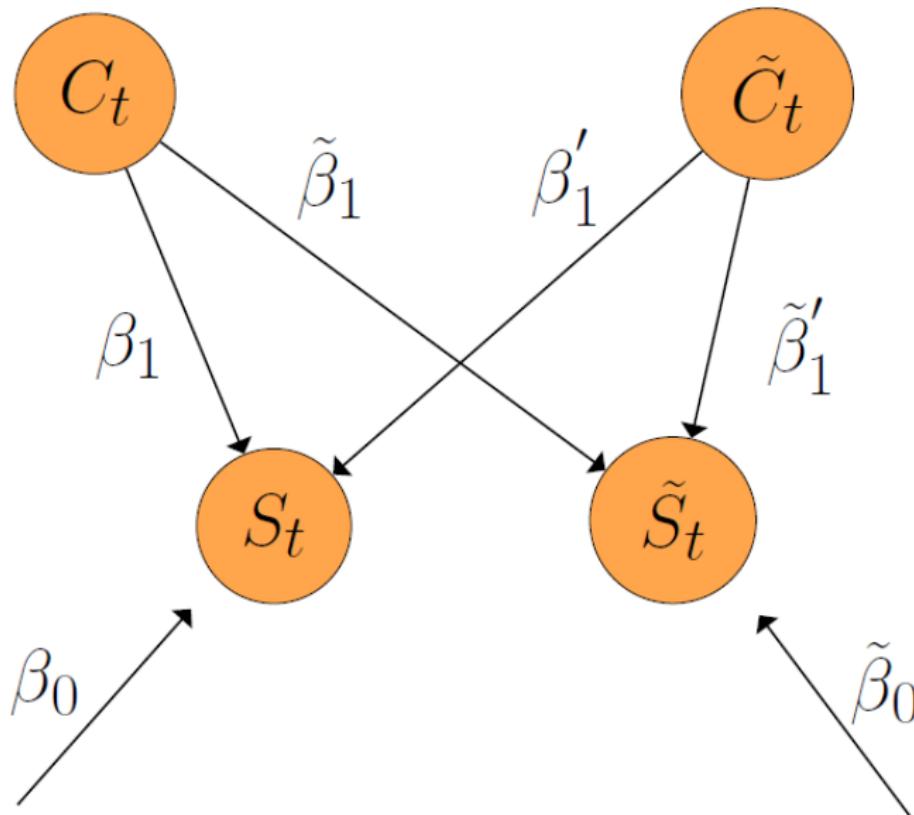
## Another Dataset: Dataset 2

- Q5: **What Effect Different Antibiotics Have?**
- Data from Guy's and St Thomas' hospital on patient's MRSA carriage levels:  $\{-, +, ++\}$  (2 different wards)
- Four-year study, 4,570 ICU admissions.



- **Note:** An individual can be on more than one antimicrobials per day

# A Transmission Model Taking Into Antibiotic Exposure



# A Transmission Model Taking Into Antibiotic Exposure

More details:

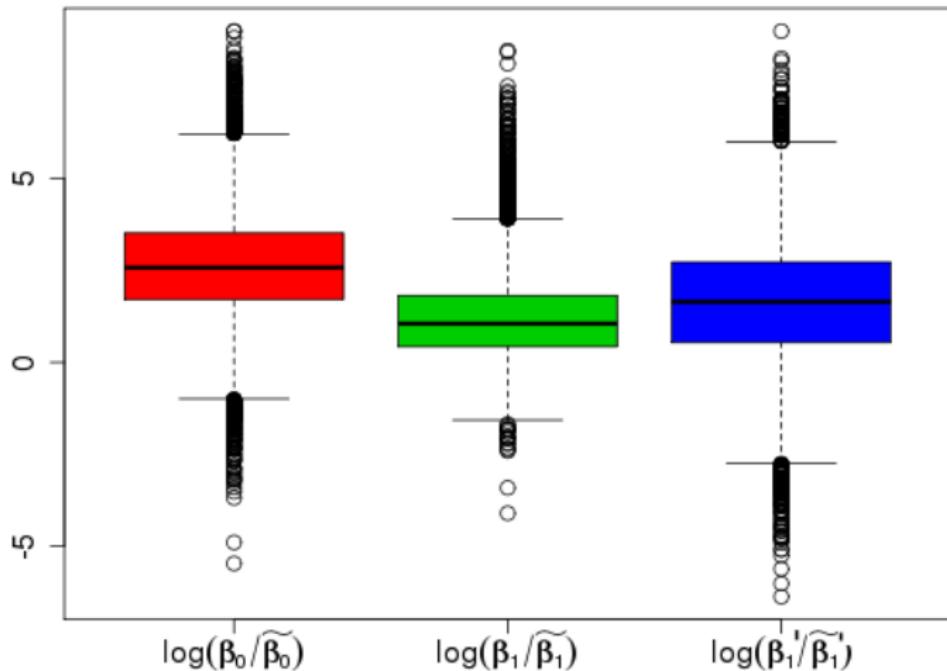
- The probability of acquisition on day  $k$  is given by  $1 - \exp(-q(k))$ .
- We define  $q(t)$  as the transmission rate from susceptible to colonised in the ward at time  $t$ .

