"ADVANCES IN INFECTIOUS DISEASE MODELLING"
FONDATION MERIEUX
MEETING REPORT

Report Version July 25, 2008

The Advances in Infectious Diseases Modelling meeting organized by Fondation Mérieux was held at "Les Pensieres" Conference Center from December 10 to the 12, 2007 in Veyrier du Lac, France. The meeting brought together foremost international experts from North America &, Europe, scientific personalities that have performed private and public research investigation on the subject

The following report summarizes the information provided during the Advances in Infectious Diseases Modelling meeting based on abstracts and speaker's lectures, all procedure specifics are not detailed in this report.

Meeting Reporter: ***Valentina Picot***
Report Editor: ***Fiona Hall***

Table of Contents

Information in this report was obtained from the lectures and abstracts given by the speakers on the scientific agenda of the Fondation Mérieux meeting "Advances in Infectious Diseases Modelling" held in France in December 2007. The information in this report was approved by the speakers as per approval signed form. All graphs, flow charts and images were obtained from the speakers' presentations to facilitate comprehension on the subject.
This report was created for meeting reporting information purposes for Fondation Mérieux; the different forms of use of this information might require further authorization by each speaker. The information provided does not constitute a manual or technical sheet on the subject; it might have omissions, we cannot assure its completeness or accuracy, and should not be used for the diagnosis or treatment of disease.

I. Introduction

The first mathematical model of an infectious disease (smallpox) appeared in 1760. Since then, modeling has evolved to today's computerized models in which scientific research and information technology work together. Such models predict the impact of infectious disease prevention, surveillance and control programs and help to anticipate the probable outcome of implementing action plans to tackle infectious diseases.

Most variables that play a role in the fate of infectious disease epidemiology—such as, the host, the pathogen, the target population, the transmission patterns and the eco-social environment—are considered, analyzed and tailored through mathematical predictions.

Infectious disease epidemiology has intrinsic aspects that are not applicable to all diseases; thus in many cases conventional epidemiological dynamics do not always address the needs of infectious diseases. The specific modeling methods and measurements developed to address these types of disease patterns are outstanding and powerful tools to evaluate and interpret data for critical decision-making and program customization for infectious diseases.

The models' applicability ranges from pharmaceuticals and vaccines (to determine vaccine strategies for current vaccines; as well as for those still to be developed, like the HIV vaccine) to help predict the likely spread of vector-born diseases such as Rift Valley Fever.

The sessions reported in this document give us a greater understanding of what a model is and its potential application. The meeting was an invaluable opportunity for us to develop a better approach to modelling infectious diseases.

Various public & private institutions and organizations involved in infectious diseases epidemiology have taken advantage of these models for public health strategy-making, and for optimizing the use of resources, among other applications.

II. Meeting Objective

The aim of this symposium is to give an overview of the different questions that the modeling approach can answer, using recent applications for different infectious diseases.

III. Summary of Scientific Agenda Lecture Presentations

The meeting was presented in sessions as follows.

2. **Session I: What is a Model?**
    a. **Chaired by: Odo Diekmann and François Simondon**
    b. **Lectures Briefings**
    c. **Discussion**
3. **Session II: What is the Expected Public Health Impact of Model Approach?**
    a. **Chaired by: Daniel Barth-Jones**
    b. **Lecture Briefings**
    c. **Discussion**
4. **Session III: Predicting the Impact of Interventions?**
    a. **Chaired by: P. Beutels**
    b. **Lectures Briefings**
    c. **Discussion**
5. **Session IV: The Future of Infectious Diseases Modelling**
    a. **Chaired by: Martin Eichnert, Ira Longini**
    b. **Lectures Briefings**
    c. **Discussion**

1. Welcome Address & Keynote Presentation

Christophe Longuet, Medical Director of the Mérieux Foundation, welcomed the speakers and participants to "Les Pensières" conference center. He presented the foundation's mission: to control infectious diseases in developing countries by supporting scientific research, sharing knowledge and supporting health structures, patients and their families. The presentation allowed participants to better understand the scope and role of the Mérieux Foundation in disease control activities, emphasizing that the Advances in Infectious Disease Modelling meeting is one aspect of the foundation's knowledge sharing activities.

Dr Bernard Ivanoff, the meeting's scientific organizer, also gave a welcome address and an overview of the agenda. He introduced the keynote lecture, which follows.

1.1. Keynote Lecture

Global Climate Patients to Model the Spatial and Temporal Distribution of Vector-Borne Diseases

Kenneth J Linthicum, USDA-ARS Center for Medical, Agricultural & Veterinary Entomology, Gainesville - USA

Urbanization transforms our landscape and our surrounding ecosystems, including climate and the fate of diseases. A glance at a map of the Earth at night gives us a quick picture of what urbanization has done to the Earth; if we were to have done this a few decades ago we would not have seen the same picture. This all has a big impact on climate, disease and ecosystems.

Some background

The Earth's oceans are essentially the engines that drive climate patterns; climate, oceans and their systems are very closely linked together.

The El Niño/Southern Oscillation (ENSO) is the most well-known phenomenon influencing global climate variability. El Niño refers to a large-scale ocean-atmosphere climate phenomenon linked to periodic warming of sea surface temperatures across the central equatorial Pacific.

During a warm phase of the ENSO, the Pacific becomes abnormally warm and the same usually happens in the equatorial western Indian Ocean, while there is cold water in the Atlantic and in the eastern Pacific Ocean. This circulation maintains the Earth in balance; a balance which is key for the Earth's survival. During a cold phase of the ENSO, the Pacific is cold, and this is usually the case too for the equatorial western Indian Ocean.

There is growing evidence to suggest that there are links between ENSO-driven climate anomalies and infectious diseases, particularly those transmitted by arthropods. A few examples are Murray Valley encephalitis (Nicholls, 1986), Bluetongue (Baylis et al., 1999) and Rift Valley Fever (Linthicum et al., 1999), among others. Evidence of the links between ENSO-driven climate anomalies and infectious diseases, particularly those transmitted by insects, can allow us to improve our long-range forecasts of an epidemic or epizootic.

Rift Valley Fever (RVF)—a vector-transmitted viral zoonosis that principally affects domestic animals (livestock)—was first described in Kenya by Daubney et al. in 1931 in a report of a fatal epizootic of sheep on a farm north of Lake Naivasha. The disease results in significant and widespread livestock losses and frequent human mortality.

Until 1977 RVF was restricted to sub-Saharan Africa, but in 1977 the disease appeared in the Nile valley and then in the delta region of Egypt, and it continues to reoccur in these regions. In 1987 there was an outbreak in the Riverine floodplain and then in 1997-98 the largest outbreak in Africa occurred in Savanna Grassland coinciding with one of the largest ENSO events. In 2000 the disease left the African continent for the first time and spread to the coastal floodplains of Yemen on the Arabian Peninsula and of Saudi Arabia along the Red Sea. The importance of this was not only that the disease was spreading outside the African continent, but also that the morbidity and mortality in humans and animals were significantly higher than ever seen in Africa.

A number of years ago we discovered that the outbreaks of RVF followed periods of widespread and heavy rainfall associated with the development of a strong inter-tropical convergence zone over Eastern Africa.

A peculiarity of RVF is that it requires rainy seasons for the disease cycle and vector transmission to be completed. There is an endemic cycle where the virus persists during the dry season/inter-epizootic period through vertical transmission in *Aedes* mosquito eggs. The epidemic cycle begins during flooding, which triggers the mass hatching of infected *Aedes* eggs and subsequently *Culex* mosquitoes, which can be important secondary vectors of RVF, leading to an RVF outbreak.

Looking retrospectively at RVF outbreaks from 1950 to 2006 following the Southern Oscillation Index (SOI) part of the ENSO, we observed that every RVF epidemic and epizootic correlates with a period of strong ENSO activity. However, not all ENSO/Niño activities result in an RVF epidemic or epizootic.

Observations show that in order to have a big RVF outbreak, the warming of the Pacific Ocean needs to be concurrent with the warming of the Indian Ocean, because this will produce heavy rainfall in the horn of Africa. For example, the 1997 RVF outbreak in the horn of Africa, which was very devastating from both an economic and a human disease point of view, showed a strong correlation with this phenomenon.

Based on this evidence, operational mapping of RVF in Africa has been done since 1999 using satellite devices which can measure, among other things, rainfall and sea temperatures. In the 1997-98 El Niño, high sea surface temperatures (SST) coincided with heavy rainfall in the Pacific and Indian Oceans. Such extremes in climate affect vector abundance in different ways and produce an elevated risk of disease outbreak.

Identifying anomalies in these climate patterns would allow us to forecast possible disease outbreaks. For example, excessive rain can affect the path of RVF disease by boosting food supply, elevating the rodent population and enhancing mosquito breeding and propagation. On the other hand, drought can suppress predators of the *Anopheles* mosquito, elevating the risk of malaria in some places. Dengue fever transmission can be exacerbated by increasing water storage and elevating temperatures that favors the mosquito's incubation period.

This knowledge allowed us to predict that the 2006 fall-winter development of El Niño conditions would have significant implications for global public health, particularly for RVF. Extreme climate events (coinciding Pacific and Indian Ocean SSTs) with above normal rainfall occurred in the Horn of Africa in the third quarter of 2006 and led to the RVF outbreak in 2006-2007. An unusual climate pattern also developed over Sudan in the middle of 2007, leading to the potential for a RVF outbreak in 2007.

Based on these observations, a disease outbreak advisory was published as usual on our institutional website; it has also been published in the *International Journal of Health Geographics* since 2006. Some predictions included the following:

- In Indonesia, Malaysia, Thailand and most of the Southeast Asia islands, drought was likely to increase dengue fever transmission; respiratory illness was also likely to increase due to haze from uncontrolled burning of tropical forest during the extreme drought.
- In East Africa predictions showed that flooding due to heavy rainfall in dryland areas would elevate the risk of cholera, as well as of RVF and malaria, due to an elevated vector population.

These disease forecasts allowed disease alerts, advisories and warnings for disease prevention to be published, and surveillance systems by different institutions to be installed.

Because of the early warning system (based on climate indicators) that was put in place, the response activity started 1 or 2 months earlier for the 2006/07 outbreak than in 1997-98. The more proactive approach in the way the information was processed and disseminated improved the control and surveillance of the disease despite the use of the same technology as

had been used previously. The response included a number of different activities, such as mosquito control measures, avoidance of meat consumption from outbreak areas, animal vaccination programs, etc.

Conclusions:

- Forecasting disease is critical for the timely and efficient planning of operational control programs.
- Global, regional and local climate anomalies can be used to forecast potential disease risks that will give decision-makers additional tools to make rational judgments about disease prevention and mitigation strategies.

Discussion

**Why do you think RVF has not been reported in Southeast Asia?*

Not sure why, there are a number of groups that are trying to prevent the escape of the disease from Africa and have been fairly successful, but I suspect that given the movement of people and animals, this risk is quite high. This is taking into consideration that you have mosquito vectors in Southeast Asia that can play a role in the RVF spread there.

2. Session I: What is a Model?

Applications of Models: Roles and Approaches

Neil M. Ferguson, MRC Center for Outbreak Analysis and Modelling, London – UK

Why use a model when there are so many uncertainties? Even for the diseases we know most about there are many uncertainties (transmission patterns, immunity, the pathogen and so on), and often only limited historical data. Given that models necessarily simplify and make assumptions, what is the value in using them?

The fundamental reason is that if we don't have a model, judgments are made on the basis of qualitative evidence, sometimes biased by opinions. A model:

- makes assumptions explicit
- optimizes the use of limited data
- highlights key factors that drive epidemiology and transmission dynamics of infectious diseases, thus helping to determine policy needs.

What is a model? And why are models so much more applicable to infectious diseases than other areas of biology and medicine? The fundamental reason is the commonality between all infectious diseases, which is transmission and its different pathways (a person infects another person or an animal to a person etc).

What all these diseases have in common is that infection spread creates a chain reaction and exponential growth of infected individuals; the classical initial phase for an epidemic to occur. The most important *quantity* governing an epidemic is *how many* other people one person infects. This is the **basic reproduction number of an epidemic (R_0)**, and it needs to be greater than one for an epidemic to take off. This also determines the intensity of control

measures required for an epidemic, as well as the growth rate and how quickly the disease spreads.

Other quantities are also important. For example, **generation time (T_g)** tells us how long a particular individual will take to infect other individuals. For example, a disease like influenza, which has a low reproduction number and short generation time, can spread just as fast as diseases like measles which have a high reproduction number but also a very long generation time.

This spread process stops when a disease begins to run out of people to infect. This may be the case, for example, because immunity builds up in the population, which causes the number of secondary cases to drop below R_0. The basic reproduction number is instead defined by R, the effective reproduction number = $s \times R_0$ (s = proportion still susceptible).

Once $S < 1/ R_0$, (so $R<1$), the epidemic goes into decline. This is important because in terms of infectious disease control you don't need to eliminate all transmissions to control an epidemic; you only need to eliminate a portion. What portions you need to eliminate to control the epidemic is determined by the reproduction number. To control an epidemic a policy needs to reduce R to below 1 so that transmission cannot sustain itself. This can be achieved by eliminating a fraction $(1-1/ R_0)$ of transmission. For example, you have to block 75% of transmission if you have an R_0 equal to 4; thus the higher the R_0 the more effort required to control an epidemic.

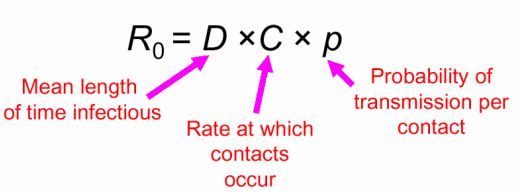
Some measures to eliminate a fraction of transmission include reducing contact through quarantine and social distancing; reducing susceptibility through vaccination and prophylaxis; and reducing infectiousness through treatment, etc.

Key aspects to be taken into consideration are: Who is targeted, how fast and how much effort is needed?

Mathematical models capture all this qualitative information in a mathematically rigorous way. **The challenges are how to estimate the R_0 and T_g for a particular disease and population, and how to estimate the effect of control measures based on these parameters.**

To do this we need to deconstruct the R_0. Although the R_0 is quoted as a given quantity, it is important to recall that it is not a fundamental biological constant at all. Instead it is composed and determined by many different aspects, such as the pathogen, the host factors, the population structure, etc. Taking all these into consideration allows the R_0 to be estimated more accurately. However, mechanistic understanding (not just curve fitting) is also needed to predict the impact of controls; thus data are necessary.

Example: This is a highly simplified example to determine the R_0, and only really applies if all contacts have an equal risk of infection and if contacts are not repeated:



As mentioned before, there are many different biological, natural history aspects etc. that need to be taken into consideration when determining the R_0. In reality these processes aren't as simple as expressed in the previous equation: diseases evolve gradually, there is incubation time, variable infectiousness, morbidity, mortality etc. For example, it was smallpox's two-week incubation period that enabled this disease to be eradicated.

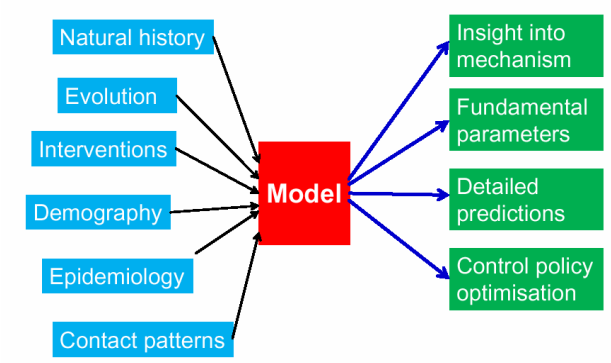
A concern is that transmission in diseases is almost never observed; there is little data on transmission mechanisms and our knowledge of the physics of transmission is still very limited. An example is influenza, for which there are only a small number of studies on transmission in small population units. Mostly transmission parameters have to be estimated by matching models to surveillance data.

Defining relevant contacts is often a challenge. Such data are easiest to gather for sexually transmitted diseases such as HIV.

In modeling infectious diseases, dynamics and population biology seem relevant to integrate our understanding of the complexity of immunity into the evolution of pathogens and genetic data. Population structure and polymorphisms are still often not well understood.
Antigenic (strain) data are often available and linked to genetic data for RNA viruses but not for many more complex pathogens, and molecular basis of transmissibility is also poorly understood.

Using this data in models could give us a much more predictive understanding of how a disease evolves over time. From the epidemiology perspective it would give us a more refined understanding of transmission patterns of disease in a population. To predict population effects we nearly always need to extrapolate data. There are usually two types of population data: clinical trials and observational studies.

A model synthesizes data from all these sources into a mathematical computation framework to produce a number of outcomes, as seen on the following graph.



It is important to mention that not all models are mathematical, as when using sources only in qualitative ways, but mathematical models can help this data reveal more explicit understandings.

The many uses of mathematical modeling include:

- quantifying risk (what might happen?);
- helping in knowledge synthesis by analysing data, extrapolating, optimizing control policies; and
- making assumptions explicit so that they can be tested/disproved.

No model can exist without data; data are the essence of a model.

In designing a model for a particular disease one needs to decide if the model should be deterministic or stochastic, compartmental or individual-based, etc. The complexity of the model should be driven by need (what does the model need to do?), and by data (what assumptions and level of detail can be justified?).

“The art of modelling is knowing what to leave out”

The history of epidemic modelling shows that very simple models can give remarkable insights, both qualitative and quantitative, into the patterns of disease spread. An example would be a model of measles dynamics.

Today’s modern computer models can go much further, particularly with emerging infections. One example is of a simulation of the emergence of a pandemic in Indonesia. This simulated 230 million people, but with only five transmission parameters. However, one needs to be cautious about the extent of need for simulation at that scale.

Model validation is also very important, regardless of the complexity of the model. Key parameters should be estimated from data and models should reproduce past epidemics to test “goodness of fit”. However, as models are rarely fully comparable because no two epidemics are quite alike, sensitivity analysis is important.

In recent years a number of trends in modelling have been driven by the availability of data and computer power; most emphasis has been placed on endemic diseases because we have been limited in solving the equilibrium properties of simple models, such as: what type of long term measures might change that ***equilibrium*** (eg. vaccination)?

HIV and other emerging epidemics, along with more powerful computers, have enabled us to model the ***dynamics*** of novel epidemics. Diseases such as SARS showed the potential for ***real-time*** modelling. In the case of endemic diseases, the focus was on seasonal and spatial dynamics. There is need for more emphasis on model fitting and parameters estimation, on the need to integrate genetics and epidemic modelling, and in demonstrating model relevance to public and veterinary health.

Why do we worry about emerging infections? We do so because of their devastating impact and their potentially profound effect on society. The risk from these diseases may be increasing due to higher animal and human population densities, declining habitats, etc. Thus predicting an emergency is a priority and can be done by detecting growing clusters of cases of a new disease, innovative analysis of surveillance data and new analytical methods to analyse cluster data, such as rapid field investigation.

Modelling scenarios can be powerful advocacy tools by showing the benefit of modelling for policy decision-making, and assessing disease preparedness and control options (vaccination, quarantine, treatment, prophylaxis).

What are the new challenges facing infectious disease modelling?

- We live in a much more mobile world where diseases spread faster, so models need to provide faster and better responses.
- The public health response needs to prioritize limited resources, so models need to deliver actual health benefits that can be directly attributed to modelling.
- We need better natural history/transmission models, we need to be able to quantify and validate proxy measures of infectious contact patterns and to understand and collect data on how interventions block transmission. In all of this we need to maintain some essential simplicity.

Discussion

**You describe all these different parameters and techniques for modelling; how do you think inference models can look like large scale models, how can we put it all together in a comprehensive package that goes from estimation all the way through to policy analysis?*
I think that is one of the key challenges for the next few years; to date we have had a very ad hoc process by which models are parameterized.
**For diseases with historical data we can predict the outcome, but for those diseases for which we have very limited data—the new diseases—will simulation data really work?*
For SARS we knew nothing when we were tracking the epidemic as it unfolded; all we can do is to track transmission. I guess this is like influenza. We don't know exactly how the influenza pandemic will unfold, but can extrapolate from the past to offer more than informed guess work about the likely pattern of spread.

Simulations: What Level of Complexity is Appropriate?

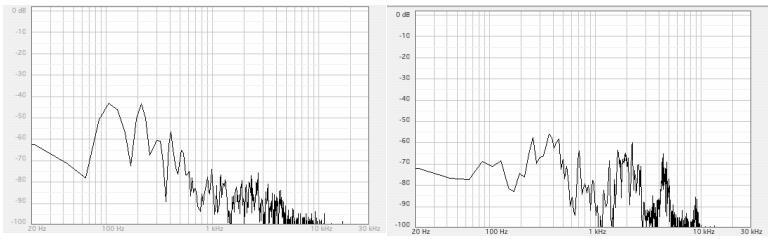
Stephen Eubank, Network Dynamics and Simulation Science Laboratory, Blacksburg – USA

Do individual based simulations represent advances in the use of models for controlling infectious disease? This talk will be based more on epidemic than endemic diseases.
Our hypothesis, or H_0, is that simulations aren't science. Usually science is described as taking a complex situation and making it simple, then building an analytical solution to that simple system. In contrast, simulations are not generalizable—they don't find a specific solution to a specific problem and they don't help you build any intuition about the system. All you get is a simulation result and you have to make sense of that. Also models are reductionist: there are no symmetry assumptions and no analytical solutions.

Models of complex systems are complex or idealized
Our hypothesis is that simulations aren't worth the effort: they cost too much (calibration, validation), they yield too little (numbers, not understandings), and they are hard to understand (only experimental sensitivity analysis).
Even in a simple model, calibration and validation are not easy. The following is an example related to music:

A Simple Linear Model
A (f,t+1) = λ(f) A(f,t)
Define R_0 ≡ λ_1 / λ_2, ratio of two largest λ's
A (f,t+1) = λ(f) A(f,t)
Define R0 ≡ λ1 / λ2, ratio of two largest λ's

T=1 T=10



A(f,t+1) = λ(f) A(f,t)
Define R0 ≡ λ1 / λ2, ratio of two largest λ's.

Q1: What is the limit after many iteractions?
A1: if R0 >1, f1 (+ harmonics)

Q2: How many are "many"?
A2: n >> log (A(f2,0)/A(f1,0)) / log R0

- **Calibration:**
  - Measure frequency response
  - Estimate R0
- **Validation:**
  - Perform "Sitting…" and listen
  - If ≠ expectations after many iterations, model is
    - Not valid? No, we know it is structurally correct
    - Out of calibration, because system is non-stationary

Even if the model is structurally valid some things might not be quite perfect in the calibration and this is because the system is non-stationary.

Calibrate here…….. **and your system will be out of calibration here..**

Infectious diseases are non-stationary systems.
For example, in the efforts to eliminate measles "failures were due to:

- Faulty data on the level of immunity required for herd protection.
- Models used did not take population heterogeneity into account."

Black and Singer, 1987

Useful epidemiological models include biology and social environment:

Transmission rates (biology) + opportunities for transmission (sociology) = epidemiological model

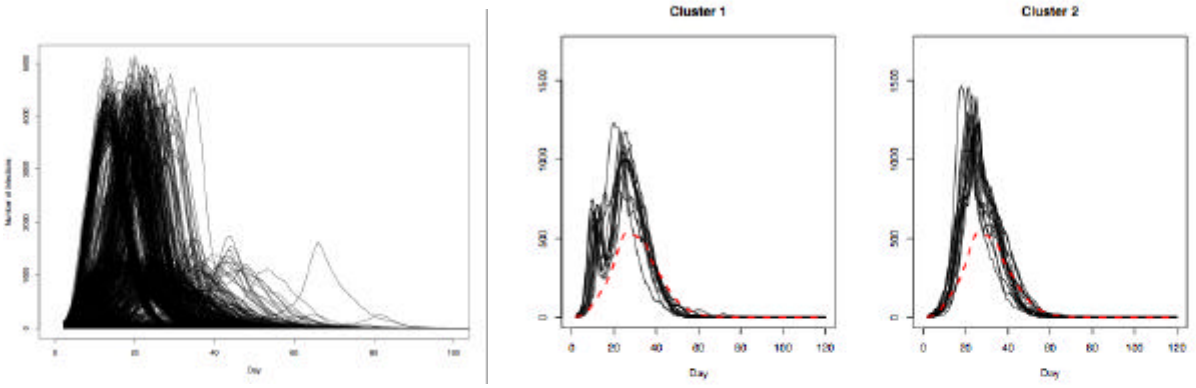
The benefits of using simulations are that:

- Their costs are not that much higher than other models.
- They allow complete representations, which are very difficult to reproduce with analytical base models. "Much relevant work remains to be done in teasing apart the social, genetic, age-related, and other complications that are smoothed out in the usual mass action assumption" (May, 2000).
- They allow you to understand not only the initial conditions at T_0 (time) and the final conditions, but also what happens in between. They provide information about intermediate times.
- They can explore important questions, after an outbreak begins but also before it is over.
- They can help in decisions about the optimal allocation of limited resources to control infectious diseases, which requires knowing who is most vulnerable, who is most critical and when.
- They allow new characterizations such as: vulnerability (the probability of infection in a set of people), and, criticality (the change in others' vulnerability when a set of people is removed). Both of these characteristics depend on the set of people, the time, the initial conditions, the transmission dynamics and the contact network.

When deciding the appropriate complexity of a model, the model itself can give you insights. Taking the following into consideration can determine if a simple model can be the outer

layer of a more complex model. Then you can decide whether the simple model will represent the requirements of the model well. The things to consider are how the sensitivity specification compares to model differences, information required on parameters, and whether the cost of improving specification outweighs the benefits.

One of the complaints about simulation models is that you just get data out, and there is not much you can do to understand it, hence there is too much variability. There are standard statistical techniques that will allow you to impose conditions at time T to reduce variability significantly at T+ ? . (see graph bellow).



One technique is to cluster results. For example, with Principle Components Analysis you can assign each run to a cluster, then find conditions determining which cluster a run is in.
Compare the information required to encode condition to information gained by reducing variability.

Simulations enable experimental epidemiology in a controlled, stationary, virtual world. The search for groups with extremes of risk is an important focus of epidemiology:

"The study of health and disease of populations and groups … The clinician deals with *cases*. The epidemiologist deals with cases *in their population*." (Morris, 1955).

It is made possible through controlled experimentation in a stationary setting.

Conclusions:
Detailed simulations

- are not necessarily harder to calibrate and validate
- yield insights into adaptive, targeted control
- can suggest the appropriate level of complexity
- enable otherwise impossible analyses.

References

F. Black and B. Singer, "Elaboration versus simplification in refining mathematical models of infectious disease," *Annual Reviews Microbiology* 41:(1987) 677-701
May, S.R. 2000, Enhanced: 'Simple rules with complex dynamics', *Science* **287** p. 601 (2000).
J.N. Morris, "Uses of Epidemiology", *BMJ* 1955 (2) pp. 395-401

Model Parameterisation, Validation Methods and Data Needs

Azra Ghani, Jamie Griffin, Tini Garske and Paul Clarke, MRC Center for Outbreak Analysis and Modelling, London – UK

Over the past 20 years, with the increase in computational power and the development of more complex models, there has also been a move towards more statistically-driven determination of parameters for models. In this talk I will review some of the common methods that are used to explore parameter uncertainty, as well as to obtain statistically rigorous parameter estimates. These methods will be demonstrated with examples from human and animal diseases. Data needs for the different methods will also be discussed.

I will start by giving some background on parameterisation and the idea of validation and fitting models through available data.

The parameters of parameterisation have varied over time; they have also varied between the different disease areas. Some of the approaches that have varied are:

- Sourcing parameters from external data, mainly used when computer power was not available (often with no validation, and performed in comparison with other outputs from the model, such as transmission, incubation period, etc.).
- There has been a general shift to more formalized methods of validation – statistical fitting. However, it is important to take into consideration that in some cases it is impossible to use exclusively formalized methods of validation. This should not be a reason not to undertake model work, thus models can give interesting hypotheses to reduce uncertainties in the output.

It is important that when you move to model fitting to estimate parameters from data that these be identifiable, the data must contain some information on the particular parameter, which is statistically identifiable from the data. If not, then sourcing from external data is the only appropriate method.

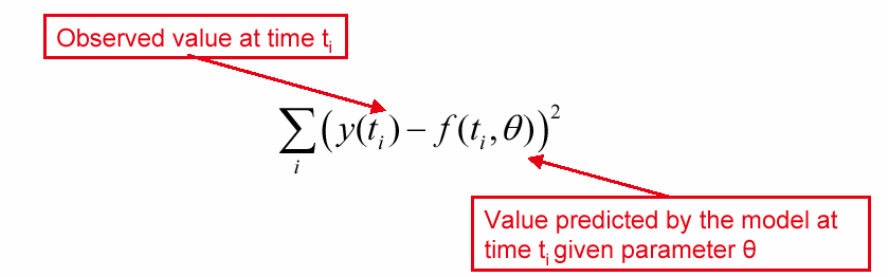
This talk focuses on statistical fitting or parameter estimation. The aim of model fitting is to alter parameters of the model so that they are in greater agreement with observed data.
Models can be validated to a certain extent by fitting all the parameters in a model, by fitting some of those parameters of the model output to some external data; or by estimating the parameters. Parameters can be sometimes difficult to determine, such as the reproductive numbers (R_0).

One can also start to improve estimates of parameters that have been obtained before, for which some prior knowledge is known and by refitting those in a model, they help us to obtain better estimates of the parameters.

Parameter estimation is not the only aspect of model validation; this talk will not cover the structure of uncertainty in a model, which is often left out when people consider validating the models.

Model fitting aims to alter the parameters of the model so that they are in greater agreement with the observed data.

The simplest approach is to use least-squares, basically by minimizing this function:

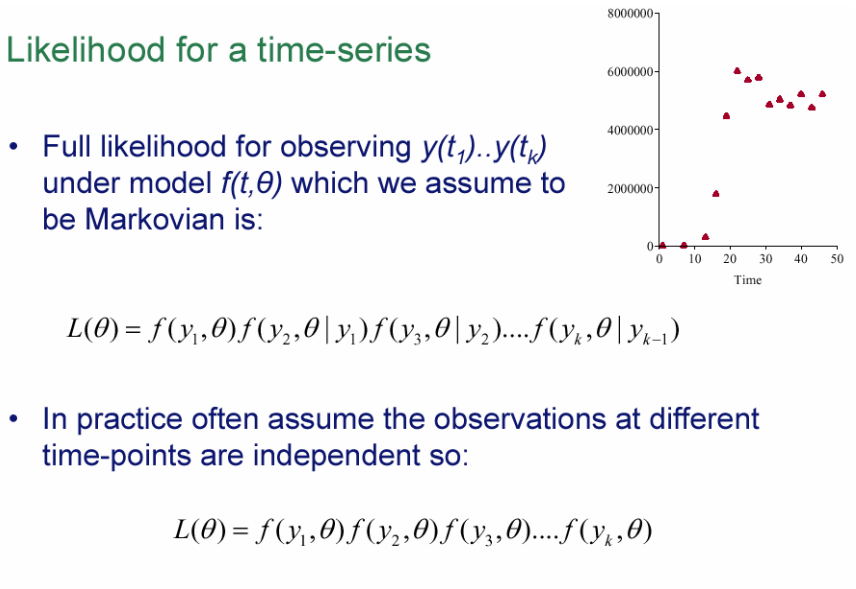


This is a very simple way to fit your model; however, this approach does not necessarily give the best estimate of the parameters.

A better approach to get those estimates is the **maximum likelihood method**. This method is common in statistics. Through it you maximize the likelihood function, ie. the probability that you observed the data under some model.

Some advantages of this approach under some weak assumptions are that it tends to be asymptotically unbiased, especially if you have enough data, the bias tends to zero for a larger n. It is asymptotically efficient, as no other estimator has a lower mean-squared error; and it is asymptotically normal, which means for large n estimate has a Gaussian distribution (useful for constructing confidence intervals).

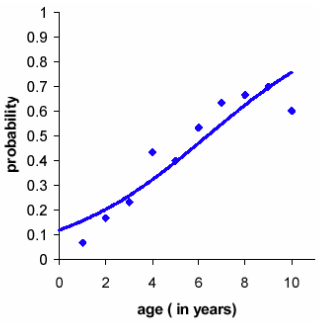
In order to implement the maximum likelihood method, you need to carefully understand the likelihood for a time-series, as follows:



The following is an example of how to apply the maximum likelihood method:

Example – Binomial Data

- For serological data our observed value $y(t_i)$ is the number of serologically positive individuals at time t_i out of $n(t_i)$ that were tested
- We have observations for $t_i=t_1 \ldots t_k$



$$L(\theta) \propto \prod_{i=1}^{k} \left(f(t_i;\theta)\right)^{y(t_i)} \left(1-f(t_i;\theta)\right)^{n(t_i)-y(t_i)}$$

$$LnL(\theta) \propto \sum_{i=1}^{k} y(t_i)\ln(f(t_i;\theta)) + (n(t_i)-y(t_i))\ln\left(1-f(t_i;\theta)\right)$$

- Equivalently minimise deviance $D = -2LnL(\theta)$

By generating different model scenarios we can then determine the overall likelihood of that scenario. Each given scenario can generate sets of different data, from which we pick the most likely data: the data that maximize our likelihood as our best estimate of parameters and of the epidemic.

The likelihood method is used even in situations with a great deal of uncertainty because it does allow us to quantify that uncertainty, to get better estimates and also to have confidence in the model's parameters and future predictions.

Bayesian Methods

These methods effectively extend the likelihood-based approach to incorporate information we already knew about some parameters. They are based on the Bayes Theorem:

$$f(y,\theta) = f(y \mid \theta) f(\theta)$$

Posterior density — Likelihood — Prior

With this model you can get posterior densities for the parameters as well as for the model predictions.

This is a slightly more integrated approach for parameter estimation and is increasingly being used in computation methods derived from Bayesian methods like Montecarlo sampling. However, the formal Bayesian method has been less frequently used; perhaps the field could move towards a wider use of this method since we now tend to have quite a lot of prior information in our parameters.

Estimation during epidemics
This is a growing area, particularly in infectious disease modelling. During epidemics, key quantities to estimate are the basic reproduction number (R_0) and the effective reproduction number.

The traditional approach to estimating R_0 is based on the early growth rate ? and the mean duration of infectiousness D.

$$I(t) = I(0)\exp(\Lambda t)$$
$$\Lambda = (R_0 - 1)/D$$

A big advantage of this method is that it is fast and simple. However, there are limitations as it is difficult to define when interventions were implemented and their effect, and it requires assumptions to be made for the duration of infectiousness (D), thereby introducing uncertainty into that value.

There has been a move towards simplifying epidemic model fitting. For any model it is fairly straightforward to write down the likelihood of the infection process given full data and conditional on first infection. However, the demanding part is actually running it through and looking at all the different interactions and dealing with some of the data complexities.

If we have infection times (t_1 …..t_m) and a total population of size N, then our likelihood is simply the product of the risk of each of those people becoming infected at times t_1 …to t_m, multiplied by the cumulative risk of being infected, multiplied by the cumulative risk of ever being infected.

$$L = \prod_{k=2}^{m} \text{hazard for } k\text{'s infection at } t_k$$
$$\times \prod_{k=2}^{m} \exp(\text{-cumulative hazard for } k \text{ being infected before } t_k)$$
$$\times \prod_{k=m+1}^{N} \exp(\text{-cumulative hazard for } k \text{ ever being infected})$$

We then calculate the R_0 from estimated parameters, and determine the impact of intervention if you know when those interventions occurred.

Good parameter estimation is key for understanding the impact of control measures.

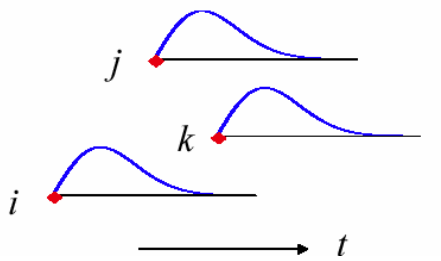
Some of the complexities observed in data during epidemics:
If you have a fairly simple model, you can write down the full data likelihood, and if you have a good data set, then you can fit your model very easily. However, one of the major challenges is "unobserved" data. Estimation methods need to deal with this type of data, which include infection times, onset of infectiousness, censoring (those that have been infected but have not yet developed the disease), missing infections/cases, contact network, etc.

One way to deal with this is to use data augmentation techniques to simulate the possible units for a given unobserved parameter, for example possible infectious time, etc. Basically you sample unobserved data as if they were parameters in a Bayesian MCMC scheme.

The generation time method (Wallinga and Teunis, 2004)
This method for model fitting estimates the basic reproductive number R_0 from observed infection times using the generation time distribution.

You label your cases in the order they were infected (j = 1,…, m) and assign them their ordered infection times t_1 ,…,t_m, with the first case assumed to have been infected from outside. Thus *g* stands for generation time density with parameters *?*. The probability that *j* infected *k* (with $k > 1$) is estimated as:

$$p(j \to k \mid \theta) = \frac{g(t_k - t_j \mid \theta)}{\sum_{i=1}^{k-1} g(t_k - t_i \mid \theta)}$$



You can estimate the number infected by *j* by adding up these probabilities

$$R_{(j)} = \sum_{k=j+1}^{m} p(j \to k \mid \theta)$$

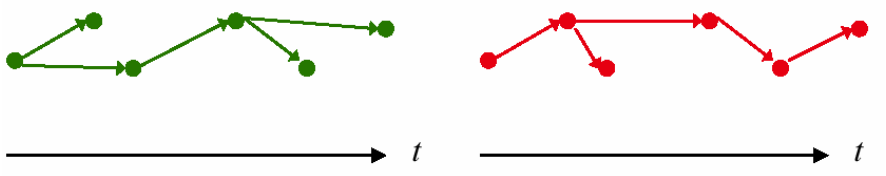
So if you know generation time and distribution, it is fairly simple to work out the probability of everyone infecting everyone else in the population. Then you can estimate the number that each individual in your population went on to infect, by summing these probabilities. That gives you an estimate of the reproductive number for each individual. You can look at these as a time series to get an estimate of how the reproduction number has changed.

This is a very appealing method to get a view of how an epidemic is progressing. This method has been applied retrospectively to data from the 2003 SARS epidemic (Wallinga and Teunis, 2004).

Estimating generation time parameters
Define the infection tree v=(v_2….,v_m), where v(k)=j denotes that *k* was infected by *j*.
The likelihood of the infection time *t*; given *v* and *?*, is:

$$L(v, \theta \mid t) = \prod_{k=2}^{m} g(t_k - t_{v(k)} \mid \theta)$$

Probability under assumed generation time distribution that k was infected by v(k)



These dots show a time series of cases and represent two ways that the infection (infection tree) could have been passed through this population.

In practice since we don't observe the infection tree, we can instead use the integrated likelihood (over all possible infection trees assuming that they are equally like):

$$L(\theta \mid t) = \sum_{v} L(v, \theta \mid t) = c \sum_{v} \prod_{k=2}^{m} g(t_k - t_{v(k)} \mid \theta) = c \prod_{k=2}^{m} \sum_{j=1}^{k-1} g(t_k - t_j \mid \theta)$$

Sum over all infection trees

Likelihood of infection tree= product of probability of each link

We maximize this likelihood to estimate the generation time distribution and subsequently the reproductive number R_j for each case. We then estimate R_0 as the mean of $R_{(j)}$ for the first x cases. Finally you bootstrap confidence intervals.

Advantages and Disadvantages of Estimation Methods

Area	Generation time method	Model-based methods
Data Needs	• Population size not required • Need infection times (removal times strictly require use of serial intervals) • ? Limited use if data have many unobserved events	• Usual to have knowledge of population at risk; can proceed without these data but becomes computationally demanding
Use	• Simplicity • No need to specify full model • Easy to calculate point estimates	• Can estimate any quantity of interest in the model • Natural incorporation of uncertainty • Computationally more demanding

Conclusions:

- Statistical methods can and should be used in modelling exercises to estimate parameters when the data are informative about the parameter.
- The choice of the method depends on the available data, the aims of the modelling exercise and slightly on preference (frequentist vs Bayesian).

- Most methods allow quantification of uncertainty (confidence or credible intervals) and even when one is not fitting a model formally some sort of uncertainty analysis is very important for modelling.
- Assessment of model fit should also be considered (visually and through formal techniques such as goodness-of-fit/AIC/DIC).

Discussion

For generation time methods, as you have mentioned, do you need to know about population size beforehand? Can we take into consideration inference for generation time methods?
For this method you don't take into consideration the population size, as it is partial likelihood.

Existence of a Dominant Network: From Global Pandemics to Small Scale Disease Spread

Marc Barthelemy, CEA, Paris - France

This talk deals with networks and their importance in epidemic spread. Infectious diseases spread among humans when individuals interact and move among their many networks (social, transportation). Characterizing these networks is therefore a crucial task for understanding the spread of infectious diseases.
In determining the relevant transportation model it is important to know what the scale of interest is. Are you interested in a pandemic or the country scale, etc.? For example: SARS would involve a global scale, while flu is of more local interest.
Beyond that, it is key to know whether or not you are in a closed system; it is usually a problem to target a system that is fully closed (a spread that is only at the country level but can be affected through networks of other levels). Another concern is that there can be a degree of subjectivity in the choice and total number of parameters, and that not all parameters are measured to the same level of detail; some might be known accurately while others are very difficult to extrapolate from the data.

We can draw parallels from the field of physics; when making a model to describe a physical system with a certain number of parameters and then compare quantitatively with the experiments. Usually when you increase the number of parameters plotted with the quantitative agreement of experiments, the curve increases towards a favorable situation. However if the parameters increase the data accuracy, the curve will go down towards an unfavorable situation.

Thus the aim is to find the ideal number of parameters for a model.

Transportation network at a global scale: the example of SARS

For this example we use the meta-population model, which considers that all cities are described by some homogeneous mixing of population and that all the urban areas are connected through an air network.

Using input data from a database, we know how many passengers travel from one airport to another etc. When describing the SARS disease we know there are many compartments and many parameters. Regarding the parameters, some were clinically estimated, but others were more empirically assumed. Furthermore, with SARS you need to take into consideration the geotemporal initial conditions—where the epidemic started.

For the purposes of our model we did a local fit in Hong Kong for the initial condition and estimated three parameters in order to reproduce the results in Hong Kong. However, the problem was that we did not have any tuneable parameters. Once the initial condition is set you can no longer play with parameters, and if you do, it can decrease the level of confidence in your model.

The results from this model showed that Hong Kong worked very well, which is not surprising since it was our local fit. It worked for many countries but for a few it did not.
However, the results are not bad since you don't expect a model to reproduce everything. For example, there might be some other transportation modes, for example Mongolia has a common border with China where transportation means other than air travel are used; the same applies for Southeast Asia, Taiwan, and so on.

A summary of predictions for the SARS model:

- Correctly predicts 23 of the 28 countries infected (5countries at no risk while infected).
- 10 countries were predicted to be at risk, while no cases reported (except Japan with about 30 cases).
- Risk and not risk were classified correctly in 205 countries out of 220.

What did we learn from this model?
The existence of some "epidemic pathways", where the epidemic is coming from. If you can identify these pathways then you can target some networks and build a strategy for control etc.

There are preferred channels for the transmission of a disease; with regards to the transportation network this is due to the heterogeneity in the number of passengers in the airline network.

Why does the model work?
Homogenous mixing in the urban areas seems to work; the model correctly captured the relevant heterogeneities; and there is a dominant network at the global scale which is air travel.

So although with this model is at a global scale it is rather simple. More complications arise at the smaller scale, as with the flu model described below.

Small scale model: flu
First we tried to identify the epidemic pathways by searching the epidemic data. Once the pathways were identified they were correlated to transportations networks in order to find dominant networks.

For example, in the USA, comparing data between states means determining a person correlation coefficient between the two profiles to obtain a value. This value intrinsically has no meaning; it needs to be compared with some other data, and you have many constraints in this system (epidemic period, spatial…etc.). In fact you need to eliminate all the correlations induced by constraints (null model) to identify the correlation that is actually induced by the movement of people. This gives you a new number for the correlation coefficient, and allows you to compute some indicator which gives you the part of the correlation induced by travel, in other words the measure of travel-induced epidemic spread.

The concern with this model is to be able to relate the pathways to transportation modes. This can be done by performing a multivariate analysis taking into consideration air travel, (distance, temperature…). After performing the multivariate analysis you can plot the information and obtain dominant networks (air travel prevails vs. ground transportation, etc). Thus you can identify actual dominating mode pathways for your model.

We can conclude that we have epidemic pathways at all scales, and if you have very strongly defined pathways this increases the predictability. However, although we have epidemic pathways, they are not always governed by one single mode. This depends on the scale (global, local…). At the global scale the mode is normally air travel; for the USA it is air travel mainly, but interstate roads to some degree. In smaller countries like France there are many equivalent modes such as train, car etc. So the model needs to determine if they are actually dominant transportation modes or not.

The key points are:

- To identify the relevant scale, and if it is really a closed system. It can be very dangerous to focus on a closed system if this one is not (focus only at the city scale if that is not the case for disease spread).
- To identify the relevant transportation mode(s).
- To acknowledge the existence of strong heterogeneities.

The last comment is regarding the convergence of models (average over different models), a technique being used today in climate change research. But it might be also useful for modelling in epidemiology.

Discussion

You mention the difference between the gravity model and the model based on air transportation data, but don't you think there is scope to mix these two types of models?
If you have all data for all transportation types, such as train, car, air etc, then that is fine. The gravity model tends to be used when you don't have data. However, the gravity model does not have strong theoretical foundations, so if you can avoid using this model then it is better to do so.

From Model to Public Health Decision: The Example of Chikungunya Outbreak in Réunion Island, 2005-2006

Pierre-Yves Boëlle, Antoine Flahault, Alain-Jacques Valleron
Inserm UMR-S707, Université Pierre et Marie Curie, Paris – France

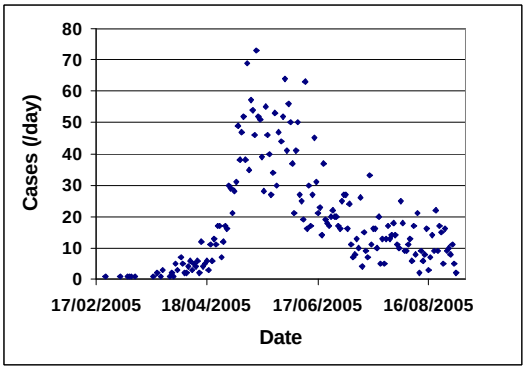
La Réunion is a small island located in the Indian Ocean with a population of 780,000. Chikungunya is an arboviral disease, transmitted by mosquitoes (*Aedes albopictus* /*Aedes aegypti*). The disease was first described in Tanzania in 1953, and was first isolated in La Réunion in February 2005.

Chikungunya is generally a mild disease and has acute and chronic presentations:

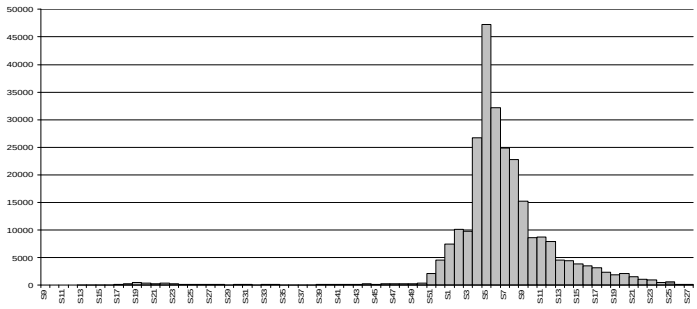
- Acute : fever, arthralgia
- Chronic : arthralgia

The acute presentation is better known than the chronic: about 10 to 20% of infected patients still report problems 18 months after the initial infection.

This epidemic graph shows the path of the disease during 2005:



The epidemic caused a little more than 100 cases per day at its peak and started to fall off as winter came. The reason for this fall was likely due to the low temperatures, but also because the mosquito in Réunion was said not to be a good vector of the disease. A second strong epidemic occurred in 2006 (see graph).



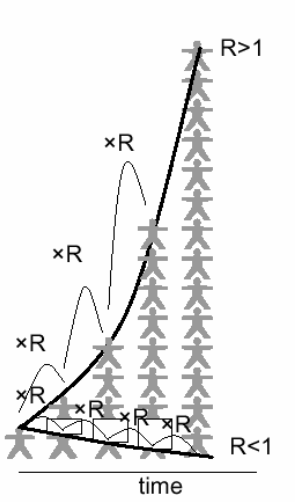
There were many cases in this epidemic and it showed presentations that were not known before, such as in children.

In summary:

- More than 266,000 cases, which is a high percentage of the total population of the island.
- 255 death certificates listing Chikungunya as the cause of death
- 40 vertical transmissions in newborns
- 2-4% of cases were hospitalised
- Deaths were due to the disease and not to the combination of any other factors.

How can we measure the epidemic potential of Chikungunya?

- Use the reproduction ratio (R) of average number of secondary cases per index case. The first R gives you an idea of the amplification during epidemics.
- We also used the generation interval (GI), which measures the lag between onsets in index and secondary cases.



R may be estimated from the epidemic curve and generation interval (Wallinga Am J Epid 2005).

Since little was known about the mosquito, in order to estimate the reproduction ratio we focused on transmission from one human host to other humans. This helped us to determine the most likely transmission, e.g if your GI is medium sized, transmission is likely to have come from someone who is not necessarily close to you.

Estimating the reproduction ratio

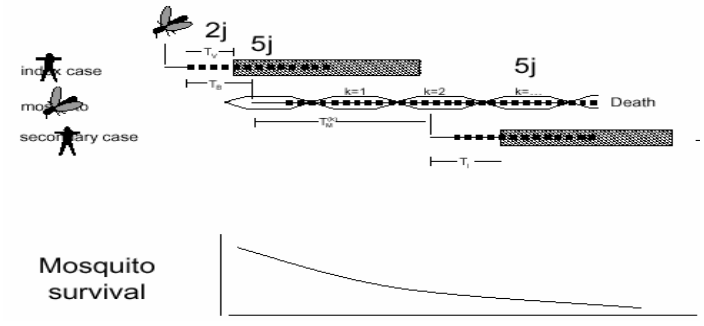
1. Bypass the mosquito
2. Probabilistically impute each transmission according to GI
3. Compute average number of secondary cases

This provides an epidemic curve almost in real time, as well as a generation time distribution (this number was constructed from what is known about the transmission process).

Calculating the generation interval (GI)

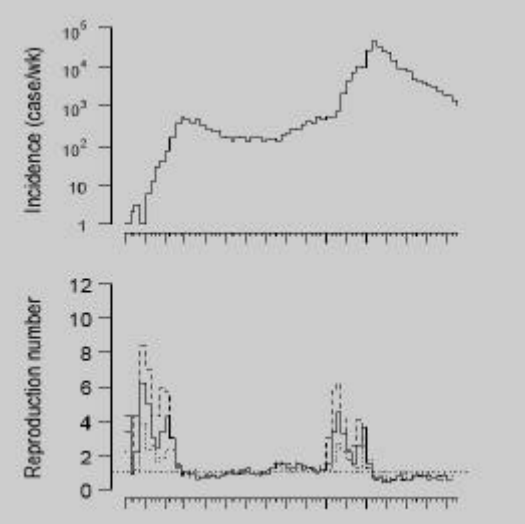
For a mosquito to become initially contaminated with the virus, it needs to bite an infected human during the viremic period, which is about 7 days. This period starts one or two days before symptoms appear. Once the mosquito is infected, it continues to bite and lay eggs in cycles. It bites about every 5 days depending on temperature; another aspect taken into consideration was the mosquito's survival.

.



Taking into consideration that the mosquito's survival will be affected by time, and combining all the different information, a number of hypotheses were made about how the GI might look between the index case and a secondary case. We used different scenarios which varied the length of the infectious life of the mosquito, as well as introducing latency into the mosquito.

This information was then plotted onto an epidemic curve to estimate the "R" for Reunion Island. This revealed that at its peak, R was approximately 4; during the winter R was slightly below one. (see graph below).