$$q(t) = \beta_0 \mathbb{1}_{OFF} + \tilde{\beta}_0 \mathbb{1}_{ON} + \beta_1 n_C(t) \mathbb{1}_{OFF} + \beta'_1 \tilde{n}_C(t) \mathbb{1}_{OFF} \\ + \tilde{\beta}_1 n_C(t) \mathbb{1}_{ON} + \tilde{\beta}'_1 \tilde{n}_C(t) \mathbb{1}_{ON}, \quad (1)$$

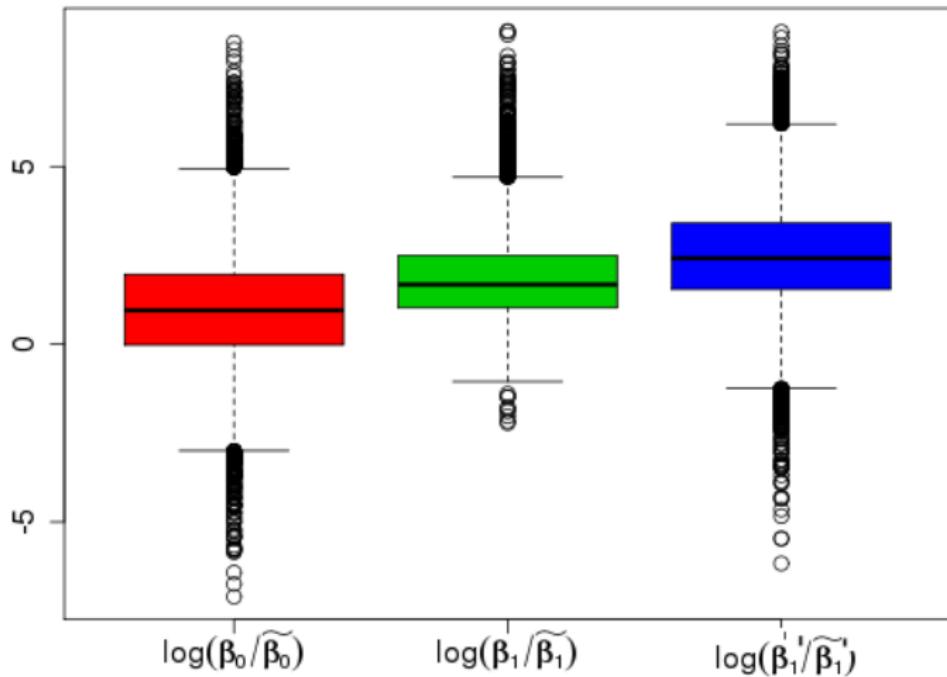
where “~” means “on” antimicrobial treatment.

- Consider one treatment at a time - thus each day, each patient is either ON or OFF the treatment.

## Some Results from Ward 1 for Antiseptics



## Some Results from Ward 2 for Antiseptics

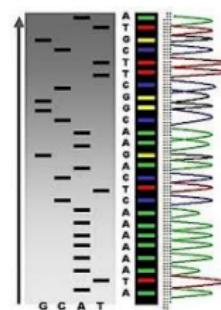
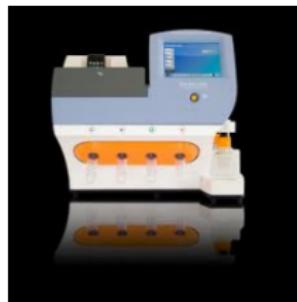


# High-Resolution Genetic Data

# High-Resolution Genetic Data (1)

“High-resolution genetic data”: what are they?