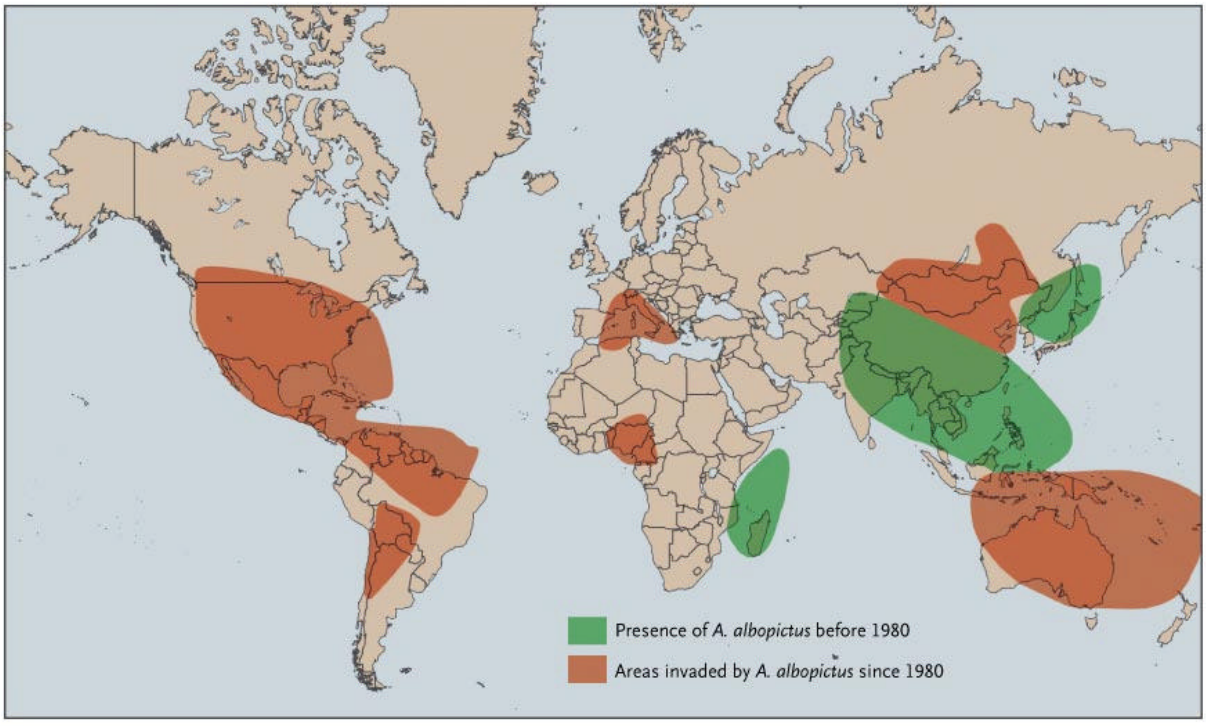


Because the magnitude of cases was so different between the 2005 and the 2006 epidemics, the genomics of the Chikungunya virus were investigated to try to explain these differences (Schuffenecker I et al, PLoS Medicine 2006). These studies found that there was a change in the genus of the virus. The virus behind the first epidemic was different from the first; there was a sequence switch from A226 to V226.

Other research was done by social scientists into perceptions of the patterns of transmission. The results showed that most people believe that the mosquito was the main transmission mode; not much credit was given to human-to-human transmission. Other social studies highlighted other beliefs in the causes of the epidemic, such as the tsunami, or the fact that some native species of flowers which accumulate water played an important role in mosquito reproduction and epidemics.

Control strategies such as pest control started late (February 2006) against a background of skepticism and hostility on the part of the population. Citizen movements accounted for programs like Kass Moustik for mosquito prevention.

An important concern is that the East-Africa strain of Chikungunya has spread since 1980 to many other geographic areas, due to vector spread as shown in the following map:



This possible evolution of the disease is very concerning and a serious public health issue. For example, the mosquito vector *A. albopictus* could survive at latitudes as high as Stockholm.

Conclusions

- Chikungunya fever was properly monitored and documented, almost in real time.
- Nevertheless, the disease still developed into a major public health problem a year later. This was because of the risk perception on the part of health authorities, coupled with the lack of proper analysis of surveillance data.

Discussion

**I was surprise to see in your map that in Central Africa the disease is only present in Cameroon. I wonder if there are some places where the mosquito might be present but which have not been investigated?*

It is highly probable that the mosquito is present in other places, but it might not have become the main species, as it is highly dynamic.

**Is this the same vector of Dengue fever? Have these populations had a problem with Dengue in the past?*

Yes, the same mosquito is competing for both Chikungunya and Dengue; yes these populations have suffered from Dengue.

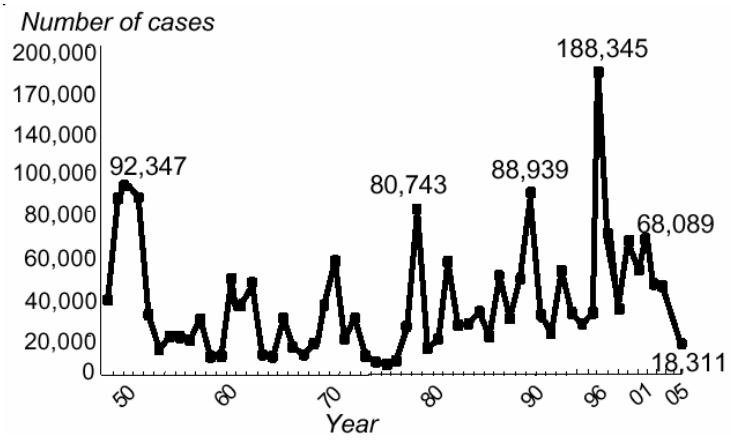
3. Session II: What is the Expected Public Health Impact of Model Approach?

Expected Effects from the Introduction of a Meningococcal A Conjugate Vaccine in Sub-Saharan Africa

Marie-Pierre Preziosi and Marc LaForce, World Health Organization, Geneva, Switzerland.

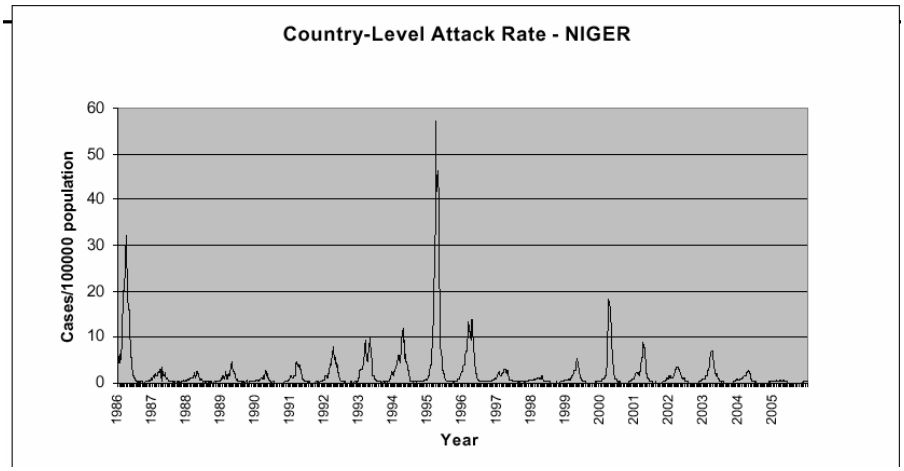
There is a specific meningitis epidemic area in Africa called the Meningitis Belt, which stretches from Senegal (West Africa) to Ethiopia (East Africa). Meningitis is caused by the bacteria *Neisseria meningitidis*.

Epidemic meningitis has been present in Africa for at least a century. The following graph shows the epidemic curve for the last 50 years.



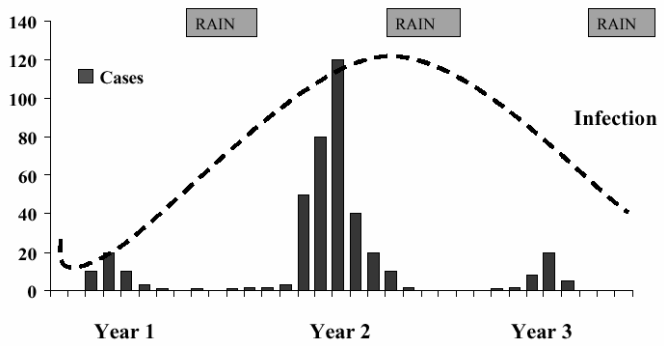
The characteristic feature of this chart is the usually strong epidemics occurring every 8 to 12 years, interspersed by yearly epidemics. The last of the huge epidemics was in the mid to late 1990s, and that is what highlighted the urgent need for a vaccine against this disease. This has led to the Meningitis Vaccine Project, which is developing a Men A conjugate vaccine.

When looking at the disease epidemic charts for a given country, for example Niger, which is at the heart of the Meningitis belt, you observe yearly epidemics and a huge epidemic peak every ten years. (See graph below).



Although the disease appears every year, within the year it is very seasonal; cases only occur during the very dry seasons and completely disappear during the rainy season. On average an epidemic lasts 2 to 3 years—the spread of a new clone through the population takes about this time—with no cases during the rainy season.

Seasonality of meningococcal meningitis



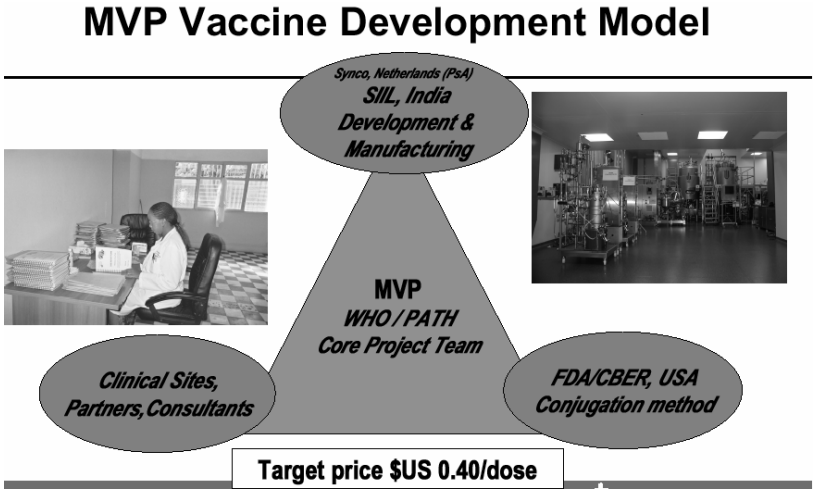
The dusty wind in Africa's dry season is very strong, like a cloud over the cities, and this might be one of the main reasons for the spread of the disease during the dry season.

The bacteria *Neisseria meningitidis* colonises people's nasopharynxes, and is found in about 10 to 30% of asymptomatic carriers. Most of its life is spent spreading from one individual to another, going through full cycles of acquisition, invasion, colonisation, release etc. Most cases probably occur during the dry season, because the bridges of the nasopharynx mucosa are more fragile due to this dry wind.

The Meningitis Vaccine Project - MVP

- Created in 2001 after the huge epidemics that occurred in the mid-late 1990s in the meningitis belt.
- Funded by a grant from the Bill & Melinda Gates Foundation, as a 10-year partnership between WHO and PATH.
- Its goal is to eliminate epidemic meningitis as a public health problem from Sub-Saharan Africa through the development, testing, licensing and widespread use of affordable conjugate meningococcal vaccines.

Figure: Vaccine Development Model



The conjugation method was develop in a US laboratory and transferred to a developing country manufacturer in India which is developing the product for trials. All clinical trials staff are located in Africa.

The titers obtained four weeks post-vaccination in toddlers, when looking at the 4-fold responders results, showed excellent response. Ninety-six percent of the subjects receiving the vaccine had the 4-fold response (see table).

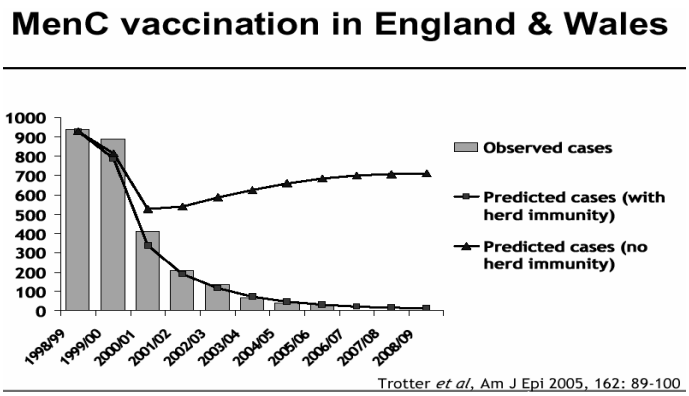
		N	%	[95% Confidence Limit]
PsA-TT	198	190	96 *	[92; 98]
PsACWY	193	123	64 *	[57; 71]
Hib-TT	194	69	36	[29; 43]

* ? = p (PsACWY) – p (PsA-TT) = -32% [-40%; **-25%**]
Primary endpoint of non-inferiority achieved ("immunological superiority" demonstrated here)

When comparing the titer results for the children that received the conjugate study vaccine with those receiving the conventional polysaccharide vaccine, a huge titer response was observed for the study vaccine. These results, obtained last summer, have given the MVP team strong hope in the quest to eliminate epidemic meningitis from this region.

Why are conjugate vaccines *so* different?
Conjugate vaccines are quite different from polysaccharide vaccines: in addition to providing individual protection from a serious disease, they also prevent the asymptomatic carriage of the disease. Reduced carriage means reduced infectiousness, lower transmission and indirect protection. It means that even unvaccinated individuals in the population are at reduced risk of disease.

There are several examples of conjugate vaccines working better and producing better immunity than the standard polysaccharide vaccines. These examples include the PNC 7 vaccination in the USA and the Men C vaccination in England and Wales. These examples show mainly that without herd immunity it would not be possible to achieve such good vaccine immunity within a given population.
.



Immunization Strategies for the MVP
Current immunization strategies are mainly reactive, since the vaccine compound is still not at its best. For the moment the strategy is based on detecting cases, confirming and vaccinating at-risk populations.

This is based on interventions thresholds, such as:

- alert threshold (5cases/100,000/week) for the confirmation of serogroups; or
- epidemic threshold (15 cases/100,000/week) for mass immunization campaigns.

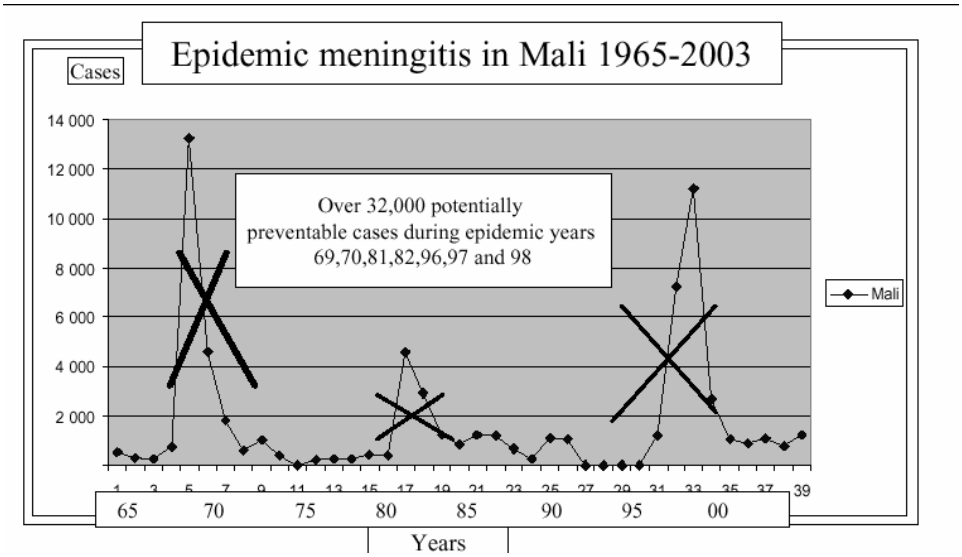
However, this is not an efficient way of preventing epidemics. Thus, the MVP expects to:

- Conduct mass vaccinations for 1 to 29 year olds with a single dose of Men A conjugate vaccine to induce strong herd immunity.
- Protect birth cohorts with Men A conjugate vaccine either by follow-up campaigns every 5 years for 1-4 years olds, or by a routine single dose at 15 months, or by routine immunization in infancy.

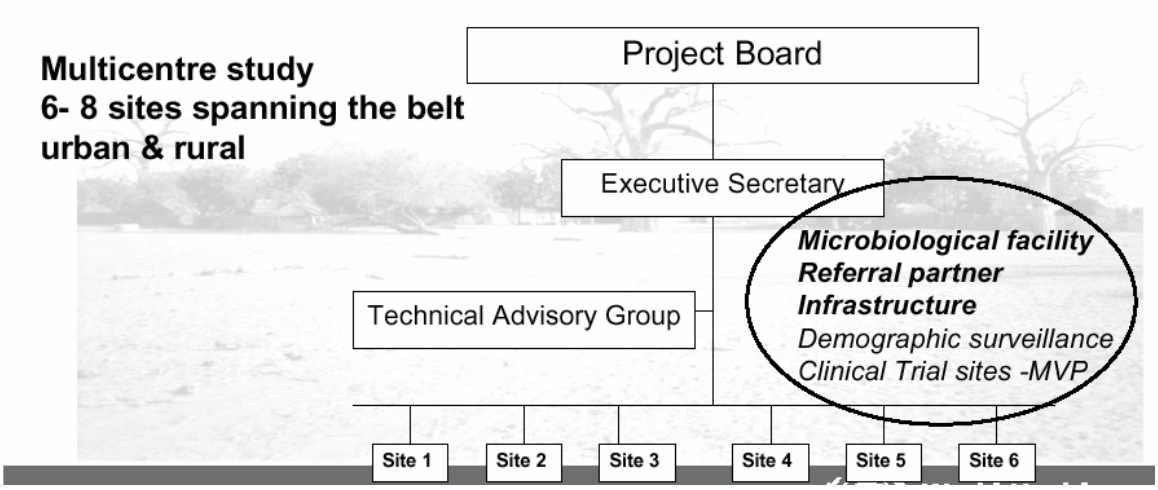
This part of the project is ready to start, as it is entering phase III. It is expected that by 2009 the vaccine will be introduced into a larger area of the belt. The project expects to achieve its objectives by 2013.

Potential public health benefits from preventing Men A epidemics

By using predictive models after the introduction of the conjugate vaccine and having a retrospective look at prior epidemics, such as the one in Mali, we roughly estimate that there were over 32,000 potentially preventable cases during the epidemic years from 1968 to 1998. (See graph below).



In order to estimate the impacts of the conjugate vaccine, we need appropriate data to do the modelling. We are therefore building a parallel project—the African Meningococcal Carriage Consortium—which intends to get insights into the direct and indirect effects of the vaccine (susceptibility, colonization, transmission, etc.).



Discussion

**What is the immunization schedule? Do you like to give this vaccine during the extended program of immunization, one month apart? Because if you give just one dose of a conjugated vaccine, even if good immunization is achieved, the beauty of a conjugated vaccine is the booster dose.*

Now it is one dose for those above one year of age. With our results from the phase II trials we think we have persistence of antibodies for long enough to introduce one dose in the older population (those above one year of age) in a mass national campaign. Then comes the question of how to maintain the immunity in the population and in the youngest. We don't have an answer for this yet.

Impact of Combined Effect of Vaccine and Decreased Antibiotic Use of S. Pneumoniae Susceptibility to Antibiotic

Guillemot Didier, Sanofi Pasteur, Paris – France

S. pneumoniae is a common human pathogen which is responsible for diseases such as otitis, pneumonia and meningitis. It causes about 3.5 million deaths a year worldwide, and up to 50% of colonized children are asymptomatic carriers.

There is widespread antibiotic resistance to this pathogen, though this varies between countries. In France, there are more than 60% penicillin-resistant strains, and more than 50% multiresistant strains (penicillins and macrolides). This is largely due to the patterns of antibiotic consumption in France, the largest antibiotic consumer in the European Community.

The conjugate vaccine has been in development since 2000.

Antibiotic-Pneumococci interactions

Antibiotics and Pneumococci interact at three different levels, leading to resistant strains:

- Gene level: leads to mutations, gene transfers etc., that lead to the emergence of new mechanisms of resistance by the strain.
- Individual level: colonization of individuals, ecosystems (gut, skin, nasal..) which leads to competition among strains. This ends in the death of susceptible strains and the survival of resistant ones, and creates selection in individuals.
- Population level: cross transmission that leads to selection and spread in populations. Susceptible and resistant strains are transmitted to individuals and depending on the level of antibiotic exposure, susceptible strains will die and resistant strains will survive.

Modelling antibiotic resistance

The first reason to model antibiotic resistance is to achieve better understanding of underlying processes in resistance selection and to predict future changes. In turn these help to prepare and evaluate control measures such as reducing antibiotic consumption, developing new therapies (vaccines, new drug molecules) and increasing prophylactic measures, etc.

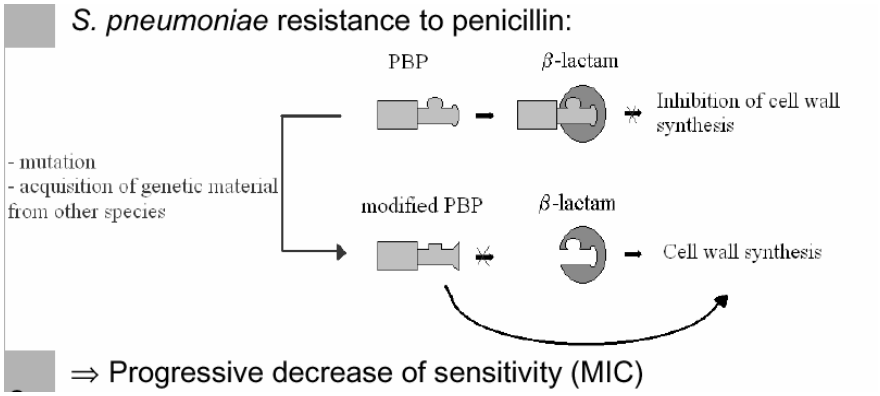
There are different types of models of different phenomena already published,[1] such as:

- resistance at the intra-individual level (bacteria evolution)
- resistance at the inter-individual level (bacteria diffusion)

However, none of the models tried to couple inter and intra-individual levels and all were developed for generic antibiotics and bacteria rather than specific antibiotic/bacteria/resistance mechanism combinations.

Thus, a model was developed for the selection of pneumococcal resistance to penicillin in France: a compartmental model (Temime L, Boëlle PY, Courvalin P, Guillemot D; Emerg Infect Dis, 2003; Temime L, Boëlle PY, Thomas G; Math Pop Studies, 2005).

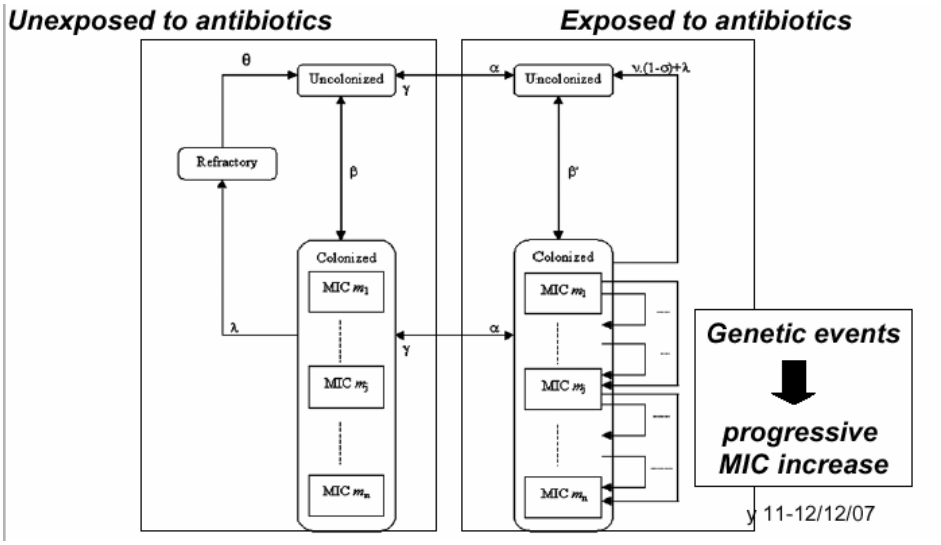
The mechanisms of pneumococcal resistance to penicillin involve the progressive modification of PBP (Penicillin B Proteins) and other factors. This implies that the susceptibility of pneumococcus to penicillin will progressively decrease, rather than suddenly becoming resistant.



The model was built to reproduce both the intra-individual selection of resistance according to the above mechanism and the inter-individual transmission of susceptible and resistant strains. This led to the creation of a compartmental model in which colonized compartments were structured according to the MIC, which describes the level of susceptibility to penicillin.

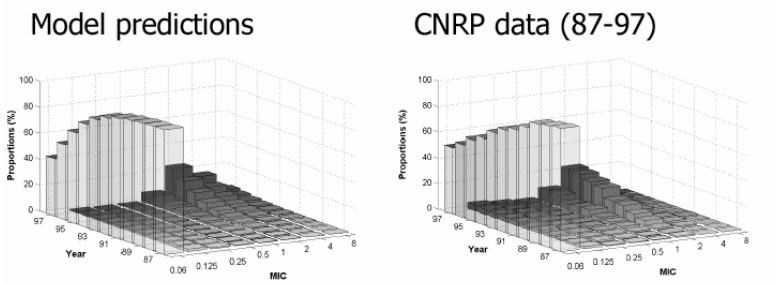
[1] Bonhoeffer S et al., 1997; Sébille V et al., 1997; Austin D & Anderson R, 1999; Lipsitch M et al., 2000; McCormick A et al., 2003.

The Model



In order to validate this model, its predictions were compared with independent historical data on penicillin resistance. The model echoed the historical emergence of the first resistant strains about 20 years ago (after the introduction of penicillin for general consumption), and the MIC distribution.

.

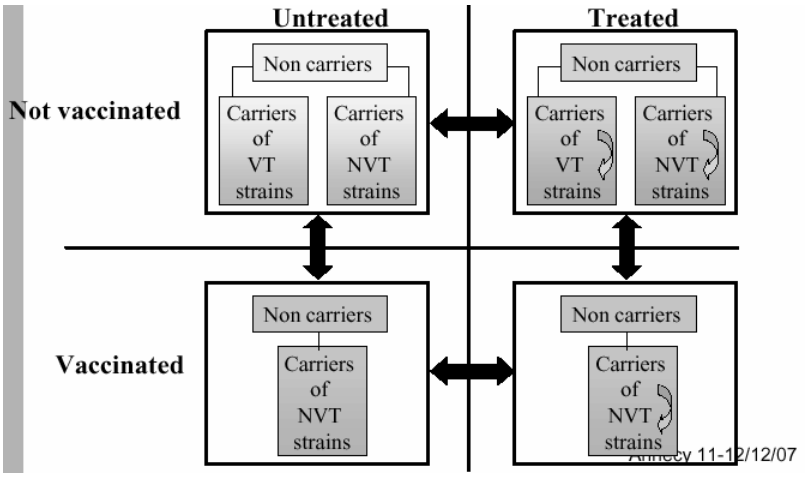


Investigating the impact of the conjugate vaccination

Conjugate vaccines:

- Protect against carriage and invasive disease, unlike other vaccines
- Cover 7 to 11 of >90 serotypes
- Decrease of carriage of vaccine serotypes/all serotypes and of pen-R strains according to efficacy studies (Dagan & Fraser, Pediatr Infect Dis J, 2000)
- Used in the US since 2000 for children <2 yrs old
- Show observed reduction in IPD incidence: -69% (Whitney et al, NEJM, 2003).

Based on this first model, a two serotype, age-stratified compartmental model was built in order to simulate the impact of this vaccine.[2] This means that individuals could be exposed or not to penicillin, also vaccinated or not, see following graph of the model.



One of the important results obtained with this model is that it predicts serotype replacement; thus in the long run the vaccine can decrease vaccine strain types, but increase non vaccine strain types (strains that are not part of the 7 serotype strains that are in the vaccine). This is applicable for carriers but not for invasive disease.

The model shows a slight decrease in overall carriage following the introduction of the vaccine; the higher the vaccination rate the more marked the transient decrease of carriage. In the long term immunization overall carriage is unchanged.

Conclusions

Need for data:

- Specificity mechanism/antibiotic/pathogen
- More complex models which require more complex data:
  - Mean characteristics in the population: duration and frequency of antibiotic exposure, infectious contact rate
  - Mean characteristics of the micro-organism: duration of colonization, susceptibility to antibiotic exposure, invasiveness, fitness
  - Resistance mechanism characteristics
  - Host characteristics: immunity, dynamics of competition

In summary, antibiotic resistance modelling can only be satisfyingly achieved through collaboration with microbiologists, physicians, etc.

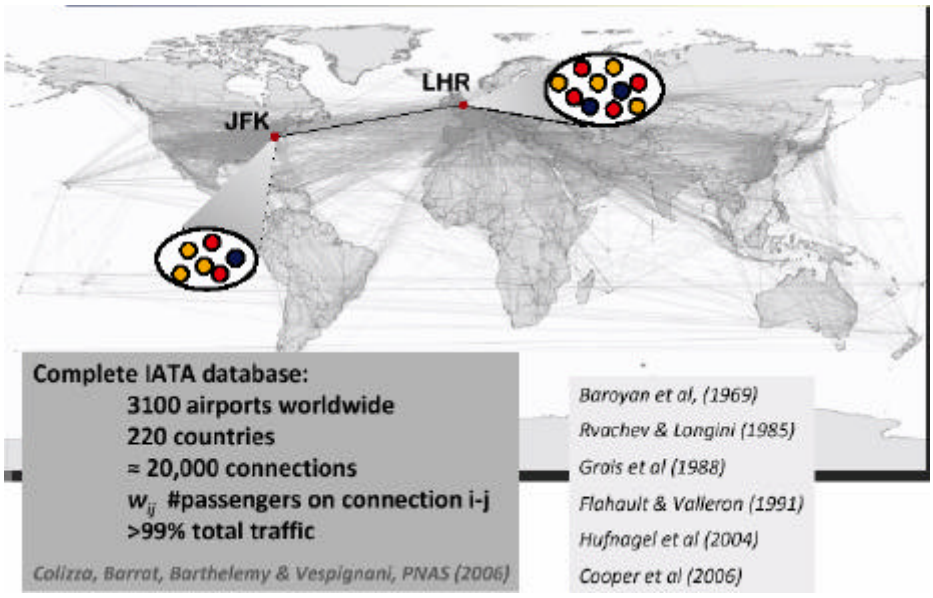
[2] Temime L, Guillemot D, Boëlle PY; *Antimicrob Agents Chemother*, 2004)
(Temime L, Boëlle PY, Valleron AJ, Guillemot D; *Epid Infect*, 2005)
(Temime L, Guillemot D, Boëlle PY; *Pediatr Infect Dis J*, 2006)

Managing Anti-Viral Resources to Control the Global Spread of an Emerging Pandemic

Victoria Colizza, Complex Networks Lagrange Laboratory, Turin – Italy

Many studies focus on containing pandemic influenza at source, and look at the different measures for containing it at source. However, we need also to be able to predict what might happen if it is not contained at source. Even if all the control strategies to contain the pandemic at source are implemented, there is still the possibility that an infected person could get on a flight and travel elsewhere.

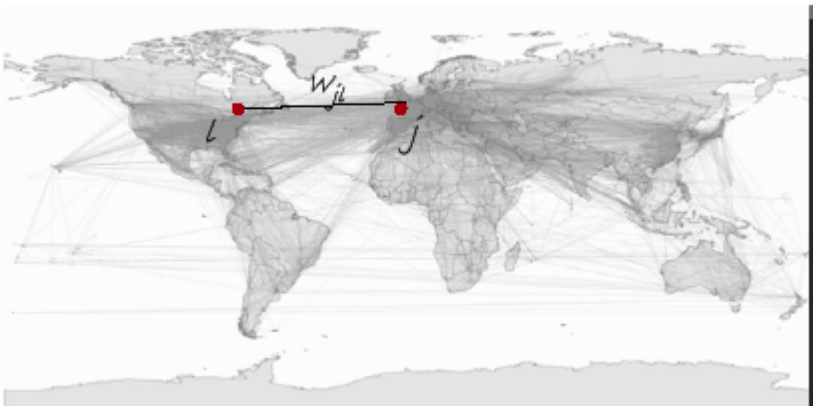
We have built a model focusing on the international spread of pandemic flu. It is a metapopulation model in which the patches of the model are the populations living in cities and the whole structure of the model is the airline transportation system.



This model was first developed in the late sixties during the cold war, and others have been used for seasonal applications of influenza, SARS etc. The differences between those approaches and the one presented in this talk are that we use the complete International Air Transport Association (IATA) database, which includes all the airports in the world, total number of passengers, etc.

This makes the model highly complicated. The question is why do we need this much information? This is because knowing this amount of information and applying it to the model will affect the epidemic pattern that will be observed, and, more importantly, the reliability of the predictions. So we took into consideration the cities surrounding each airport and all possible flights leaving from that airport.

Inter-cities discrete stochastic air travel model



The W_{jl} represents the number of passengers given from the input data that are flying from point *j* to *l*. So the probability that an individual in compartment X travels from *j* to *l* looks like this:

$$p_{jl} = \frac{w_{jl}}{N_j}$$

census data for 3100 urban areas surrounding IATA airports

The probability is proportional to the traffic flow and inversely proportional to the population.

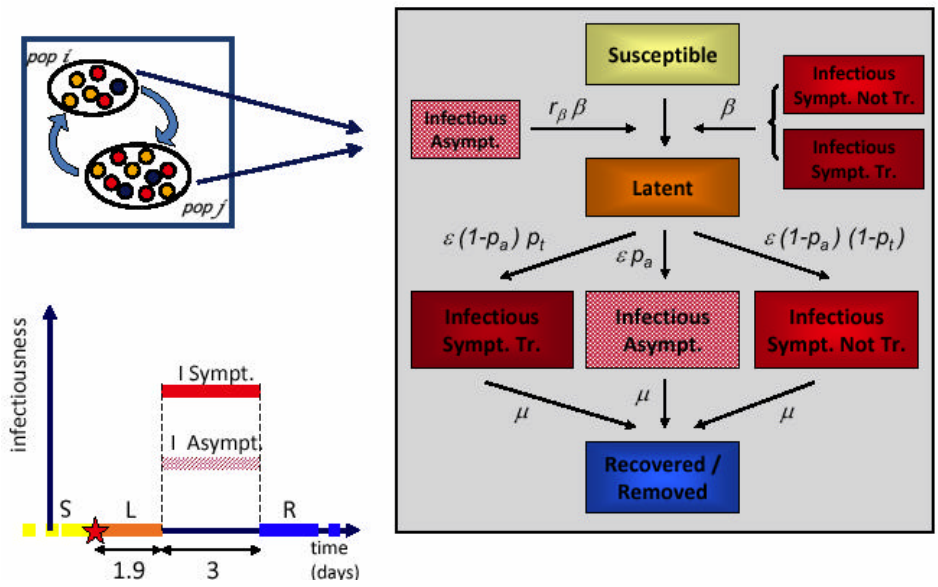
When considering multiple possible destinations, the multinomial extraction of travelers $?_{jl}$ in each compartment is as follows:

$$P(\{\xi_l\}) = \frac{X_j!}{(X_j - \sum \xi_{jl})! \prod \xi_{jl}!} \prod_l p_{jl}^{\xi_{jl}} \left(1 - \sum_l p_{jl}\right)^{(X_j - \sum_l \xi_{jl})}$$

All this simply defines a stochastic travel operator, which is the total sum, the total number of passengers going to destinations and coming from destinations. Stochastic travel operator: net balance of influx – outflux.

$$\Omega_j(\{X\}) = \sum_l \left(\xi_{jl}(X_j) - \xi_{lj}(X_l)\right)$$

Intra-City: Discrete Stochastic Infection Dynamics



In this model we tried to incorporate every possible detail of travel behavior within a compartmental model (see above figure).

The model distinguishes between people who are symptomatic and treated, infectious but asymptomatic. and infectious and symptomatic but not treated. It also distinguishes between those who are allowed to travel and those who are not allowed to travel due to their symptoms.

The model's predictions are based on probabilities, such as the probability of being asymptomatic (33%), probability of travelling (50%), etc.

Below is an example of the equations involved in the metapopulation pandemic model:

metapopulation pandemic model

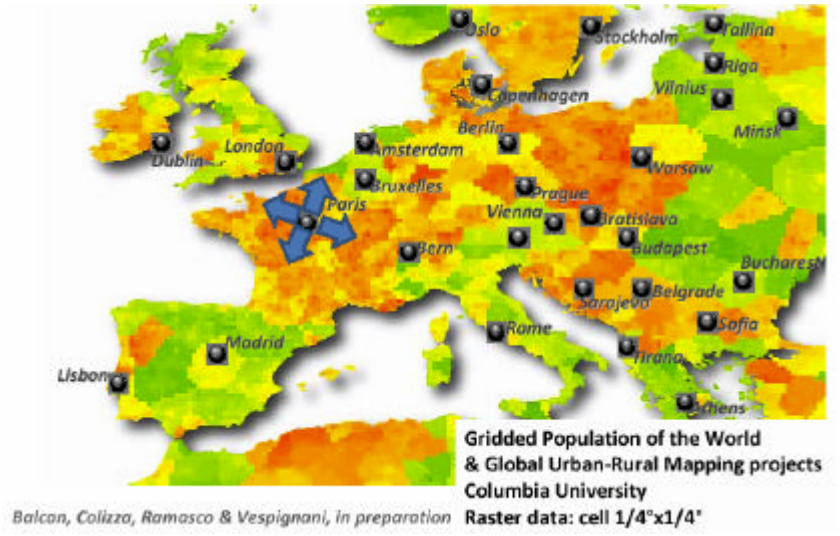
e.g. equation for latents:

$$L_j(t+1) = L_j(t) +$$
$$+ mnldev\left(S_j(t), \beta I_j^t / N_j, \beta I_j^{nt} / N_j, \beta r_\beta I_j^a / N_j\right)$$
$$- mnldev\left(L_j(t), \varepsilon p_a, \varepsilon(1-p_a)p_t, \varepsilon(1-p_a)(1-p_t)\right)$$
$$+ \Omega_j(\{L\})$$

coupling terms among cities == air travel

3100 x # compartments coupled stochastic discrete difference equations for the spatiotemporal evolution of the disease

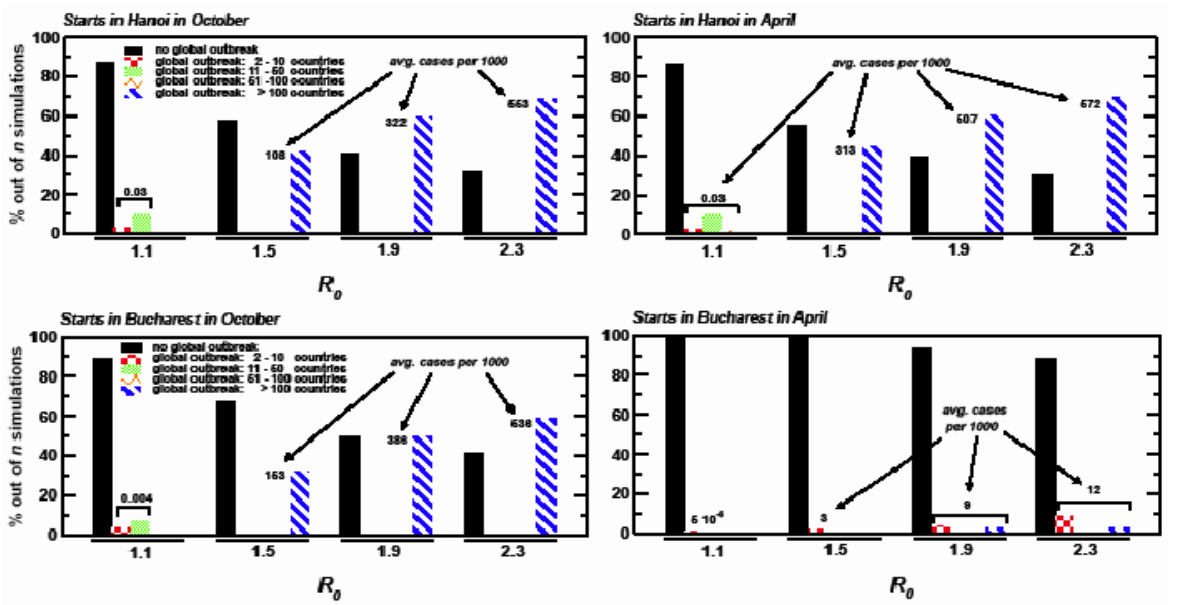
The following results are based on UN census data information, which considers only urban areas. These global data are good when considering air travel at a more macro level, but when more precise data are needed on the means of transportation, census data from Columbia University are much more detailed as they place the world population into gridded cells.
These cells are of different sizes and give a sense of distance through color coded zones.



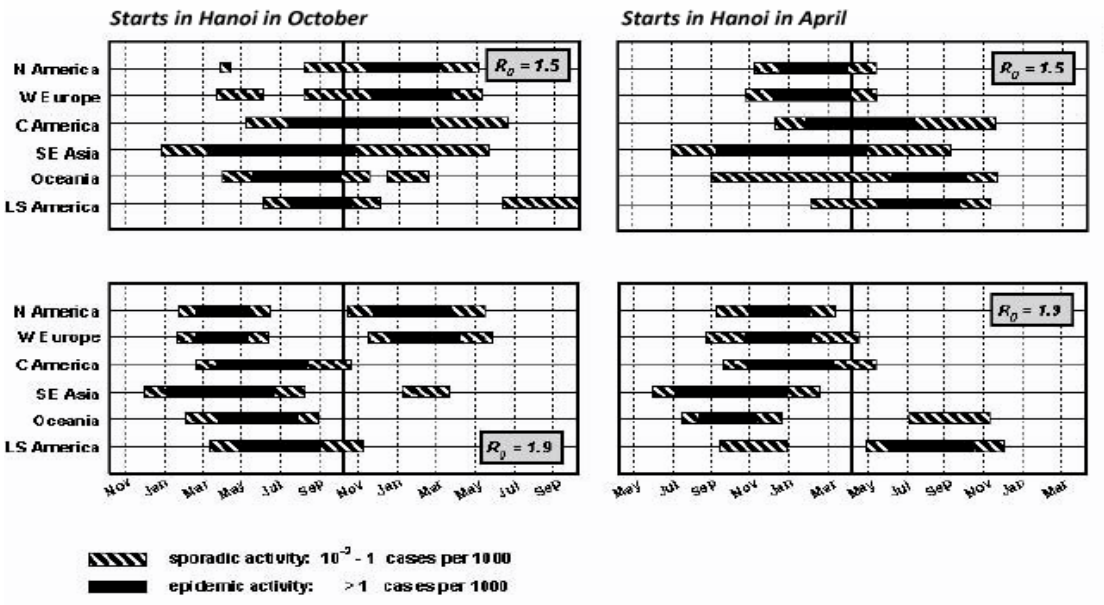
With this kind of detail it is possible to perform a more precise simulation. You will also need to include additional means of transportation other than air travel in order to include rural areas.

Another feature that needs to be included is seasonality; for this we divide the world into zones corresponding to the tropic of Cancer, the equator and tropic of Capricorn. The world's cities are then divided among these different geographical areas.

The model allows us to decide in which season and in which city we would like to start from. We can produce graphs showing the probability distribution of having an outbreak that is contained at source (the first bar in the graphs below); the second bar represents an outbreak affecting 2 to 10 countries; the third an outbreak in 11 to 50; 50 to 100, and more than 100 countries:



The model also allows us to determine the evolution of epidemic activity in different countries. When will each of these regions be affected by pandemic flu?

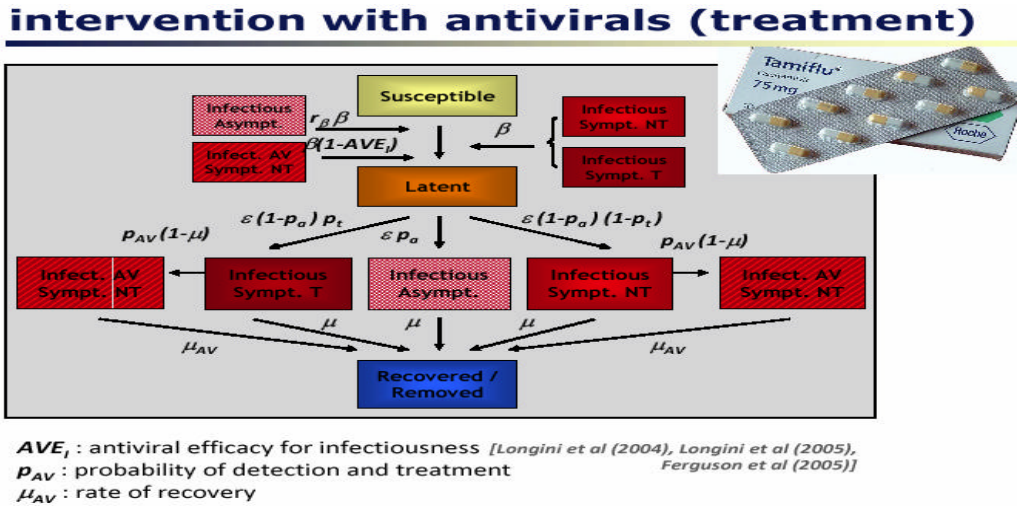


Seasonality plays a very important role in the path of the epidemic, as well as travel, which includes the number of connections and other details related to travel.

The above examples are baselines; they assume that no control or prevention measures are being taken by the government, etc. This baseline case can be used as a reference scenario for assessing the impacts of possible interventions.

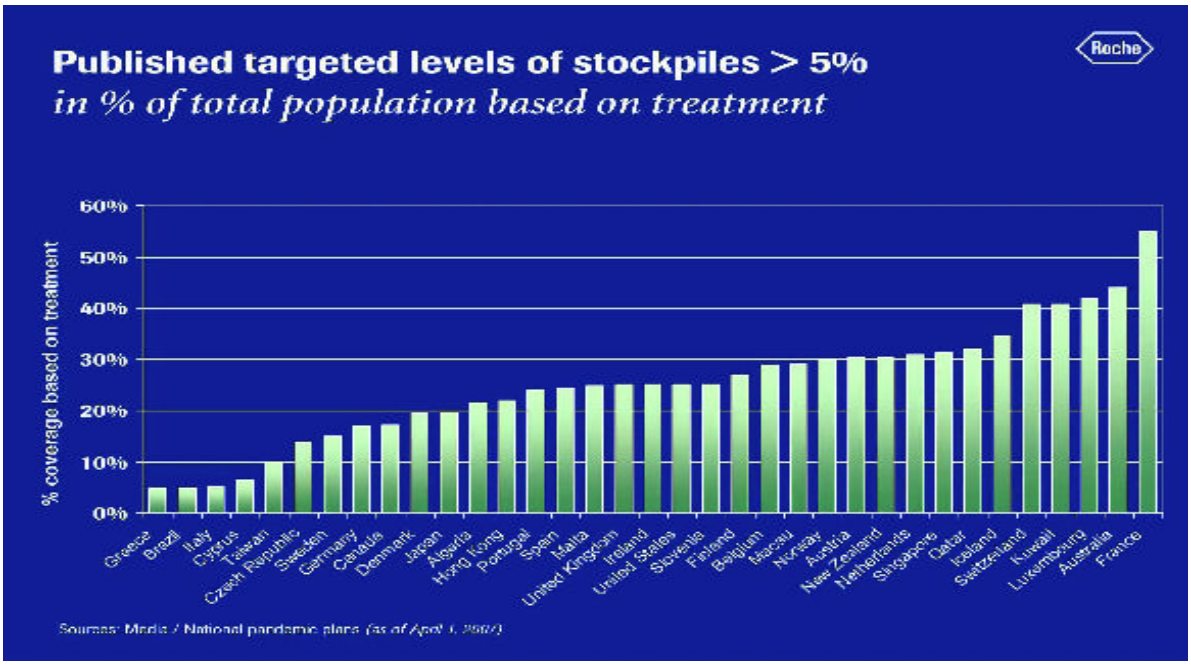
Intervention with antivirals

One intervention that can be explored in this model is antiviral treatment. For this purpose the compartmentalization of the model needs to be modified so as to include the fact that some of the clinically infectious individuals are detected and given treatment. This implies reducing infectiousness, and also reducing the total period of infectiousness for a treated person by one day.



PAV is a very important feature of the model. Once this information is included, a simulation can be run starting with baseline conditions as before and then comparing the results.

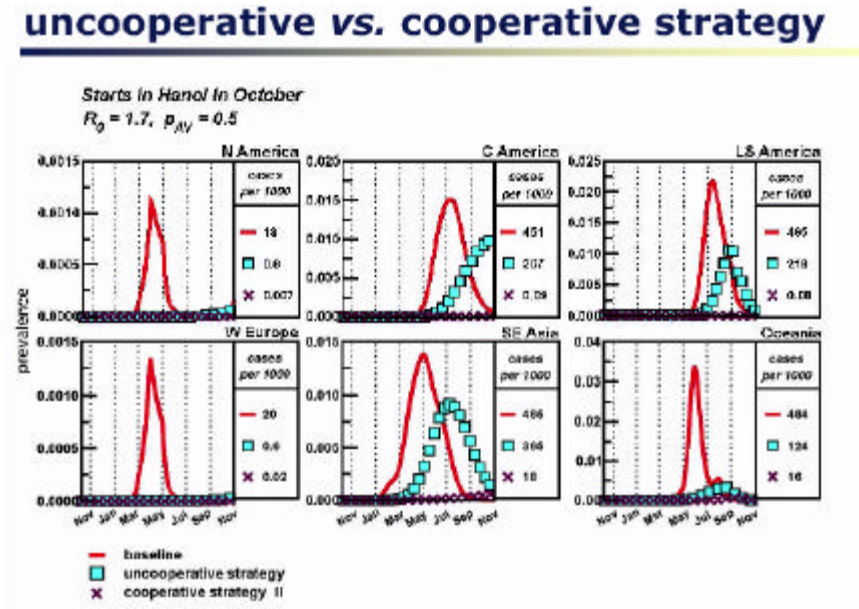
In order to run the model appropriately it is first necessary to have information on antiviral stockpiles and how they are distributed among countries.



To model a situation that is more conservative than that implied from the above data, certain countries were assumed to have fewer stockpiles of antivirals.

Cooperative vs. uncooperative strategy

In this simulation, "cooperative" implies that countries with bigger stockpiles will be willing to give part (5-10%) of their stockpiles for use in other countries if necessary. The "uncooperative" scenario is when a country will only use their stockpiles for their own population.



A question for stockpile donor countries is what do they gain by giving away their stockpiles? The answer is that there is a strong reduction of cases and of the chance of a pandemic outbreak.

The final step is then to see how to implement these strategies in reality.

Discussion

**Making the homogeneous mixing assumptions within cities, which I think you are doing, those epidemics grow faster, so it seems like you are a little pessimistic?*
Yes is true, we are little bit pessimistic taking this type of population mixing.
**Your simulations seem to suggest that the best strategy even for donor countries is to cooperate, but in reality do you know how to address a government in order to see the advantage of given away part of their stock piles?*
No we have not done that.

Health Economic Evaluation of Vaccine: The Example of Varicella-Zoster Virus

Benoît Dervaux, CNRS, Catholic University of Lille, Lille-France

When looking at resource allocation in health care, the four most important aspects to consider are safety, efficacy, quality and efficiency. In economic terms, efficiency is the "value for money" of new interventions like vaccines.

There are different types of economical evaluations for determining how health consequences are valued:

- Cost-effectiveness analysis (CEA): uses natural units such as number of deaths, or hospitalizations, etc.
- Cost-utility analysis (CUA): uses QALYs (quality-adjusted life years), for example to determine the life path of individuals in relation to indices of quality life standards.
- Cost-benefit analysis CBA): evaluation in monetary terms.

Each type of analysis gives you a different answer and a different perspective.

For example, the answers obtained from a cost-benefit analysis are more general and can guide decisions about whether or not to do something. Cost-effectiveness analysis is more precise: you do something this way because it is the best way to do it.