- individual-level data on the pathogen;
- can be taken at single or multiple time points;
- high-dimensional e.g. whole genome sequences;
- proportion of individuals sampled could be high/low;
- becoming far more common due to cost reduction;



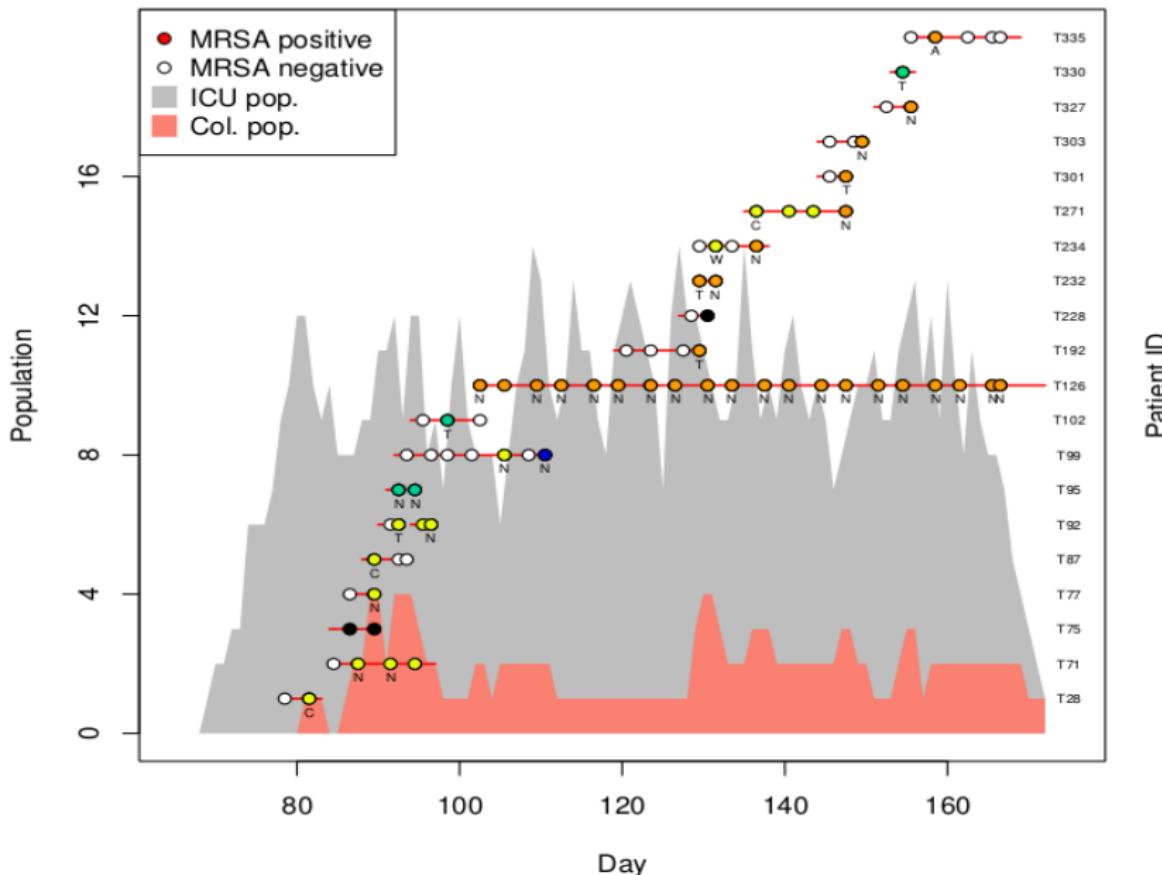
## High-Resolution Genetic Data (2)

"High-resolution genetic data": what use are they?