Are vaccines different from any other health care intervention such as drugs? In some sense vaccination has direct and indirect effects. These effects may be positive, such as creating herd immunity; or negative, such as exerting selection pressure on certain pathogen strains. The epidemiological impacts occur over the very long term, and this time period has to be taken into consideration in analysis (sensitivity of the results to discounting rules).

Economic evaluations are done through modelling. Modellers and economics professionals can work together to build economic values into the model.

One challenge with vaccination models is the difficulty of measuring them within the population as a whole, as often it is the population of children to which QoL (Quality of life) applies. For this reason the model needs to be able to value indirect costs. Another challenge in the evaluation is the fact that some diseases are eradicable.

So these are some key elements to take into consideration in the economical evaluation of vaccine models:

- Herd immunity, externalities
- Discounting, choice of end-points and time windows: prevention vs. cure
- Risk of underestimation of the value of vaccines, if we don't take into account the two above factors.

Discounting

There is a big debate on what to take into account and what not to. In discounting we normally give more weight to short term consequences than long term ones.

- How we discount when evaluating a vaccination intervention? Should the health consequences be discounted?
- Do time preferences depend on decision time horizon? Should we use exponential or (quasi) hyperbolic discounting models?
- What is the real objective of preventive programs? Should we consider risk reduction or long term consequences on morbi/mortality as endpoints of vaccination programs?
- Are time preferences identical for all commodities? Should cost and health consequences be discounted at the same rate?

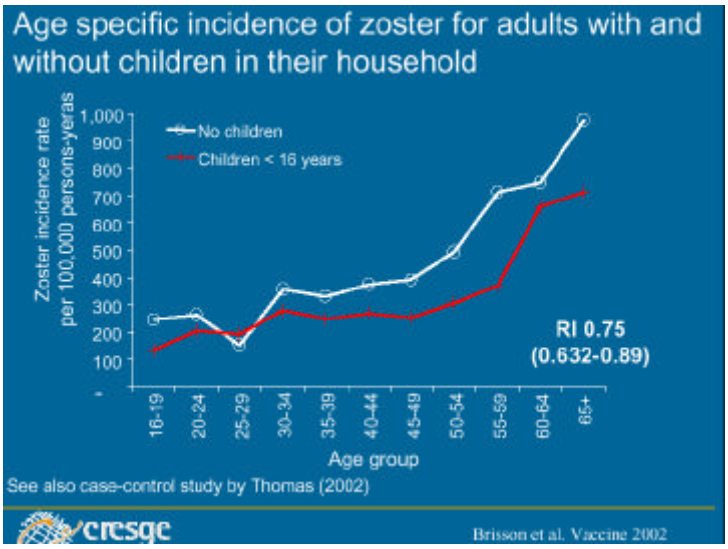
Some background on the Varicella-Zoster virus (VZV)

- Varicella is a mild childhood disease in 90% of cases; complications increase with age.
- Reactivation of dormant VZV results in herpes zoster (shingles). This disease is much more complicated than varicella.
- The immune response to VZV can be boosted via two mechanisms: endogenous boosting (sub-clinical reactivation of the virus) and exogenous boosting (exposure to infectious individuals).

What vaccination strategies have been analyzed in the literature?

1. Mass childhood vaccinations with or without catch-up strategies. A concern with this strategy is that it may lead to an increase in adult cases—the severity of the disease increases with age. If the loss of exogenous boosting leads to an increased incidence of zoster disease, vaccine-induced immunity may lead to a pool of susceptible older individuals.

The following chart shows the incidence of zoster in families with and without children:



This shows that people without children have a higher incidence of zoster than people living with children. So this is an argument for the exogenous boosting hypothesis.

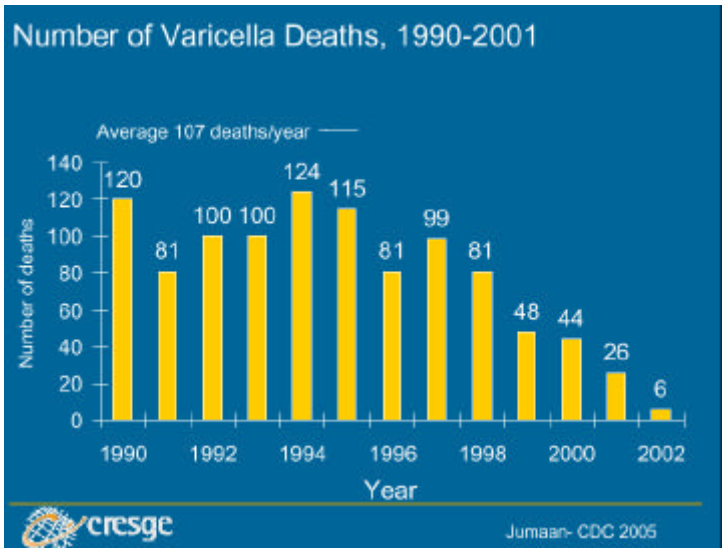
2. Targeted vaccinations for susceptible adolescents and adults, health care workers, immune compromised individuals, thus avoiding children as carriers of the virus.

3. More recently a combined vaccinations strategy has been used involving varicella vaccination in childhood and zoster vaccination at an older age.

Evaluating the patterns of varicella vaccination programs in the USA

The USA started routine childhood vaccination programs in 1995. This was strengthened in 1999, with the vaccine being required for school or daycare entry. In June 2006 the vaccination schedule changed to two doses.

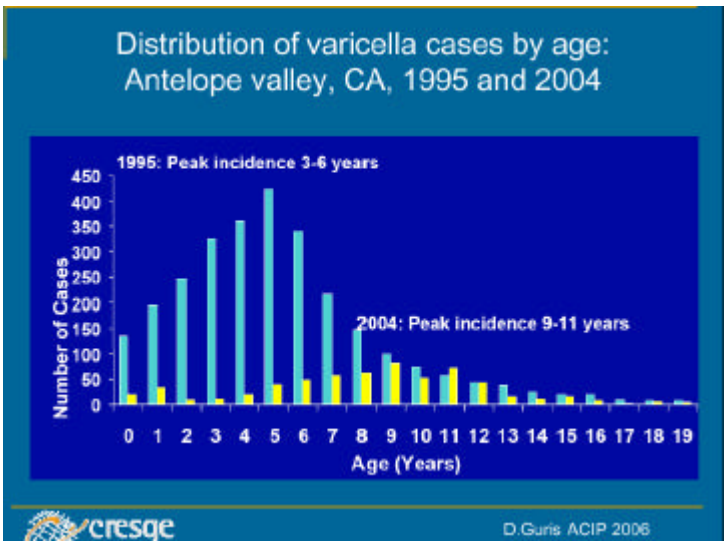
Thus there is very good vaccination coverage in the USA, and the incidence of varicella is decreasing sharply, child hospitalization has decreased and there had also been an important decrease in the number of deaths due to varicella.



But some adverse effects of this vaccination program have also been observed:

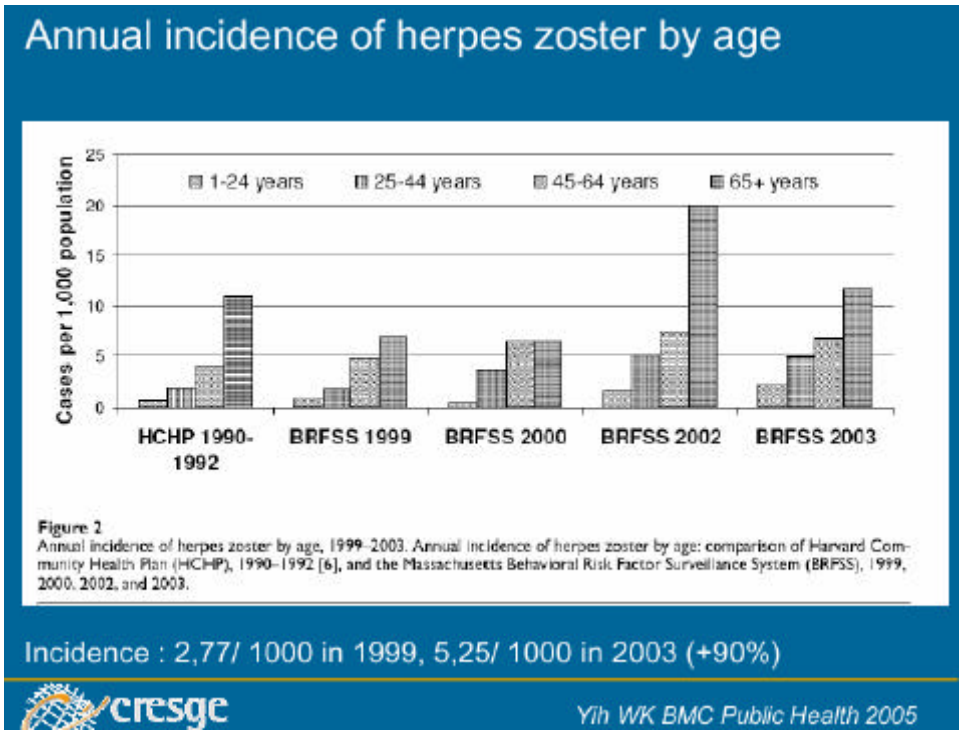
1. An increase in breakthrough varicella for children between 12 months and 12 years, and this breakthrough occurred a long time after the vaccination.

2. An increase in the average age of children becoming infected, as expected. For example, the peak incidence in 1995 was in the 3 to 6 age group; in 2004 it was the 9 to 11 age group. This implies an increase also in the severity of the disease; see graph below.



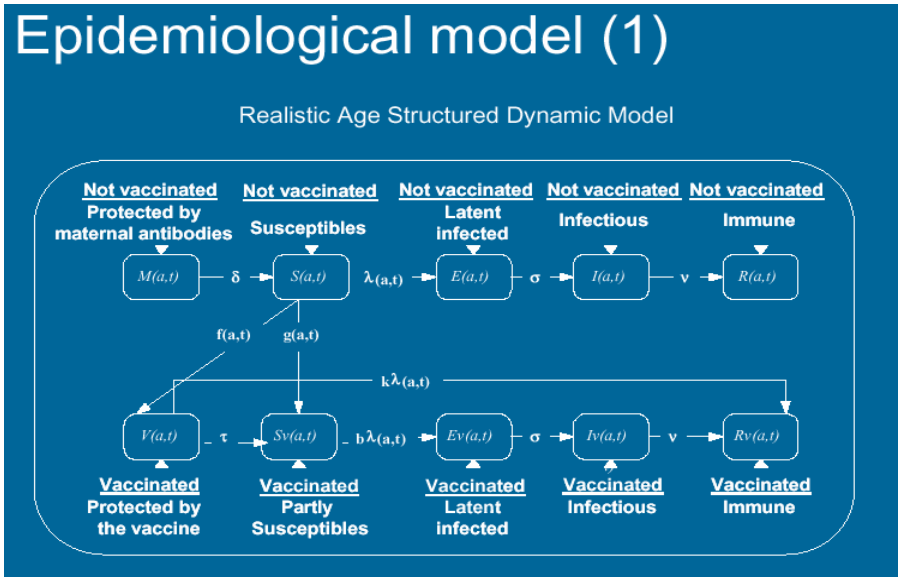
3. A clear increase in the incidence of Zoster of about 90% between 1999 and 2003 (following data from Massachusetts before vaccination in 1992 compared to data up until 2003). The graph below shows this. This is another argument for the exogenous hypothesis put forward above.

.

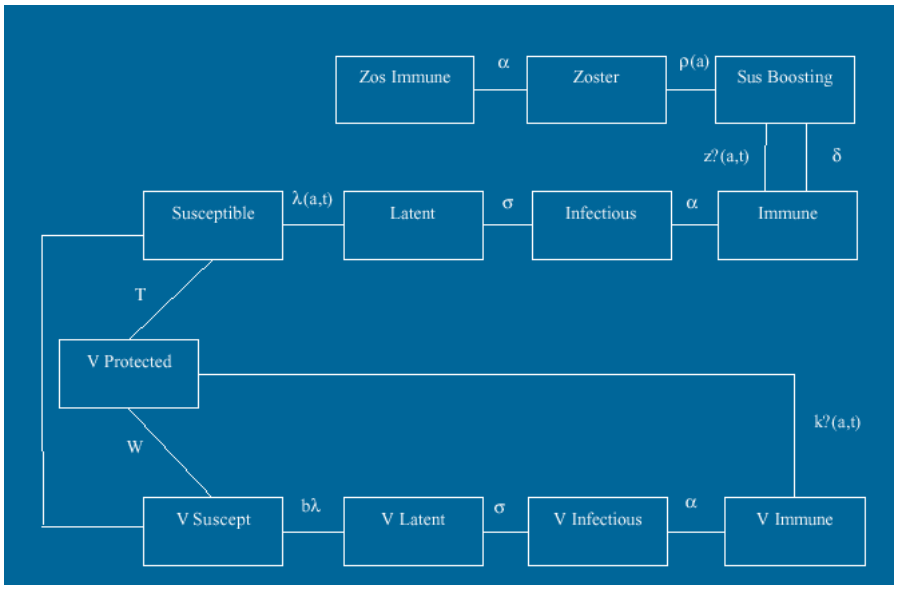


How can we take into consideration all these factors in modelling the benefits of the Varicella vaccine? There are both positive factors (reduction of incidence, hospitalizations etc) and negative factors (including indirect effects such as the increase of Zoster incidence due to Varicella vaccination).

In the initial model there is no interaction between Varicella and Zoster:



This model was revised to allow varicella and zoster to interact through exogenous boosting, as follows:



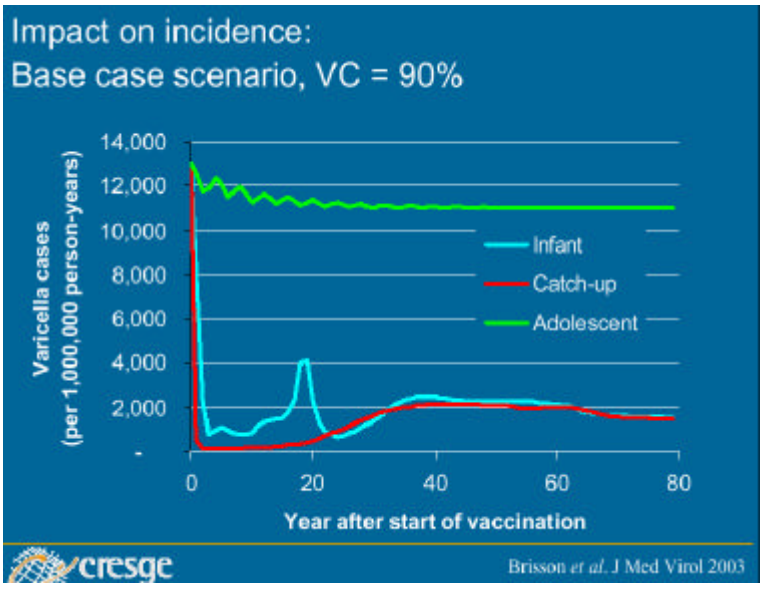
There are many models run in different countries based on the published data of epidemic evaluations. The different models have been ranked according to whether they take zoster, waning immunity and herd immunity into account, along with other factors. Most of the models are dynamic models; this makes a huge difference from static models that do not take into account these other aspects.

Since 2000	Country	Herpes zoster	Waining	Herd immunity	model
Scuffham et al. (2000)	Australia	No	No	No	Static
Brisson et al (2002)	**Canada**	**Yes**	**Yes**	**Yes**	**Dynamic**
Getsios et al (2002)	Canada	No	Yes	No	Static
Banz et al. (2003)	Germany	No	Yes	Yes	Dynamic
Hsu et al. (2003)	Taiwan	No	No	No	Static
Coudeville et al. (2004)	Italy	No	Yes	Yes	Dynamic
Coudeville et al. (2005)	France & Germany	No	Yes	Yes	Dynamic
Lenne et al. (2006)	Spain	No	Yes	Yes	Dynamic
Hammerschmidt et al (2007)	Germany	No	Yes	Yes	Dynamic

What does the model predict for Varicella?

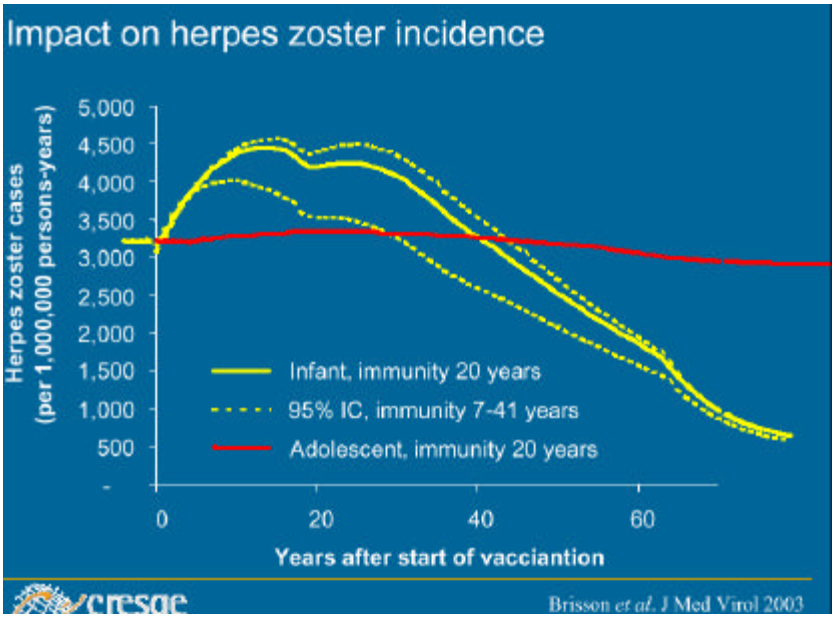
- Incidence of Varicella will rapidly decline after the implementation of vaccination.
- After this "honeymoon period", a rise in incidence would occur, there could be post-honeymoon epidemics.
- Post-vaccination equilibrium is always lower than the pre-vaccination level.
- The age at infection is predicted to increase but the large reduction in incidence in children is likely to outweigh any increase in incidence of chickenpox in adults.

The following graph shows the different strategies—infant vaccination, infant vaccination plus catch up and adolescent vaccination—in the baseline scenario. One can observe the post-honeymoon epidemic, although this does not occur if catch up is implemented. Evidently in adolescents there is not a huge incidence of varicella because most people have had varicella before adolescence.



Similar results were obtained for hospitalizations.

The interesting feature of the new model is that it takes into consideration, among other things, varicella and zoster incidence. The graph below shows that due to a reduction in varicella exogenous boosting there is an increase in zoster incidence in the short run. After that the incidence declines as the vaccinated cohorts begin to reach the age at which most zoster occurs. However, the incidence is expected to remain above the pre-vaccination level until 30 to 44 years after the introduction of vaccination. On the other hand vaccination of 11 year olds will have little to no effect on zoster.



What is the impact of varicella vaccination on zoster incidence?

Disadvantages

- Varicella vaccination is less cost-effective from the health care payers' perspective than from the societal perspective (time/leisure costs).
- Routine vaccination of preteens is the preferred strategy from the health care payer's perspective.

Advantages

- Increase in zoster after routine vaccination of infants would render immunization highly cost-***in***effective.

Persp	Infant	Catch-up	Preteen	Infant (BV)	Infant (zoster)
HC	0.61	0.60	0.73	0.59	0.16
Society	5.24	4.90	4.44	5.09	
ICER ($ / LYG)	44503	50866	**18508**	46896	118188

Brisson et al, J Med Virol 2003

When looking at the results in terms of the cost-effective analysis benefit ratio, there is no strategy in the model that is cost saving because if you look for any one unit invested we only get 60% return. But when looking at the safety side benefice, there is a big return —about 5 times what you have put towards vaccination.

When looking at the incremental cost-effective ratio, the only strategy that seems to be cost-effective is the preteen vaccination.

The above results are sensitive to different factors as follows:

- Epidemiological factors such as vaccine efficacy, vaccination coverage.
- Economic factors such as vaccine price, value of time/leisure lost (indirect costs), discount rate for health consequences and time horizon.
- Predictive value of anamnestic screening for targeted vaccinations.

Indirect costs

These are a very big issue in the ethical and economical evaluation of vaccines. This raises many questions, such as: Should they be included in the cost-effectiveness analysis? If the answer is yes, then how to measure productivity loss?

There are some proposed methods:

- The human-capital method: refers to welfare economics (wage rate = opportunity cost of time). It is easy to implement and gives potential productivity costs.

- The friction-cost method: a more realistic estimate of productivity costs, but requires more data to be collected.

The two methods lead to similar results when the period of absenteeism from work is short (as for chickenpox).

There is another aspect of indirect costs with regards to varicella, and they are likely to be the driver for varicella vaccine demand. Thus parents can ask for vaccine so as not to lose work time; thus it is necessary to take into consideration the influence of this kind of phenomenon on vaccination demand.

Other questions with regards to the economic evaluation of vaccination concern the transferability versus the generalisability of studies of the disease. "Transferability" means you can take the studies done in one setting and apply them to another setting. "Generalisability" implies you take the model and apply it anywhere. For example, models are not always transferable between different countries, such as Germany and France, because of the different settings, health care structures and cost structures.

Conclusions

- The cost-effectiveness of a varicella vaccination program depends on both the direct and indirect effects on varicella and zoster epidemiology. It is important to include in the model both economical and epidemiological aspects.
- Mixing patterns as well as time windows can have major implications.
- Extensive sensitivity analysis on key parameters is needed.
- A step further? Epidemiological models with endogenous behaviors (eg. vaccination coverage and breakthrough varicella, impact of MMRV on vaccination coverage).

Discussion

**Can you explain how you related Zoster to Varicella in the model?*
In the model we assume that people who are not in contact with the varicella virus are not robust in terms of immunity, so can get zoster later in life more easily.

Evaluating New Pertussis Vaccination Strategies for the US

Annelies Van Rie, University of North Carolina, Chapel Hill, USA

Pertussis is an endemic vaccine-preventable disease. It causes an estimated 50 million cases and 300,000 deaths worldwide each year. Childhood vaccination leads to a fall in pertussis incidence rates, but does not result in adequate control of pertussis, despite the high coverage (95%).

Since 1976 the incidence of this disease has steadily increased in all age groups, even though it is thought to be only a childhood disease. Infant pertussis accounts for > 60% of pertussis related complications, 86% of hospitalizations and 90% of deaths.

There has been an underestimated incidence of severe pertussis. This is because we cannot diagnosis the disease that well; childhood symptoms are more identifiable but in other age groups symptoms are more atypical.

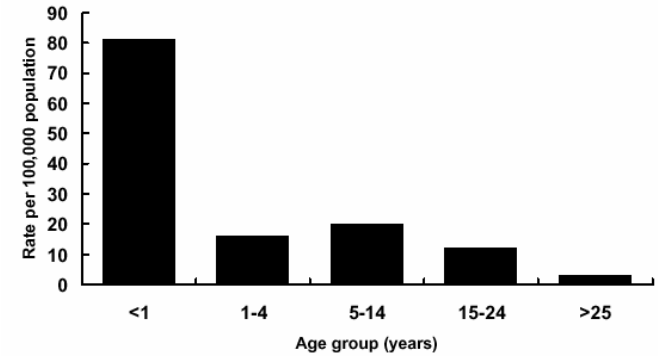
The increase in pertussis in recent years is minor compared to the 1940s, but there is still a gradual increase in the disease presentation; this was especially marked in 2004-2005.
Pertussis has a cyclic epidemic presentation every 3 to 4 years; however the increase in recent years has sparked a debate over whether this is due to virus sensibility, or due to greater awareness of the disease. It is likely to be a combination of both factors.

When looking at the graph of pertussis by age, we can see that more than half of the cases are adolescents and adults. Thus, is this a true increase or is it that there is now more awareness of the disease in other age groups?

This raises another question: Is pertussis a problem for adults?
Pertussis causes a prolonged cough, for 3 months or longer, sometimes with post-tussive vomiting and other complications. Regarding medical costs, it requires multiple medical visits and extensive medical evaluations, and work absenteeism.

The disease can be transmitted by adults to young infants (parents accounted for 55% of source cases, followed by siblings: 16%), thus adults are reservoirs of pertussis in the community.

When looking at the age specific pertussis incidence, the incidence (even in the error of vaccination) is the highest in young infants and lowest in adults over 25. When looking at the absolute number of cases, there are as many people over 25 as there are infants of less than one.



N=25,172. *MMWR* 2004; 53(53):30.

When considering complications, the pneumonia and hospitalization rates are much higher in young infants (mainly those under six months, and especially those under 3 months), and much lower for children over 1 year of age.

This underlines the critical need to protect infants. There are two ways to approach that:

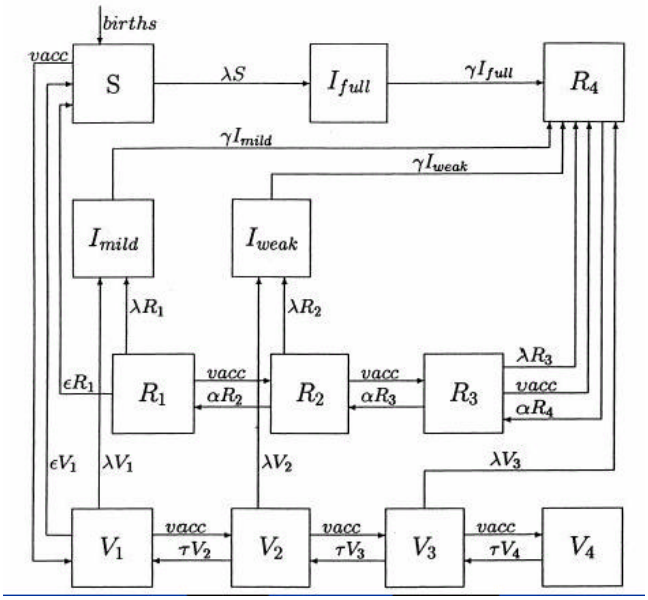
1. Develop vaccines that are immunogenic at birth (the current vaccine cannot be given at birth so children up to 3 months are exposed, despite this being the age with the highest incidence rate). Progress in developing such vaccines is slow, although there are recent encouraging data for premature infants.
2. Boost the immunity of adolescents and adults using new vaccination strategies (boosters are available for these age ranges).

Adult and adolescent vaccination

The primary objective is the direct protection of the vaccinated adolescent or adult. The secondary objective is to reduce the reservoir of B. pertussis and thus the incidence of pertussis in infants.

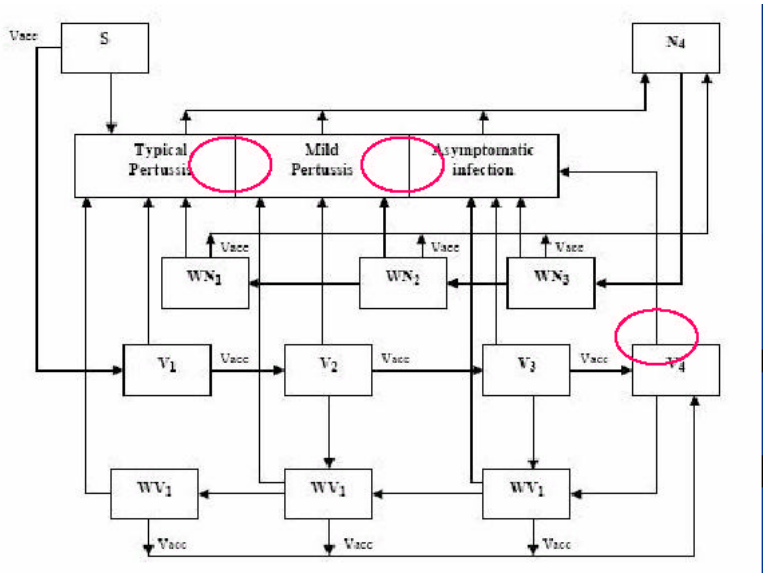
Can the introduction of new vaccination strategies achieve this and will this be cost-effective?

The starting point for establishing this is the Age Structured Model for Pertussis Transmission developed by Hethcote in 1997. It is an eighth structured model that was not specifically created to look at adult vaccination. This model has been used or built upon in the last decade.



In 1999 the model was updated by Hethcote to take into consideration the rising incidence of the disease, which the previous model was unable to take into account. However, the changes made were still limited for the requirements of disease evaluation.

In 2004, Hethcote and Van Rie evaluated the model for five adolescents and adult strategies. The structure is pretty much the same except that it allows people in certain categories to have different presentations of the disease such as typical pertussis, mild pertussis, or asymptomatic infection. It also includes four doses of vaccines so as to make the model more realistic.

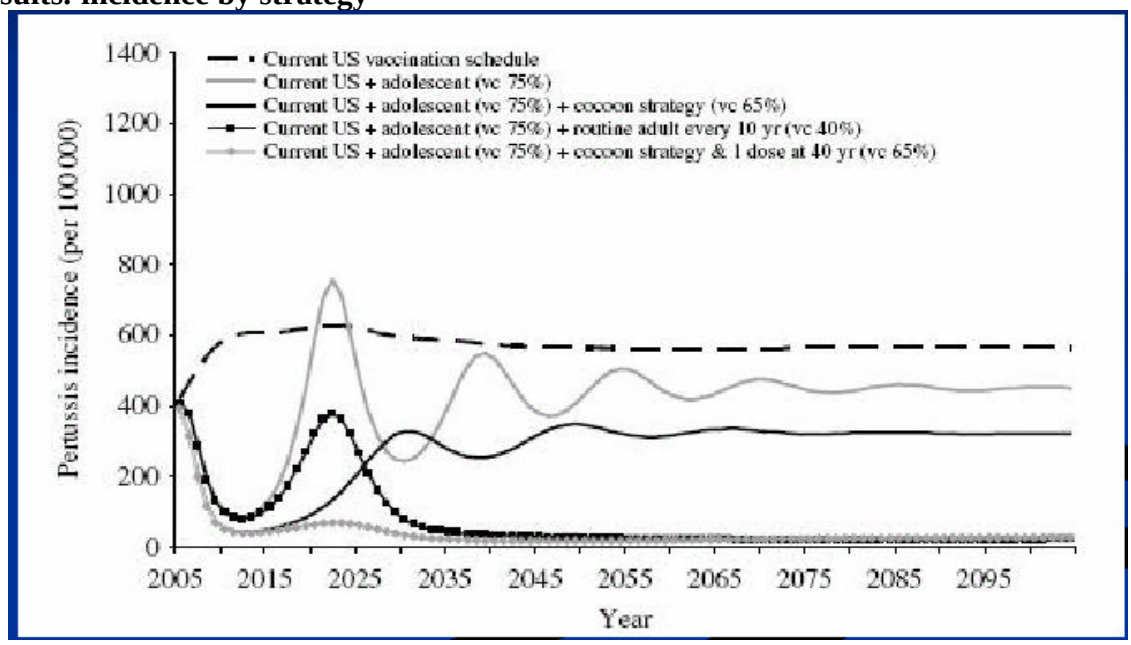


In 2006, the model went through an epidemiological update (Coudeville, Van Rie, Andre). The model structure was kept but many of the parameters were revised based on the data that have become available over past years.

The main changes are:

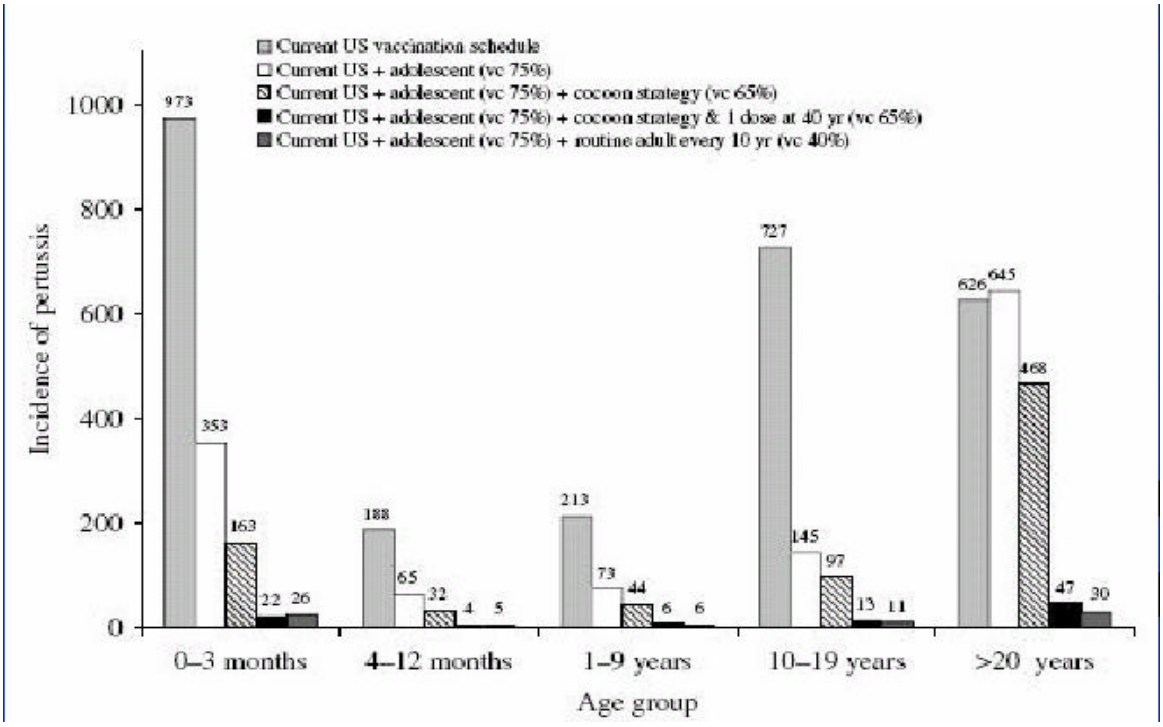
- The probability of developing pertussis after exposure to B. pertussis are calculated using recently published data of the efficacy per dose of vaccine administered.
- The age specific forces of infection are based on recent US incidence data instead of pre-vaccine era seroprevalence data.
- The role of asymptomatic infection, which was revised with regards to the transmission of B. pertussis (simultaneous calibration based on an expectation minimization algorithm).
- The use of recent data on sources of transmission of B. pertussis to young infants.
- Taking into account US population growth over time.

Results: incidence by strategy



Results: incidence by age group (herd immunity)

If you introduce an adolescent vaccine there is a large direct effect and also a decrease in infant incidence etc.



Is adult vaccination economically viable?
In the model, adolescent vaccination seems to give a good result initially, but in the long term it might produce a honeymoon period followed by an increase of incidence. This is not observed with adult vaccination, which seems to mitigate that effect later on. The literature shows that adult pertussis is an economic problem at least in the developed world, with high adolescence pertussis that results in high costs.

Cost-effectiveness analyses
In 2004, evaluations were done of the health and economic benefits of 7 strategies for pertussis booster to adolescents and adults (Purdy 2004). These found that the most economical strategy was to immunize adolescents from 10 to 19 years of age, and that routine adult booster vaccination every decade would be more expensive and more difficult to implement.

In 2005, Caro used an epidemiologic model of routine pertussis immunization in adolescents in the US. The evaluation considered both direct and herd immunity and found that the conditions required for adolescent immunization to be economically warranted are realistic.

In 2007, Lee published an article based on use of the Markov model to calculate health benefits, risks, costs, and cost effectiveness of 1) no adult vaccination, 2) one-off adult vaccination at 20-64 years, and 3) adult vaccination with decennial boosters. Routine vaccination of adults aged 20 to 64 years with combined TD is cost effective if pertussis incidence in this age group is greater than 120 per 100,000 inhabitants.

However, none of these studies of adult vaccination fully included an in-depth assessment of herd immunity, and the benefits and costs over time. The solution is a cost-effectiveness analysis that builds upon a compartmental age-structured transmission model of pertussis and looks at different time points (Coudeville et al).

Taking into consideration costs over time by strategy, the analysis shows that costs are driven by disease incidence and not by vaccination.

In the sensitivity analysis the parameters assessed were: vaccine efficacy, pertussis incidence, vaccine coverage, disease associated costs, discount rate and transmission to infants. Similar to the study done by Lee, the results are sensitive to vaccine efficacy, pertussis incidence at baseline.

Conclusions

- Pertussis remains endemic, even with high coverage rates of a highly effective vaccine in children.
- Analysis of pertussis epidemiology is complicated by lack of a diagnostic gold standard, wide range of disease presentation, and waning of immunity.
- Changes in model structure and parameters have an important impact on results.
- Most recent models seem to indicate that pertussis could be eliminated as a public health problem if both adolescent and adult vaccination are implemented. This strategy may be economically viable.

Discussion

**In your talk were all the vaccination strategies that you mention cost-saving?*
Yes
**In fact the key point in this is the incidence, if you consider reported incidence in the USA. You will say no, nothing is cost-effective, so the conclusions are really driven by the incidence that you are considering (Comment).*

4. Session III: Predicting the Impact of Interventions

Strategies for Containing an Emerging H5N1 (or something else?) Pandemic

Ira M Longini, Vaccine and Infectious Disease Institute, Fred Hutchinson Research Center, Washington, USA

For the first time in human history we have the capacity to stop pandemic flu before it spreads around the globe. H5N1 is being watched very closely and we have plans for containment, etc. The methods developed look closely at the clusters for detecting and estimating infection transmission at source:

- Is transmission person-to-person ?

- If yes, determine the estimates of important transmission parameters for more complex models by using TRANSTAT, a simple stochastic model.
- Determine control measures for transmission at source (social distancing, antivirals, vaccines) using a large-scale stochastic model.

Real time detection, estimation and control
The problem is that the illness is observed at onset times. Thus, we need to decide rapidly whether the disease is infectious or not, if it is spreading person-to-person or via another common source such as chickens, and if so, what are the estimates of transmission parameters, such as the secondary attack rates and the reproductive number. Afterwards this we can assess

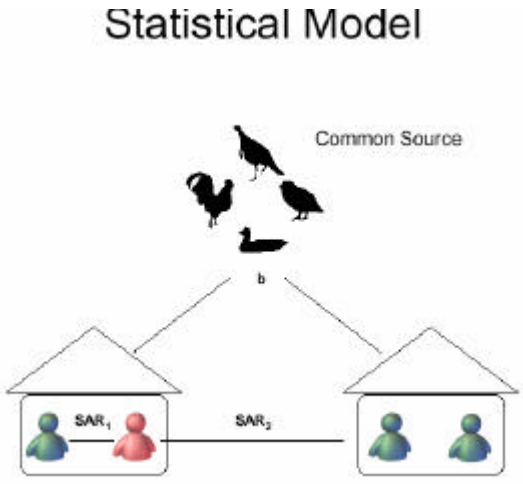
the effectiveness of the interventions, and use the information to calibrate more complex models to help with control and containment strategies.

The information needed

- Data. Part of the battle is the initial data; when an outbreak occurs a team is sent out to gather data. Often the right data is not gathered completely; its focus is mainly on illness onset times. However, other information, such as crude exposure information, not only on the cases, but also on all exposed non-cases, is also very important. Other key information includes treatment, prophylaxis, hospitalization, deaths, infection information, covariates and illness serial interval distribution.
- Natural history. This includes basic information on incubation and infectious periods.

The model described in this lecture considers a number of components such as:

- Close contact within households: the probability that a susceptible person is infected by someone in the same household in one day is p_1 (household = any close contact)
- Casual contact within community: the probability that a susceptible person is infected by someone in the same community, but different household in one day is p_2.
- Common source of infection (eg. zoonotic source or visiting infective from outside of the community): the probability that a susceptible person is infected by the common source in one day is b.



SAR = Secondary Attack Rate
SAR1 = p1
SAR2 = p2

The model is statistically simple; the hypothesis to be tested is as follows:

H_0: $p1 = p2 = 0$ vs.
H_1: $p1 > 0$ or $p2 > 0$

The model does not consider person-to-person transmission p_1 and p_2 are both zero, the only thing driving the cases is the source b.

A likelihood is set for the model based on the symptom onset times of the cases:

- Test statistic

$$\lambda = -2 \log \frac{\sup_b L_0(b|\tilde{t}_i, i = 1, \ldots, N)}{\sup_{b,p_1,p_2} L(b, p_1, p_2|\tilde{t}_i, i = 1, \ldots, N)},$$

where $\tilde{t}_i$ is the symptom onset time for person i, and N is the population size.

Some assumptions are made such as:
-Random mixing in the households and in the community.
-The latent period coincides with the incubation period.
-It is necessary to know the distributions of the latent period (d), and infectious period (?):

$* \ \delta \sim g(l) = \Pr(\delta = l), \ l = \delta_{min}, \delta_{min} + 1, \ldots, \delta_{max}.$

$* \ \eta \sim f(l) = \Pr(\eta \geq l), \ l = \eta_{min}, \eta_{min} + 1, \ldots, \eta_{max}.$

-Observation starts from day 1 to day T, and exposure to the common source starts from day 1 to day S = T.
* When S >T – d min. the asymptotic method does not work.
-The probability that an infective *j* infects a susceptible *i* on day *t* is:

$$p_{ji}(t) = p_1^{I(j \in H_i)} p_2^{I(j \notin H_i)} f(t - \tilde{t}_j),$$

*where H_i is the set of household members of person i.

-The probability that subject I escapes infection from all infective sources on day t is:

$$e_i(t) = (1 - b)^{I(t \leq S)} \prod_{j=1}^{N} (1 - p_{ji}(t)).$$

-A likelihood for *b*, p_1 and p_2 contributed by person is:
:

$$L_i(b, p_1, p_2|\tilde{t}_j, j = 1, \ldots, N)$$

$$= \begin{cases} \prod_{t=1}^{T} e_i(t), & \text{not infected,} \\ \sum_t g(\tilde{t}_i - t)(1 - e_i(t)) \prod_{\tau=1}^{t} e_i(\tau), & \text{otherwise,} \end{cases}$$

-When $p_1 = p_2 = 0$, (2) reduces to

$$e_i(t) = (1 - b)^{I(t \leq S)}.$$

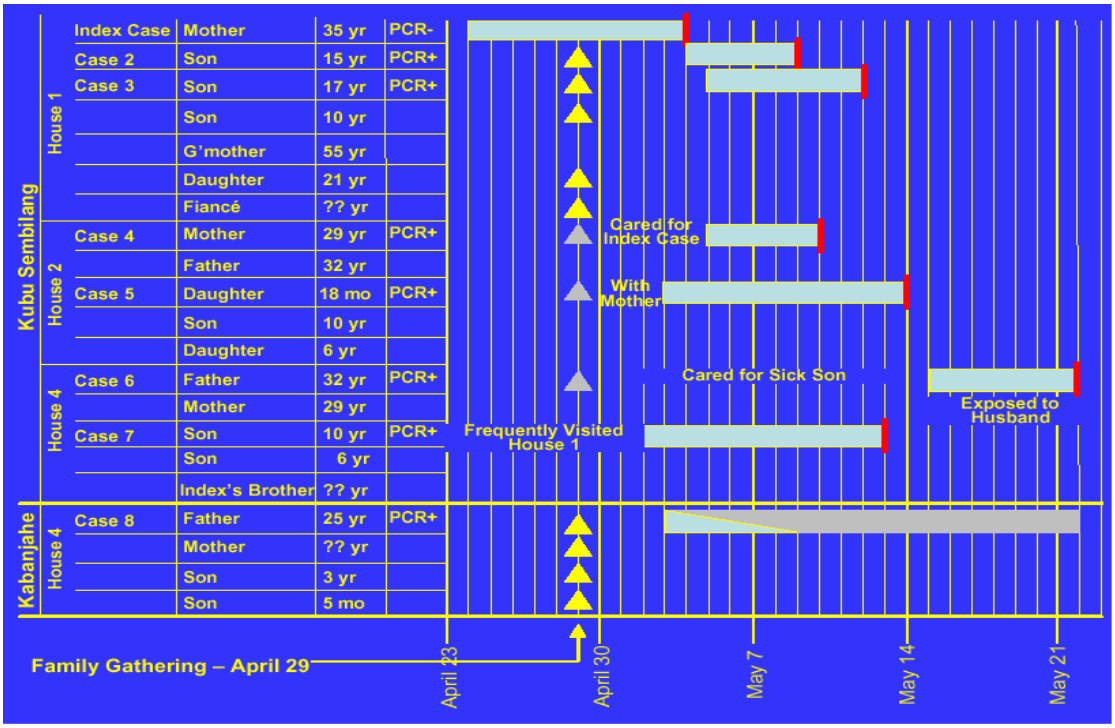
The important parameters estimated in the model are:

- Household $\mathrm{SAR}_1 = \sum_l f(l)\left(1-(1-p_1)^l\right)$
- Community $\mathrm{SAR}_2 = \sum_l f(l)\left(1-(1-p_2)^l\right)$
- Local $R_0 = (M-1)\times \mathrm{SAR}_1 + (N-M)\times \mathrm{SAR}_2$

So SAR_1 stands for the secondary attack rate probability that a person gets infected within a household or close contact during the entire infectious period. SAR_2 stands for the same probability but within the community due to casual contact during the entire infectious period. To estimate the local reproductive number, use M as the average size of a household and N as the average number of all the contacts the person has in the community. Once you have estimated SAR_1 and SAR_2, you can get a rough estimate of the local R_0.

H5N1 influenza in family cluster in North Sumatra, May 2006
Having looked at the essence of the model, we can now follow an outbreak sample.

This was the biggest cluster seen of H5N1 cases, together with one of equal size presented in Turkey. The main question asked was whether the virus was being transmitted person to person?
When following the data on the index cases presented as follows…



***The yellow triangle means close contact**

The above data implies that potentially, person-to-person transmission has occurred. However, nothing is really known unless analysis and statistics are performed. All the information is put into a package called TRANSTAT for analysis.

The tests and estimates are as follows:

- there is statistical evidence of person-to-person transmission (p=0.009);
- the household SAR is 29%, meaning that in the household if you have a close contact you have a nearly 30% chance of getting infected, which is pretty much the % of seasonal flu;
- the lower bound on the local R_0 is 1.14 (i.e. a little above 1, which is a 12% chance of the virus not being spread further in this case).

There is strong evidence that this virus is evolving to the point that it is widely capable of person-to-person transmission. So why did it not spread more in this case? This is because the prophylaxis and therapeutic strategies were implemented very rapidly in all exposed individuals, with strict household quarantine, etc.

The outbreak that happened in Turkey was also analysed and had similar patterns to the one in Sumatra. But for this outbreak we could not reject the null hypothesis, which was that the virus was not spreading person-to-person.

Control of transmission at source

The containment of pandemic influenza at source is the best strategy. This strategy is supported by mobile stockpiles of antiviral agents (Roche donated 5 million courses to WHO), and by mobile stockpiles of vaccines (GSK donated 50 million doses to WHO).
The question is how to use these resources for containment?

Pre-pandemic vaccines and antiviral agents will slow the spread of a pandemic. In order to determine their efficacy there are ways to measure it. The following are measures of vaccine efficacy:

- $\{VE(t)_S , VE(t)_P , VE(t)_I\}$
- $VE_{SP}(t) = 1 - (1 - VE_S(t))\,(1 - VE_P(t))$

$VE(t)_s$, $VE(t)_p$, and $VE(t)_i$ are all functions of time; $VE(t)_s$ is how well the vaccine protects against infection; $VE(t)_p$ is how well it protects against illness caused by the infection; and $VE(t)_i$ stands for how much it reduces transmission.

For example, if $VE(t)_s$ is 0.1 that means that the per contact probability of transmission of a vaccinated person getting infected is reduced by a 10% factor compared to a person who has not been vaccinated., etc.

In many vaccine trials the measuring point is the VE_{sp} which is the reduction probability that someone will get clinical illness and infection if they get exposed, compared to an unvaccinated person. The VE_{sp} is usually measured in phase III trials, which correlate immunity with transmission.

What can we conclude at this point about VE parameters for H5N1 pre-pandemic influenza vaccines for heterologous virus?

VE parameters

- Based on clinical studies in humans

– Overall: VE SP = 60-70%

- Get about halfway there after 1st dose

• Based on ferret challenge studies

– VE I = 0.70

– e.g., viral shedding in the URT: (throat or nasal swab) 26.1% of treatment group, 91.7% of controls

• VE I = 1 - (26.1/91.7) = 0.72

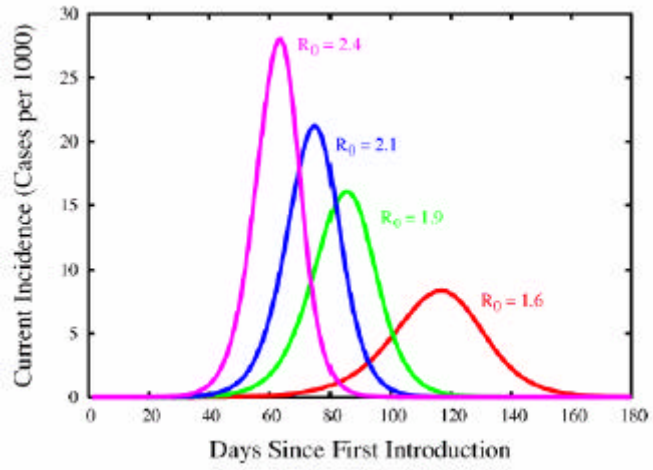
Large-scale, individual-based, stochastic simulation models

Transmission models

The four key elements of our models:

1. Disease natural history model and parameters.
2. Community-level transmission between people, through various contact groups (household, workgroup, school).
3. Census demographics (where people live) and worker flow data (where they work), at tract-level resolution.
4. Transportation statistics on long-distance travel.

The higher the R_0, the earlier and the higher the peak incidence of the pandemic



Germann et al. PNAS 2006; 103 (15): 5935-5940

Pre-pandemic vaccination strategies

• Mass pre-vaccination

– Two doses at least five weeks before initial case.

• Reactive mass vaccination

– Begin vaccinating X days after first case in a geographic region.

• Ring vaccination

– Begin vaccinating X days after first case in ring, then in a ring after each subsequent case.

Antiviral efficacies used in the model: Oseltamivir

• Antiviral efficacy of reducing susceptibility to infection:

AVE_S = 0.48, [0.17, 0.67] 95% CI *

• Antiviral efficacy of reducing illness given infection:

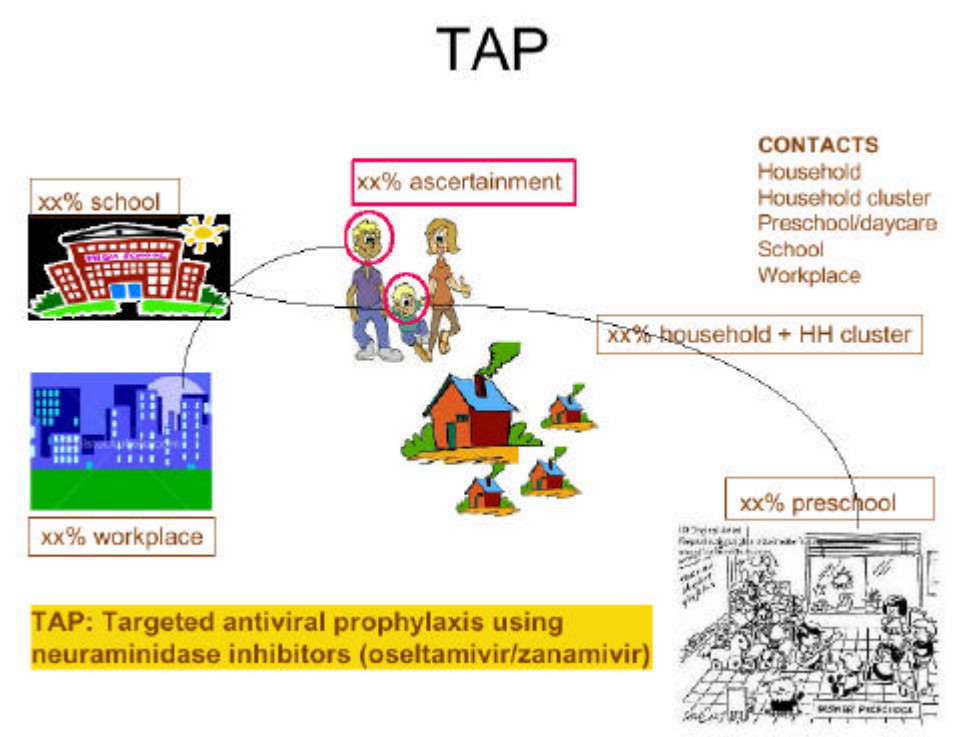
AVE_P = 0.56, [0.10, 0.73] 95% CI *

• Antiviral efficacy of reducing illness with infection:

AVE_{SD} = 0.80, [0.35, 0.94] 95% CI *

– Mult.: $AVE_{SP} = 1 – (1- AVE_S) (1-AVE_P) = 0.77$
• Antiviral efficacy of reducing infectiousness to others:
$AVE_I = 0.80$, [0.45, 0.93] 95% CI *

* data from Welliver, *et al*. *JAMA* (2001); Hayden, *et al*. *JID* (2004); analysis from Yang, Longini, Halloran, *Appl Stat* (2006); Halloran, *et al.Am J Epidemiol* (2007).