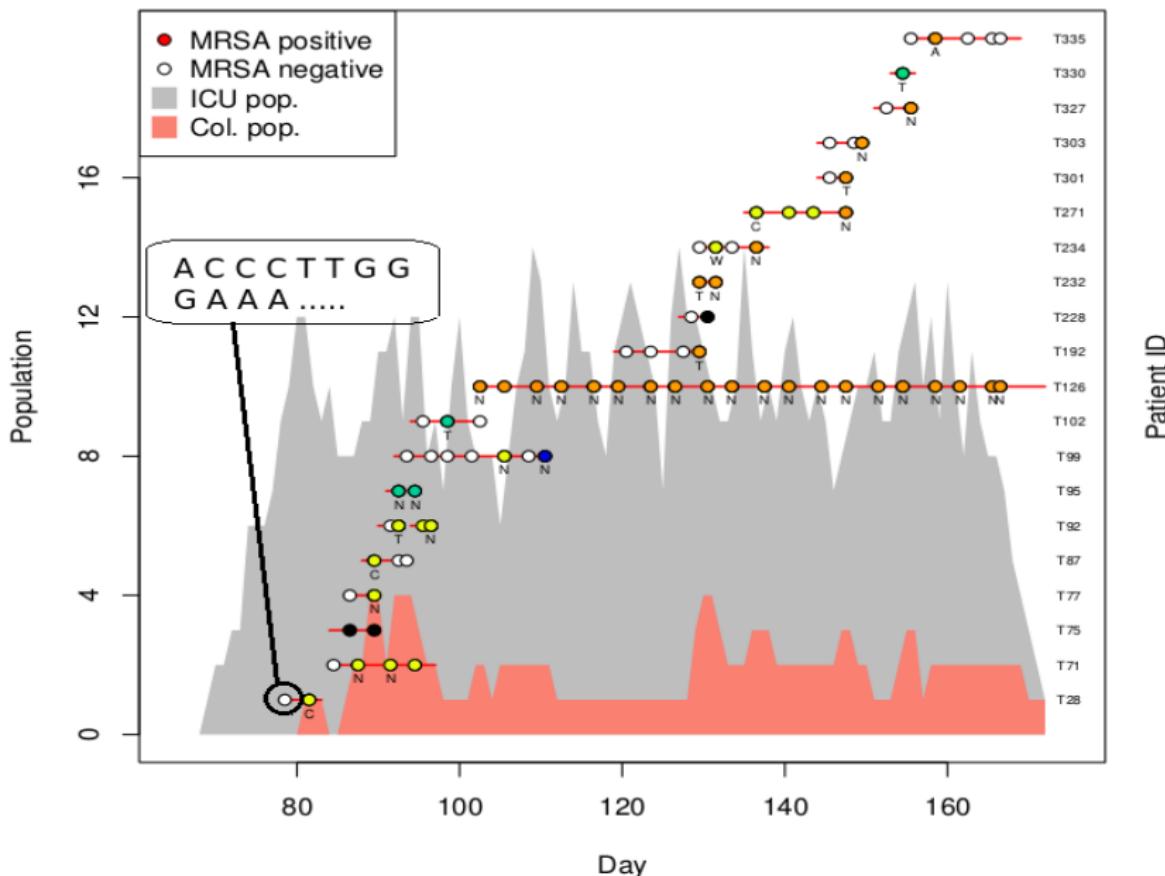
Can provide much insight into the dynamics of transmission:

- better inference about transmission paths (who infected whom)
- more reliable estimates of epidemiological quantities (e.g. the effectiveness of infection control precautions);
- understand evolution of the pathogen.

# MRSA positive patients, ICU 1



# MRSA positive patients, ICU 1



# Existing Work/Studies

This is a **very hot** research topic at the moment!

At least two kinds of approaches exist:

1. **Separate genetic and epidemic components:** For example,

- estimate phylogenetic tree;
- given the tree, fit epidemic model.

or

- cluster individuals into genetically similar groups;
- given the groups, fit multi-type epidemic model.

2. **Combine genetic and epidemic components:** For example,

- model genetic evolution explicitly;
- define model featuring both genetic and epidemic parts.

# Existing Work/Studies (Pros and Cons)

## 1. Separate genetic and epidemic components:

- + “Simple” approach;
- + Avoids complex modelling;
- Ignores any relationship between transmission and genetic information.

## 2. Combine genetic and epidemic components:

- + “Integrated” approach.
- Is modelling too detailed? [mutation, recombination etc]
- Initial conditions: typical sequence?

+/- Model differences between individuals instead?

## Our Proposed Framework

- Develop a more generalized approach to transmission network reconstruction;
- model the distribution of genetic distances observed between each pair of sampled isolates.
- allow multiple independent introductions of the pathogen;
- account for within-host diversity;
- make no assumptions about the evolutionary dynamics of the pathogen;
- do not consider the phylogenetic relationship between isolates.

## Our Proposed Framework: Genetic distance matrix

We define the genetic distance between isolates  $X_1$  and  $X_2$  to be the number of SNPs between the isolates,  $\psi(X_1, X_2)$ .

Since we are interested in the genetic distance between isolates, rather than the composition of the genome itself, we define  $\Psi$  to be the matrix of pairwise genetic distances between all isolates.

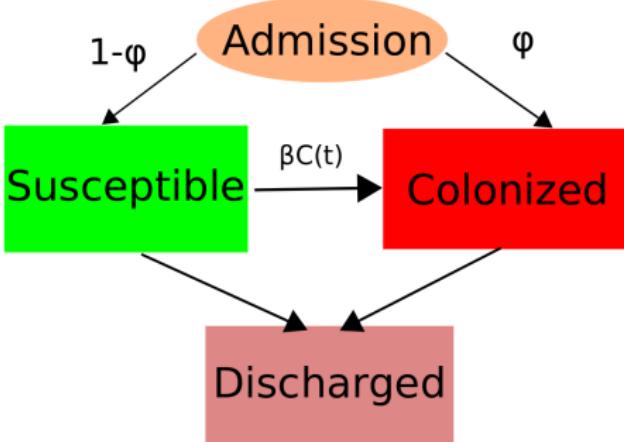
In other words, that means that each new colonised patient ( $i$ , say) needs to have distance  $\psi((i, k)$  to all existing colonised patients  $k$ .

We draw  $\psi(i, k)$  from a probability distribution according to “type”: Each new colonised patient is either:

1. An importation (i.e enter ICU already colonised)
2. An acquisition (i.e colonised by another patient)

# Our Proposed Framework: Putting it altogether

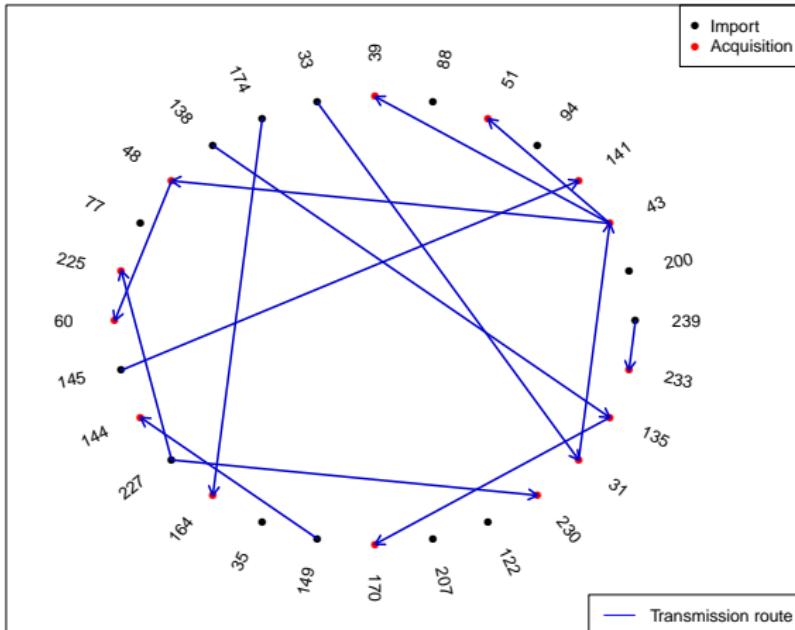
- We develop a probabilistic model for the pairwise distances by taking into account many things (eg lenght of chain, importation etc).
- We then combine this model with a model for the transition dynamics.



- We fit this model to data (not easy, but do-able!: Bayesian, data-augmentation, MCMC ...)