Containment at source
• Stochastic, individual based simulations of Southeast Asian population of 500,000 individuals.
• Transmission occurs in households, schools, workplaces, clusters of households, social places, and community.

Examples:

Simulated pandemic influenza outbreak

$R_0 = 1.7$

80% TAP

80% TAP
+ 50% Pre-vacc

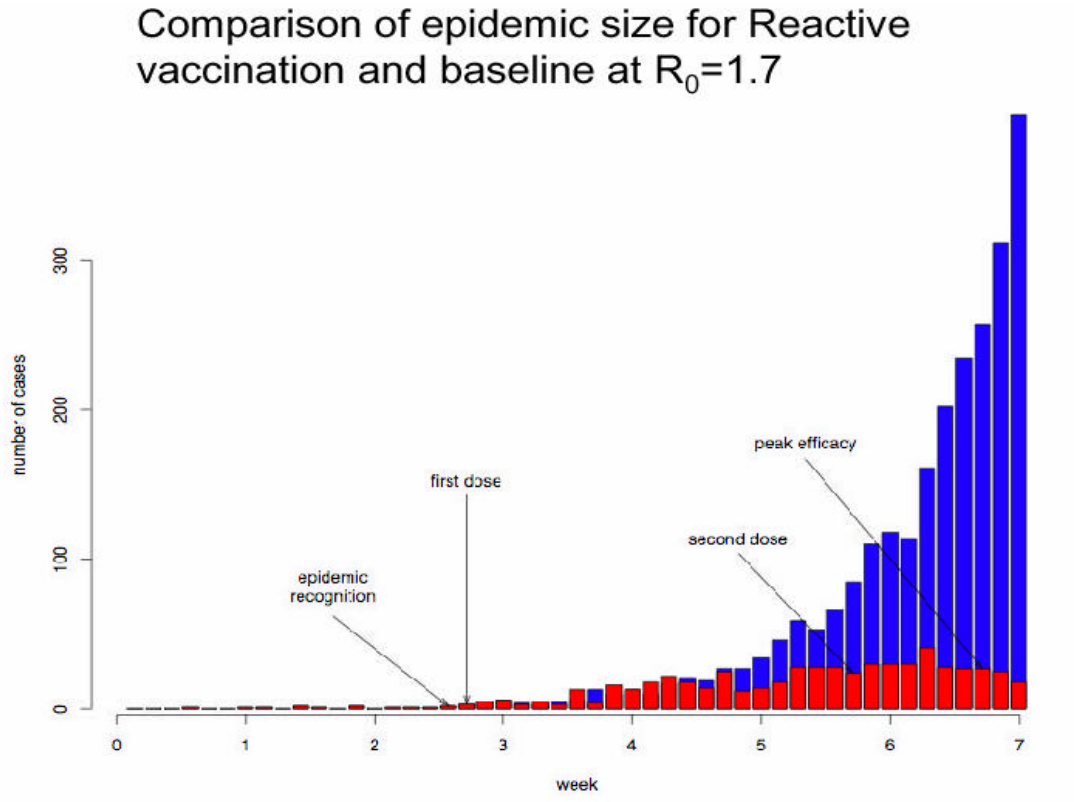
Longini et al. Science 2005; 309: 1083-1087

Simulated pandemic influenza outbreak

$R_0 = 1.7$

70% reactive vaccination

No intervention



Conclusions for Containment

- TAP alone is sufficient for $R_0 = 1.6$
- Yearly pre-vaccination with pre-pandemic vaccine would be best but obviously impossible.
- Reactive mass vaccination is somewhat better than ring vaccination.
- Reactive vaccination would only work for $R_0 = 1.3$
- TAP + reactive vaccination would work for $R_0 = 2.1$
- Reactive and ring vaccination should be started no less than two weeks after the initial case; one week would be best.

Discussion

**Concerning your test of person-to-person transmission, I wonder whether this is really a test situation or whether you could use a marginal decision series, taking into account that the cost of intervening to contain the disease is very high, than to give treatment?*

There are many ways to give the test. We could do what you suggest but I am not sure it will be more efficient.

**Have you tried to measure different levels of person-to-person infection household exposure?*

In the model I just showed you can describe as many layers of contact as you wish and just sign transmission probability to it.

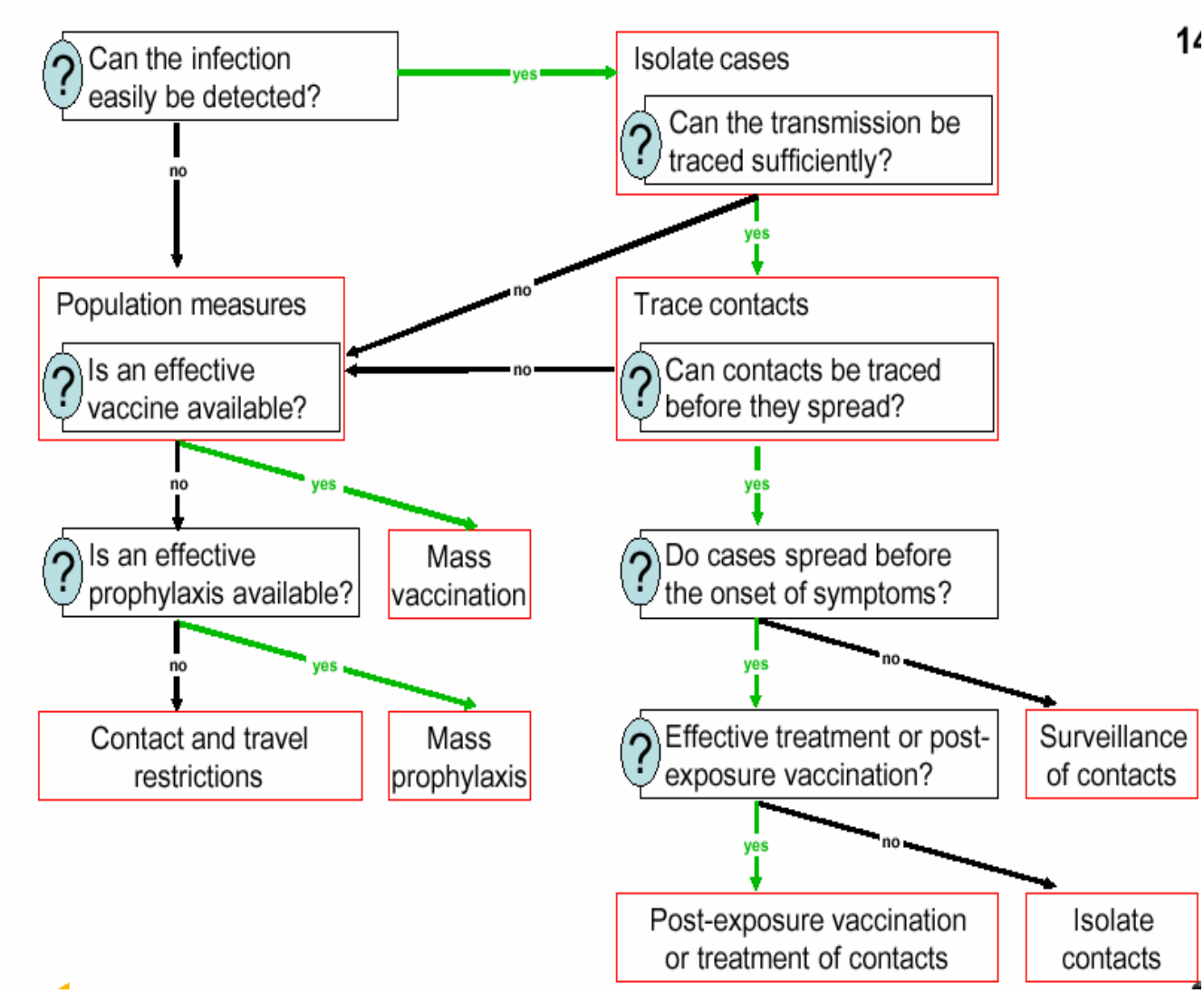
References

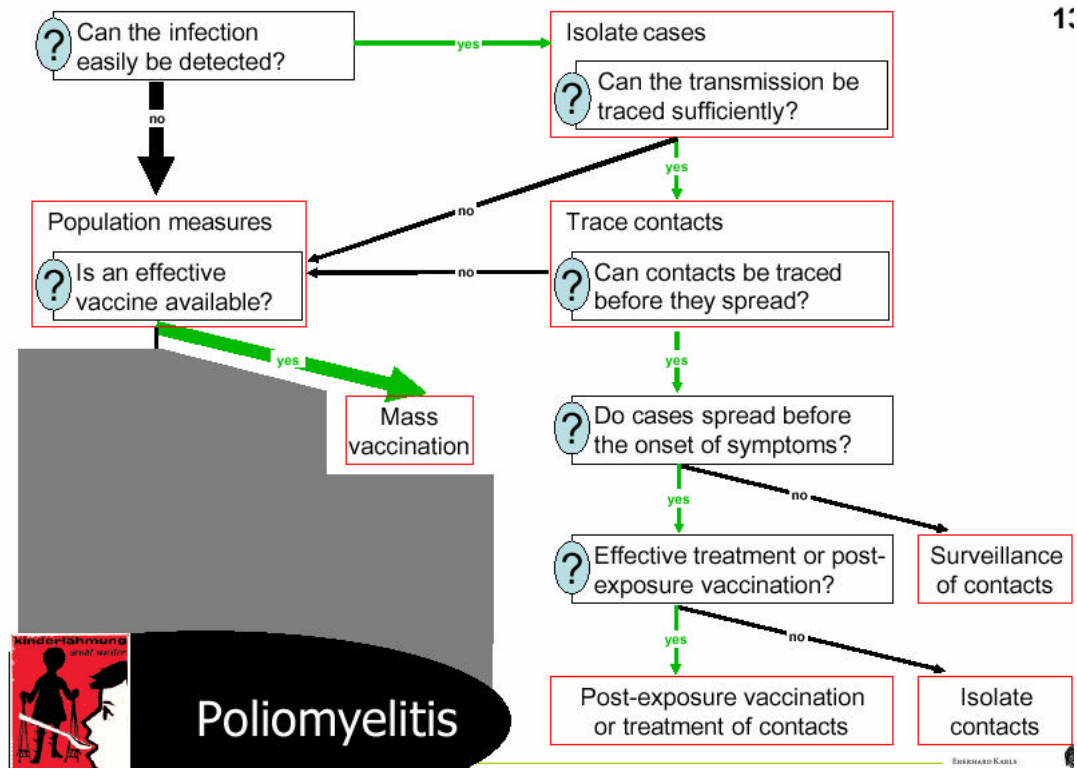
Longini, I.M., et al.: Containing pandemic influenza at the source. Science 309, 1083-1087 (2005)
Longini, I.M. and Halloran, M.E.: Science 310, 1117-1118 (2005)
Ferguson, et al.: Nature 437, 209-14 (2005)
Yang, Y., Longini, I.M. and Halloran, M.E.: Applied Statistics 55, 317-330 (2006)
Yang, Y., Longini, I.M. and Halloran, M.E.: Annals of Applied Statistics 1, 211–28 (2007)
Yang, Y., Halloran, M.E., Sugimoto, J. and Longini, I.M.: Emerging Infectious Diseases 9, 1348-1353 (2007)

Effectiveness of Interventions Against Infectious Diseases – A (Nearly) Non-Mathematical Model

Martin Eichner, Department of Medical Biometry, University of Tübingen, Germany

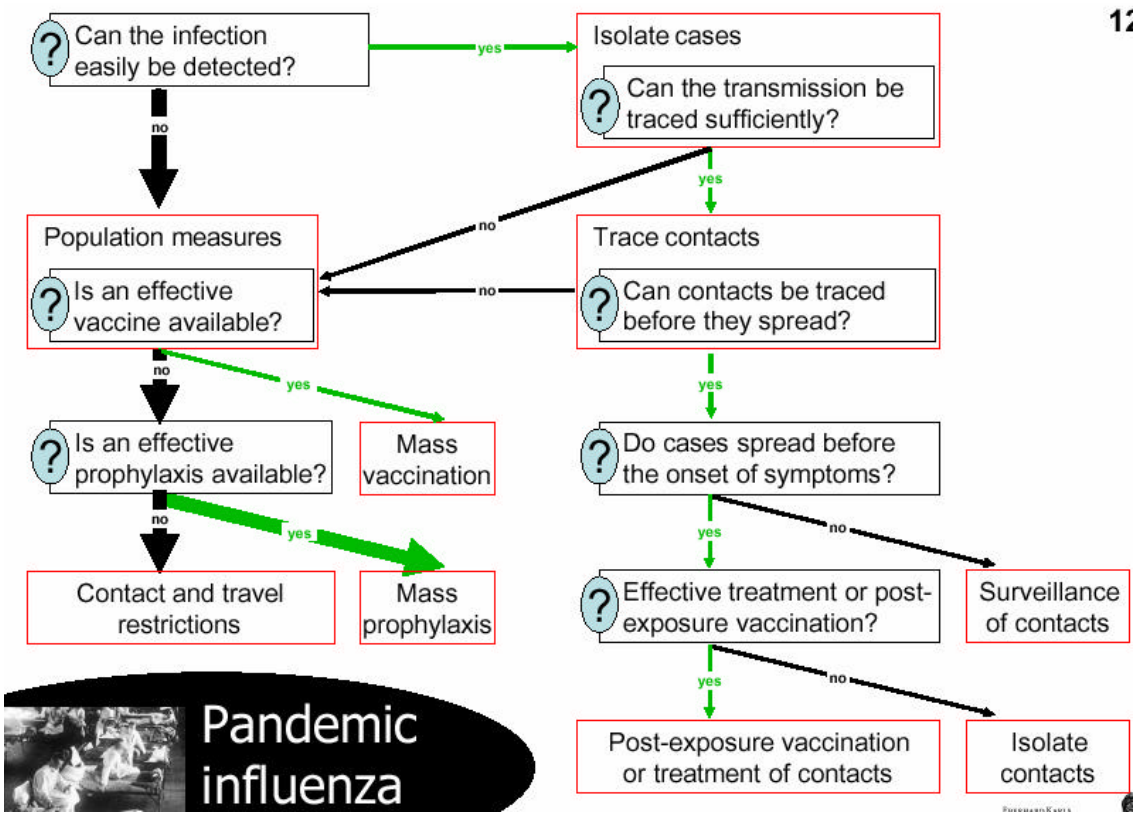
In an outbreak of an infectious disease we ask questions such as: what are we going to do? What are we hoping to achieve? Non-mathematical questions like these guide our decisions, as shown in the following graph:

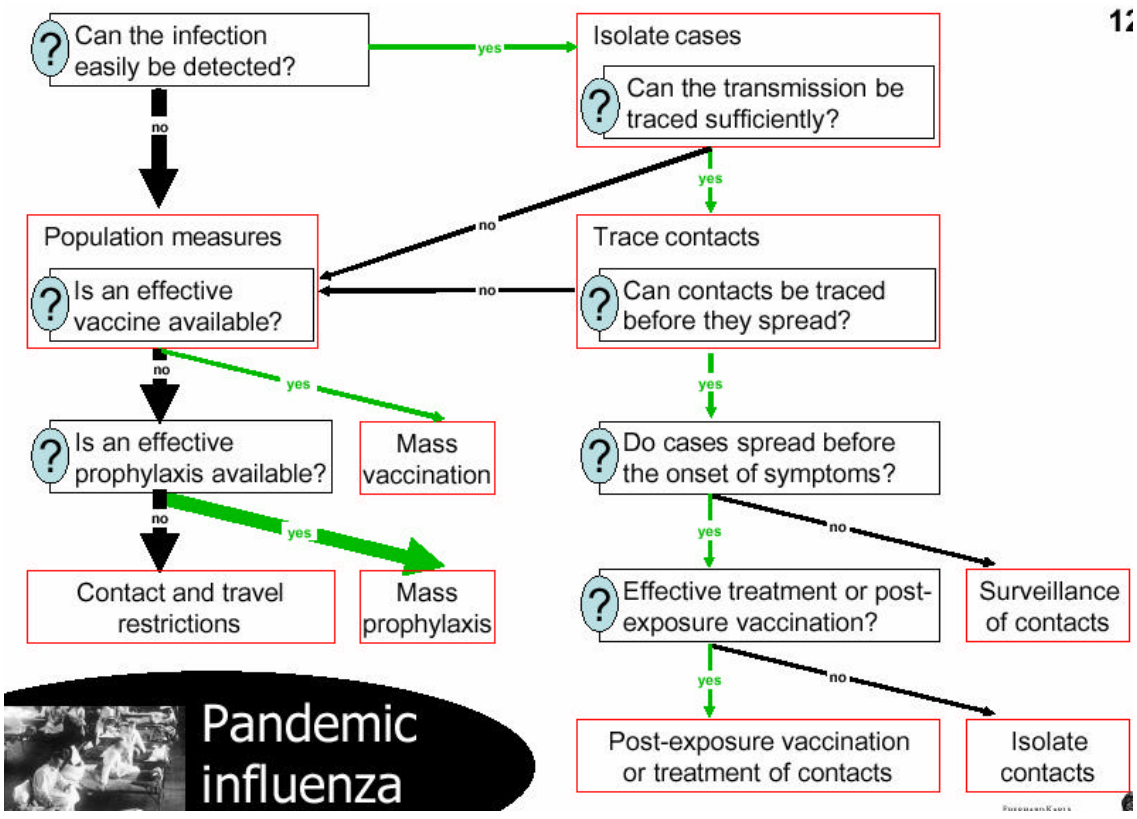




The above decision model can be converted into mathematical modelling to predict a potential reality. For example, when using this model for Poliomyelitis disease we would ask: Can this disease be easily detected? If we answer "no" (because there are hundreds of non-symptomatic cases), we then ask if there is an effective vaccine. If the answer is "yes", then the model points to the mass vaccination strategy. In this way the model helps us move towards some answers.

The model below is for pandemic influenza.





Effective reproduction number

To turn this into a more mathematical model, we can consider the effective reproduction number (ERN). The basic reproduction number (BRN) is the expected number of secondary infections caused by one index case during the whole course of his or her infectious period (in a completely susceptible population, where no interventions have taken place).

The ERN is similar, but without the above two restrictions (in parentheses).

How does the effective reproduction number change when an intervention takes place?

No intervention

If there are no interventions and the BRN: $R_0 = 3 > 1$, the outbreak is out of control.

When interventions take place, the ERN changes as follows:

1. *Vaccination*

Fraction v_p means the fraction of immunized individuals, so for example if

V_p = 70% immune: $R_0 (1-V_p) = 0.9$, gradual fade-out of the outbreak

V_p = 50% immune: $R_0 (1-V_p) = 1.5$, out of control

2. Contract prevention (masks, social distancing)

Fraction r_p of prevented contacts, so for example if

r_p = 70% prevented: $R_0(1-V_p) = 0.9$, gradual fade-out of the outbreak

r_p = 50% prevented: $R_0(1-V_p) = 1.5$, out of control

-Cases are contagious form mi = 14 days

-Cases are isolated after mp days

mp = 4 days: Ro mp/mi = 0.86, gradual fade out of the situation

mp = 7 days: Ro mp/mi = 1.5, out of control

3. Fully effective case isolation

Cases are contagious from m_I = 14 days

Cases are isolated after m_p days

m_p = 4 days: $R_0\ m_p/m_I = 0.86$, gradual fade-out of the outbreak

m_p = 7 days: $R_0\ m_p/m_I = 1.5$, out of control

4. Partly effective case isolation

This situation is more complex, but the calculations are pretty much the same.

Cases are contagious for m_I = 14 days

Cases are isolated after m_p days

Isolation prevents a fraction $r_н$ of contacts.

Assuming m_p = 4 days before isolation, $r_н$ = 90% effectiveness

$R_0\ m_p/m_I =$ 0.86 infections in the population

$R_0 (m_I - m_p)/m_I (1-r_н) =$ 0.21 infections in the hospital

Total 1.07 infections - out of control

Calculating the final size of the outbreak

This can be done as follows.

Take into consideration that the expected number of secondary cases = the effective reproduction number R_e

If $R_e > 1$, the outbreak is out of control
If $R_e < 1$, the outbreak will gradually fade out

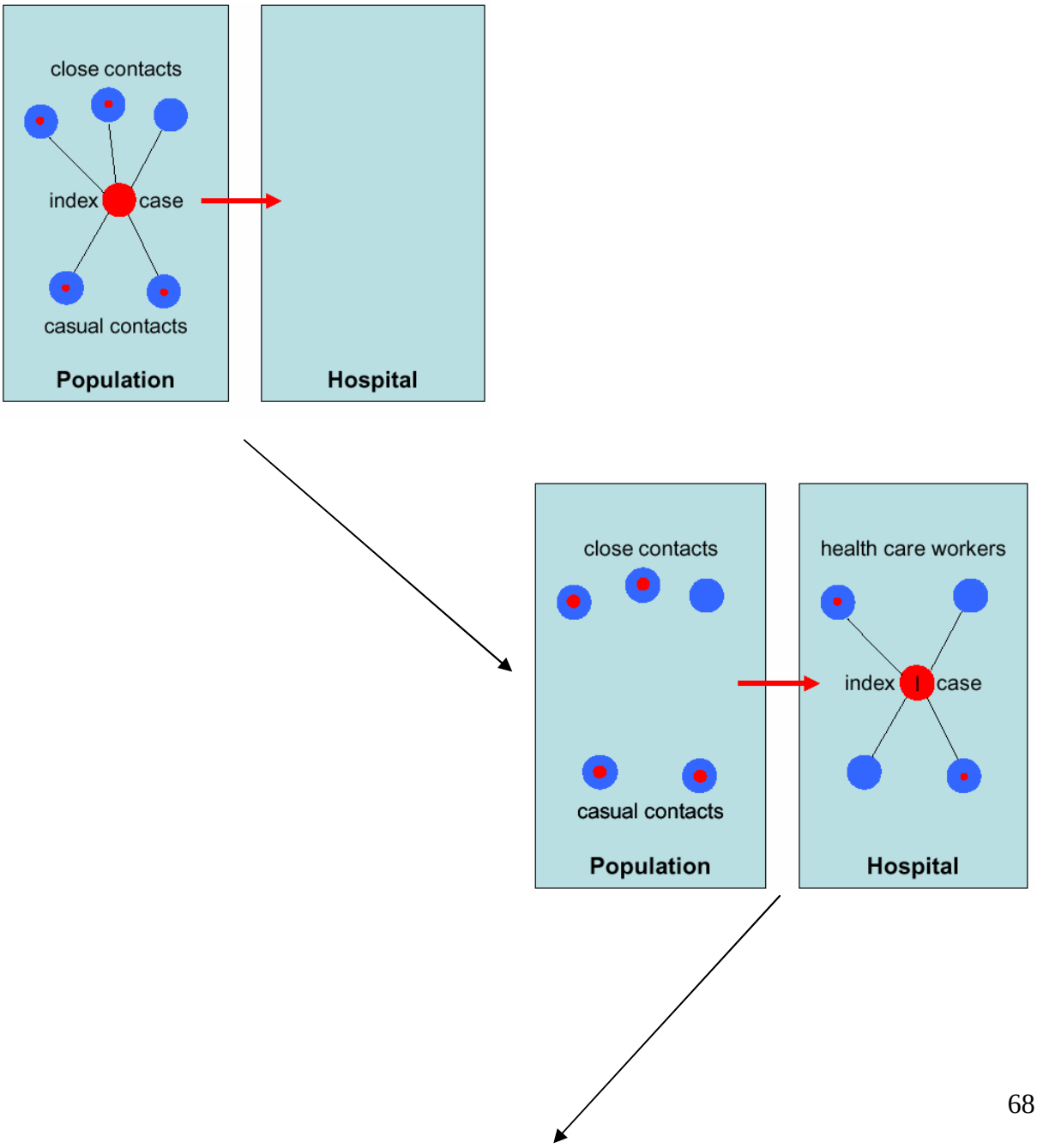
The total outbreak size (for $R_e < 1$) can be calculated as a progression.
So starting with 1 index case + expecting less than 1 secondary cases in the first generation + each of those in the first generation will create less than 1 secondary case in the second generation and so on. This can be written as a mathematical series up to infinity, ending up with the simple result of $1/(1-R_e)$:

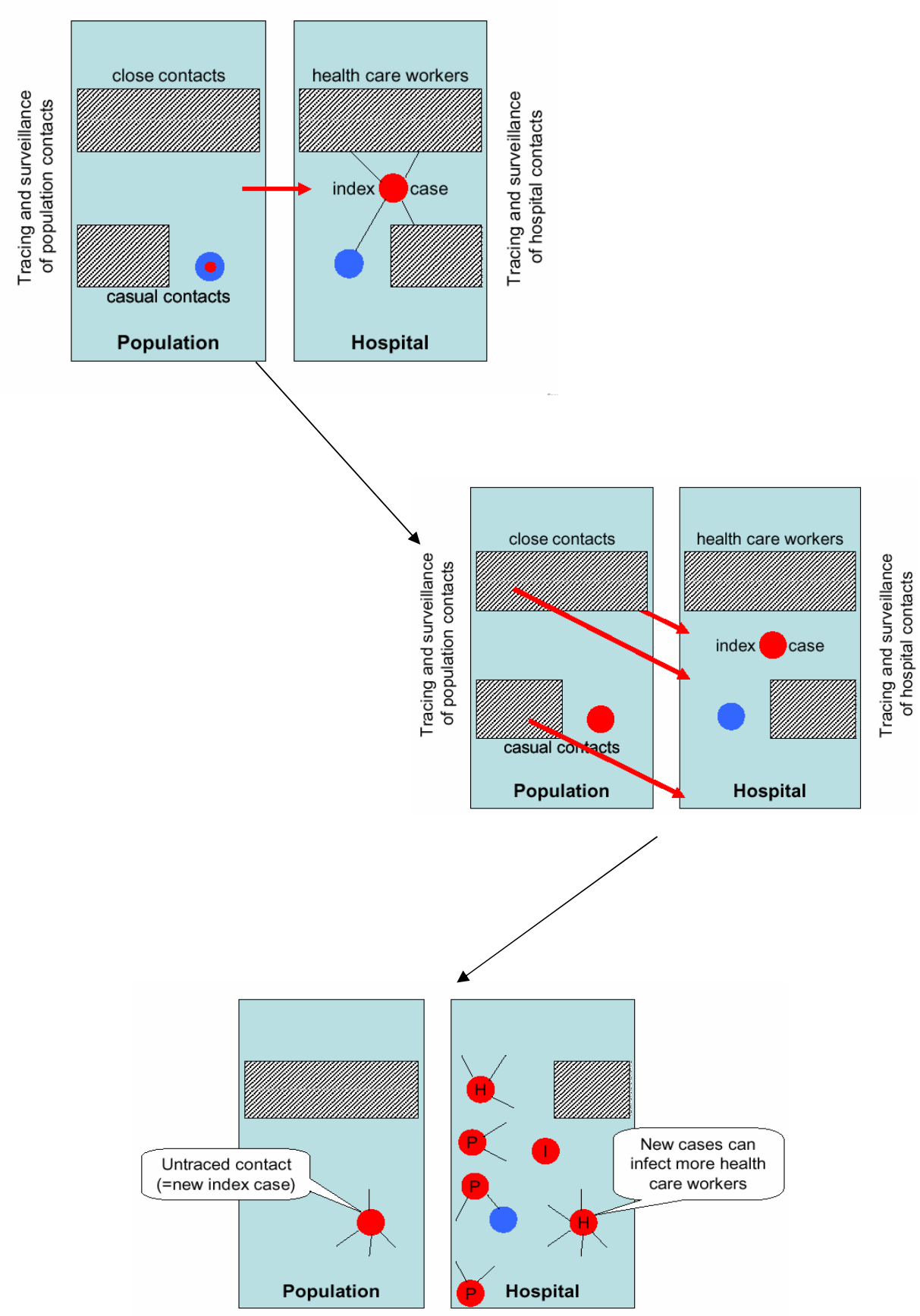
$$1 + Re + Re^2 + Re^3 + Re^4 + \ldots = 1/(1-R_e)$$

Partly effective isolation and contact tracing

This is a little more complex because we are dealing with two different types of people in the population.