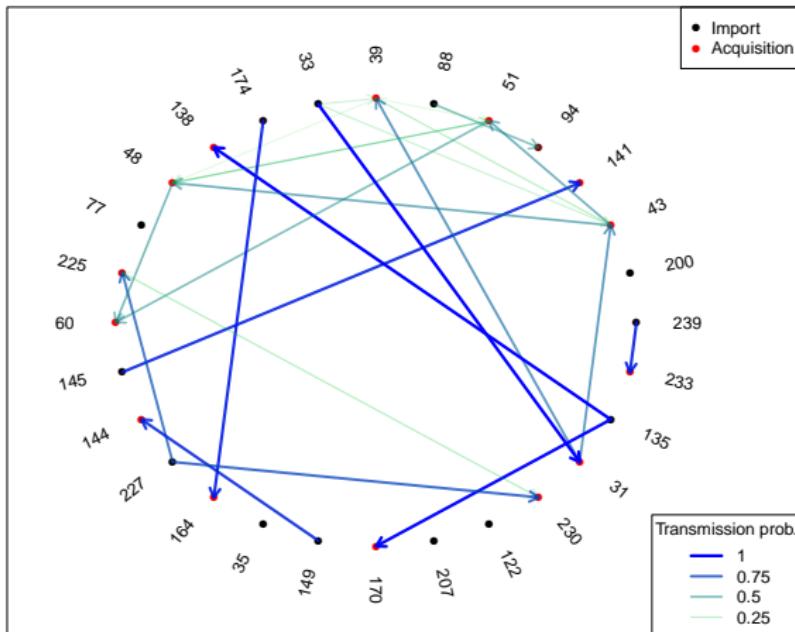
# Simulated patient network

ICU 501: True transmission network



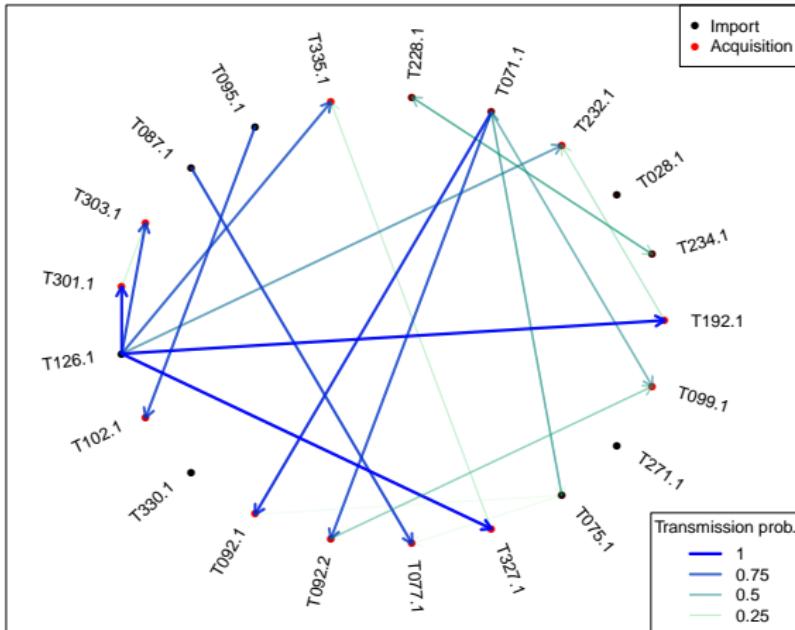
# Estimated patient network

ICU 501: Inferred transmission network



# WGS samples from Thai ICU 1

ICU 1: Inferred transmission network



# Limitations

- Only a small sample of sequences to work with — little indication of scale of within-host diversity.
- Imported strains may be related due to some external source.
- Multiple colonisation is not taken into account — it may be possible for a patient to acquire a second, genetically distinct colonisation which either replaces, or coexists with, the initial colonisation.

## Conclusions & Future work

- A generalized approach to reconstructing infection transmission routes using densely sampled genomic data.
- Although the model might be quite simplistic, provides a framework to incorporate additional complexity to the dynamics of transmission or genetic diversity.
- Within-host diversity makes it harder to resolve network.
- Mechanism to incorporate reinfection would be beneficial.



Worby, C.J., O'Neill P.D., Kypraios, T., Robotham, J.V., De Angelis, D., Cartwright, E.J.P, Peacock, S.J, Cooper, B.S. (2016). Reconstructing transmission networks for communicable diseases using densely sampled genomic data: a generalized approach. *Annals of Applied Statistics*, 10 (1): 395417.

Verykouki, E., Kypraios T., O'Neill P.D. (2016). Modelling the effect of antimicrobial treatment on carriage of hospital pathogens with application to MRSA. *Biostatistics* , 17 (1):65-78.

Worby C., Jeyaratnam D., Robotham J.V. , Kypraios T., O'Neill P.D., De Angelis D., French G., and Cooper, B.S. (2013) Estimating the effectiveness of isolation and decolonisation measures in reducing MRSA transmission in hospital general wards:a model-based analysis. *American Journal of Epidemiology*. 177(11):1306-1313.

Kypraios T., O'Neill P.D., Huang S., Rifas-Shiman, S.L., Cooper B.S. (2010) Assessing the role of undetected colonization and isolation precautions in reducing Methicillin-Resistant Staphylococcus aureus transmission in intensive care units *BMC Infectious Diseases*, 10:29.