We can visualize it as follows:





What it all comes down to is that tracing the cases and applying the appropriate prevention and surveillance strategies are key for outbreak control.

5. Case isolation, contract tracing, surveillance

Case isolation:
-Untraced cases are isolated m_p days after the onset of symptoms
-Traced cases are isolated immediately after the onset of symptoms, before they create any secondary infections.

Contact tracing (and surveillance):
-Close contacts amount to a fraction *c* of all contacts (all close contacts can be traced).
-A fraction t_p of casual contacts in the population is traced.
-A fraction t_h of contacts in the hospital is traced.

Taking all the above into consideration we can deduct the following with regards to secondary infections:

Secondary infections caused by untraced cases	
A fraction *c* of the secondary cases produced by an index case belong to the index case's close contacts who can easily be traced. Of the other secondary cases caused in the population, only a fraction t_p can be traced.	
Before isolation, an untraced case causes traced cases and untraced cases	expected number $R_0(1-v_P)(1-r_P)\,(c+(1-c)t_P)\;\; m_P/m_I$ $R_0(1-v_P)(1-r_P)\,(1-c)(1-t_P)\;m_P/m$
After isolation, an untraced case infects traced HCWs (health care workers) and untraced HCWs	expected number $R_0(1-v_H)(1-r_H)\;\; t_H\;\; (m_I-m_P)/m_I$ $R_0(1-v_H)(1-r_H)(1-t_H)(m_I-m_P)/m_I$
Secondary infections caused by traced cases (these will be put under tight surveillance so that they cannot infect anybody before their isolation)	
Traced cases infect traced HCWs and untraced HCWs	expected number $R_0(1-v_H)(1-r_H)\;\; t_H$ $R_0(1-v_H)(1-r_H)(1-t_H)$

The number of secondary cases can also be calculated through different formulas and also the final size of the outbreak as follows:

Number of secondary infections

01

An untraced index case causes on average C_{uu} untraced and C_{tu} traced secondary cases:

$$C_{uu} = R_0(1-v_P)(1-r_P)(1-c)(1-t_P)m_R/m_I + R_0(1-v_H)(1-r_H)(1-t_H)(m_I-m_R)/m_I$$

$$C_{tu} = R_0(1-v_P)(1-r_P)(c+(1-c)t_P)m_R/m_I + R_0(1-v_H)(1-r_H)t_H(m_I-m_R)/m_I$$

A traced index case causes on average C_{ut} untraced and C_{tt} traced secondary cases:

$$C_{ut} = R_0(1-v_H)(1-r_H)(1-t_H)$$

$$C_{tt} = R_0(1-v_H)(1-r_H)t_H$$

The next generation matrix is then given by $\mathbf{M} = \begin{pmatrix} C_{uu} & C_{tu} \\ C_{ut} & C_{tt} \end{pmatrix}$

R_e is the maximum Eigenvalue of the next generation matrix:

$$R_e = \left(C_{uu} + C_{tt} + \sqrt{C_{uu}^2 + 4C_{tu}C_{ut} - 2C_{uu}C_{tt} + C_{tt}^2}\right)/2$$

To calculate the final size of the outbreak, we start with an initial number of 1 untraced of 0 traced cases in the first generation:

$n_u(i)$ = number of untraced and the number in the i^{th} generation
$n_t(i)$ = number of traced cases in the i^{th} generation

$$\begin{pmatrix} n_u(i) \\ n_t(i) \end{pmatrix} = \mathbf{M}^i \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

In both series converge (otherwise the outbreak will get out of control), we get the final size of the epidemic as:

$$\text{Final size} = \sum_{i=0}^{\infty} n_u(i) + \sum_{i=0}^{\infty} n_t(i) = \frac{C_{tt} - C_{ut} - 1}{C_{uu} + C_{tu}C_{ut} + C_{tt} - C_{uu}C_{tt} - 1}$$

Discussion

**In your flu diagram, prophylactic measures can be instantaneous while vaccines such as for pandemic flu take about 4 to 5 weeks to develop full immunity, so I was just wondering why you put vaccine as the first choice?*

You are right, I should probably have made the flow towards prophylactic measures, as the first choice of intervention.

**There are highly simplified assumptions behind the effective R calculations, could you say something about that?*

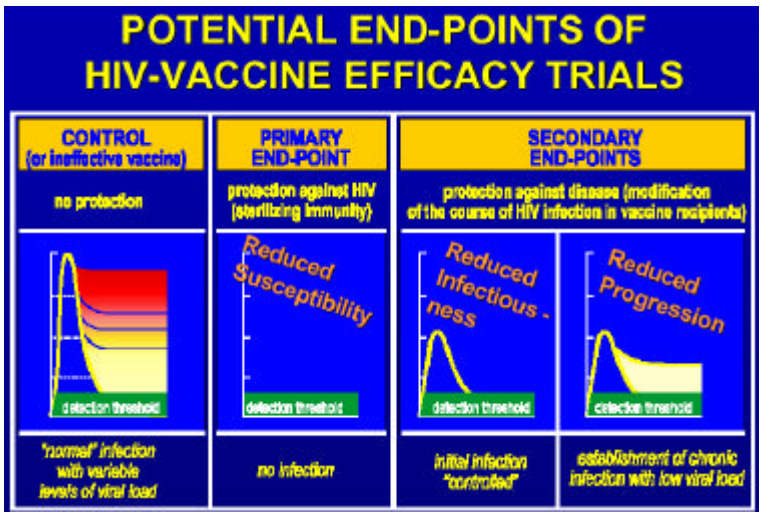
Some of the simplified assumptions in that regard are for example that cases are infectious in the first and last days of the infectious period. Another assumption is that I only worked with expected values. For the different generations I assumed that each person has a specific

number of secondary cases etc, and that an individual is identical in the population, and I didn't think about exhaustion of close contacts in the population.

Advances in Infectious Diseases Modeling: Modeling HIV Vaccines

Daniel C. Barth-Jones, Department of Epidemiology, Mailman School of Public Health, Columbia University, NY, USA.

This talk assumes that there is an HIV vaccine. In HIV there is a primary viremia which then settles down to a set point. Initially the vaccine that was being developed aimed to prevent infection, so as to reduce susceptibility. If we can prevent the primary viremia from occurring, lowering this set point will reduce people's infectiousness and delay the amount of time that it takes for the HIV disease to develop.



Potentially important HIV vaccine effects

A comprehensive evaluation of the HIV vaccine's effects should include measurements of how vaccination affects the following:

- Susceptibility to HIV infections.
- Progression of the disease in those infected despite vaccination.
- Infectiousness of those infected despite vaccination.

It should also account for important aspects of:

- how the vaccine effects are distributed in vaccinated individuals (e.g., "take" and "degree")
- any important modifiers of the effects (e.g., gender, host genetics, STIs, mode of transmission, sex acts, circumcision, etc.), and
- how effects change over time.

Measures of vaccine efficacy

- **VE_S** - vaccine effect on susceptibility
- **VE_I** – vaccine effect on infectiousness
- **VE_P** – vaccine effect on progression
- There is also a need for measurements of vaccine waning and boosting
- **VE_R** – combined vaccine effects on the reproduction number of the HIV epidemic

The combined VE_S and VE_I effects can be used to derive a vaccine effect on the reproduction number R under a set of theoretical conditions:

- random sexual mixing
- 100% of the population vaccinated
- vaccine effect does not:
    - change the duration of infection ($VE_P = 0$), and
    - wane with time

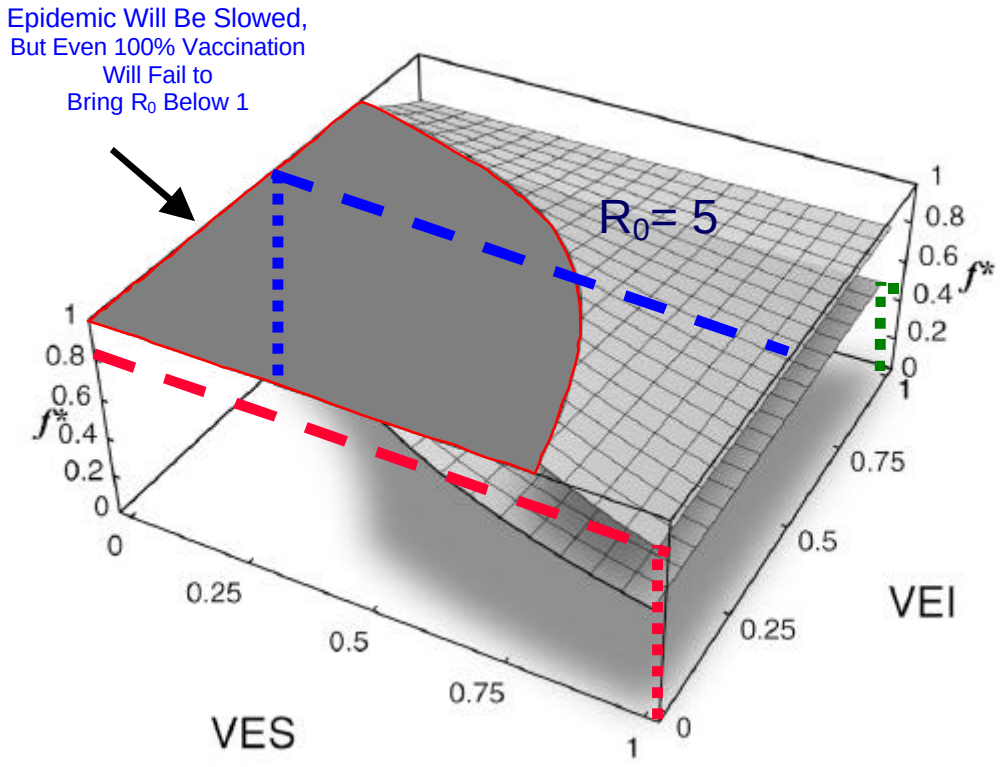
This can be placed in a formula as:

$$\mathbf{VE_R = 1 - ((1 - VE_S) (1-VE_I))}$$

This can be interpreted as the proportional reduction in the reproduction number caused by vaccination.

From this we can determine the critical vaccination fraction: f^*. This is the fraction of the population that would need to be randomly vaccinated in order to bring the R_0 below 1, thus halting epidemic HIV transmission.

The graph below shows the vaccine effect in infectiousness (VEI) on one side and the vaccine effect on susceptibility (VES) on the other. There are two different points: one for the reproduction number of R_0=2 and the other R_0=5. This shows that even if 100% of the population is vaccinated the R_0 cannot be below 1. However, any strategy that can slow down the epidemic is good.



As observed on the graph, the higher the R_0 the less effective the vaccination strategy.
It is possible to have a vaccine that has no effect on susceptibility but a high effect on infectiousness, which is also a viable public health response.

Besides the determinants of vaccine impact already mentioned (VE_S, VE_I, VE_P), there are other important determinants, such as:

- **D_v**: the average vaccine duration before waning.
- **R_{cv}**: the relative contact rate increase for vaccinated (disinhibition).
- **P_v**: the proportion of the population that will be vaccinated.
- Epidemic stage at which you are vaccinating (early, middle, late).

The following table summarises the determinants of vaccine impact in terms of (1) combined vaccine effects on the reproduction number of the HIV epidemic (VE_R) and (2) who benefits:

VE_R	Benefits
↑ VE_S - ↑ VE_R,	Individual +, Population +
↑ VE_I - ↑ VE_R,	Individual 0⁺, Population +
↑ VE_P - ↓ VE_R,	Individual +, Population –
↑ D_V - ↑ VE_R,	Individual +, Population +
↑ R_{CV} - ↓ VE_R,	Individual –, Population –
↑ P_V - ↑ VE_R,	Individual +, Population +
↑ Epidemic Stage - Later Stage, ↓ Prevention	

By default a high VE_S increases the VE_R, which is good for the individual and the population. A high VE_I increases the VE_R but it might or might not have effects on individuals, etc.…

Models have become essential in the evaluation of infectious disease interventions because they allow the estimation of indirect protection by the intervention. It is often the case that many individuals within the population can be protected by the intervention even if they actually did not receive the intervention, e.g. through herd immunity.

The HIV Vaccine Project

The original objective was to determine the optimal distribution of an HIV vaccine to limit the epidemic in scenario countries. For this many assumptions were made, such as:

- a moderately effective vaccine;
- limited quantities available; and
- model HIV vaccine distribution to heterogeneous population with distinct risk and vaccination-eligible groups.

This project evolved into a WHO-UNAIDS collaborative group on cost-effectiveness, delivery and future access to HIV vaccines, with different project partner countries such as Thailand, Kenya, Brazil, Peru, and China.

The findings from this investigation are intended to help public health policy-makers and planners to:

- assess the potential epidemiologic impacts of future HIV vaccines in their country's context; and
- determine robust and cost-effective HIV vaccination strategies.

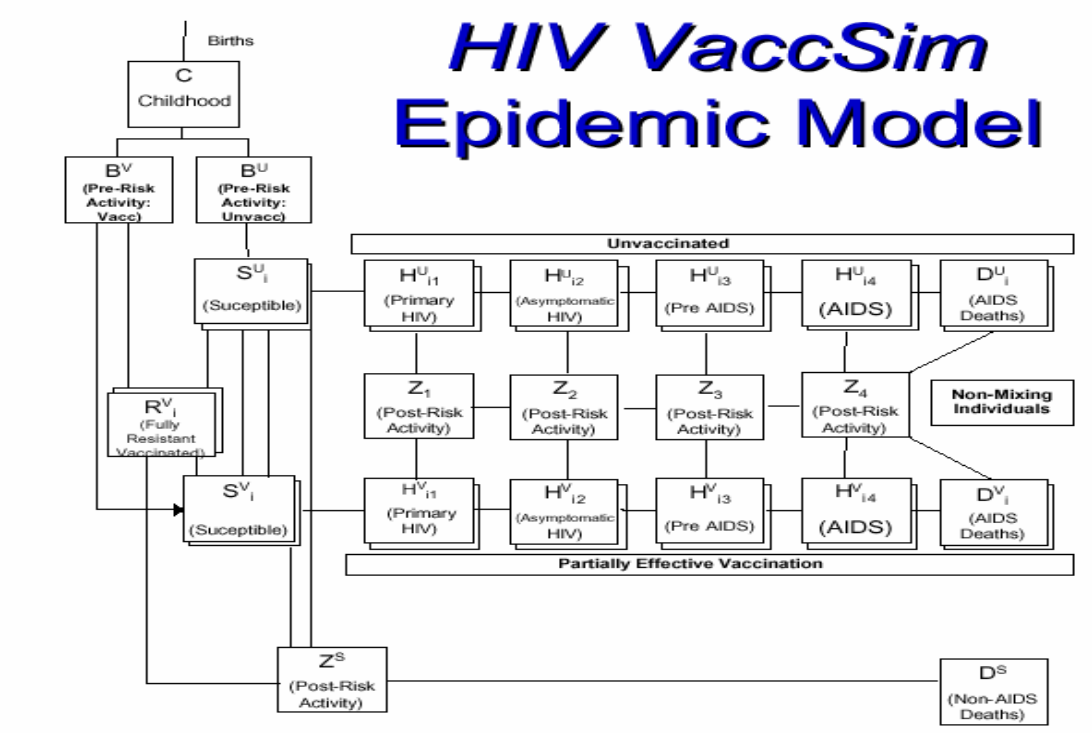
The HIV VaccSim Model

The idea of this model was to create a very user friendly program to put into the hands of policy-makers and epidemiologists without prior epidemic simulation modelling backgrounds, so that they could interact with the model.

The model integrates demographic, epidemic surveillance, policy-maker survey data and cost data for the country-specific cost-effectiveness analyses. The model/program is designed to assist policy-makers in making evidence-based decisions on the potential epidemiologic impacts of future HIV vaccination strategies, taking into account country-specific HIV epidemiology, costs, ongoing HIV/AIDS prevention and treatment programs.

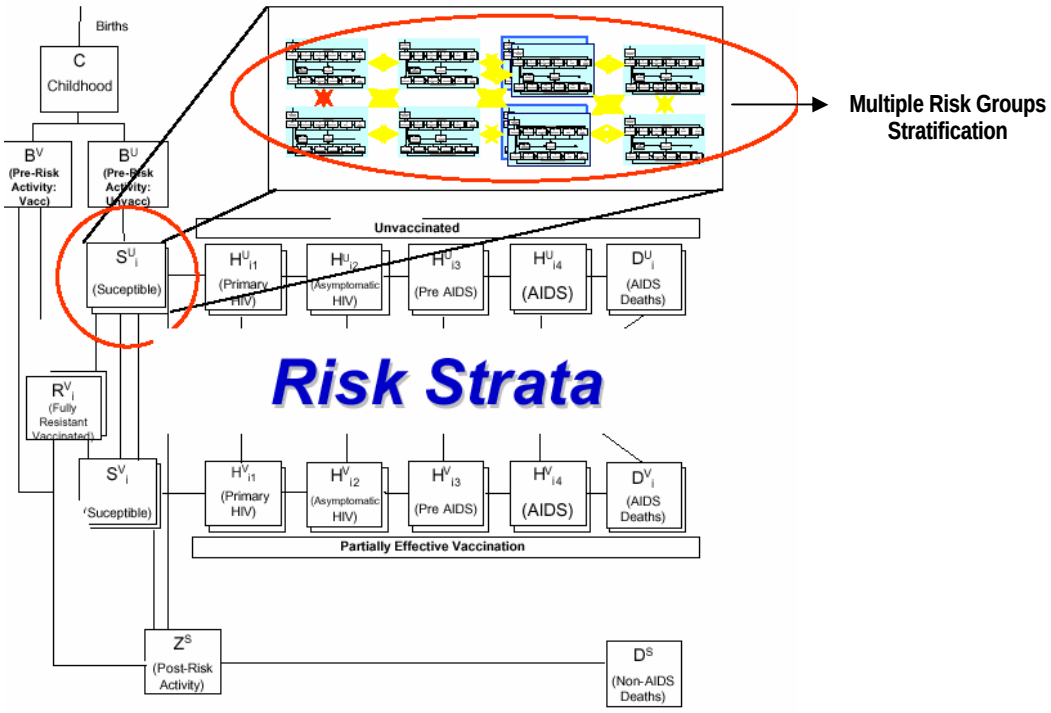
The HIV VaccSim is a fairly large deterministic mathematical model consisting of a non-linear system of differential equations and is used to dynamically model the HIV epidemics. It relies largely on uncertainty analysis methods to determine the impact of uncertainty about the model input parameters.

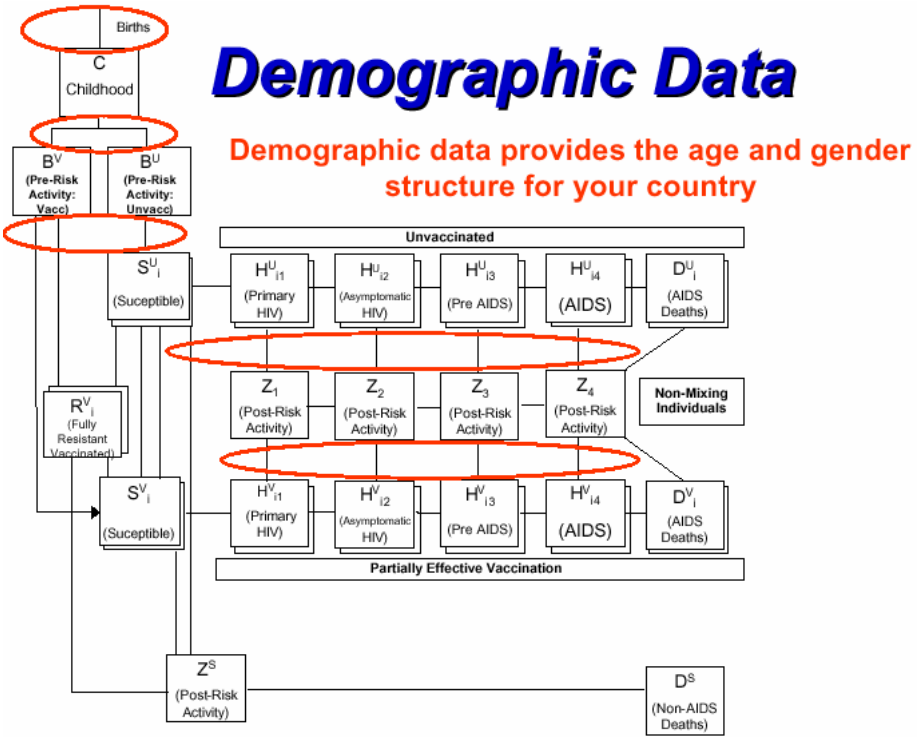
The HIV VaccSim takes into consideration the synthesis of data from multiple sources such as: demographic data; estimates of at-risk population sizes; behavioral survey data; research literature on HIV transmission probabilities; national cost data for vaccination programs and HIV monitoring, treatment and care; epidemiologic surveillance data such as prevalence, incidence, number of cases, deaths; and national HIV vaccination strategies to be considered in the given country.

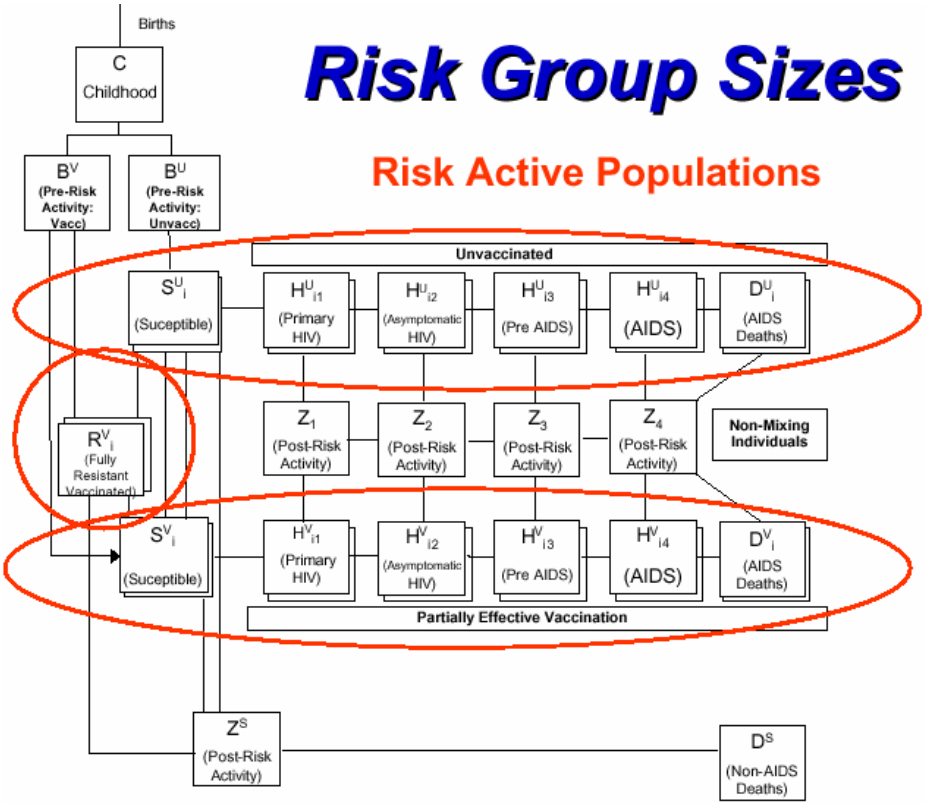


Understanding the model and data

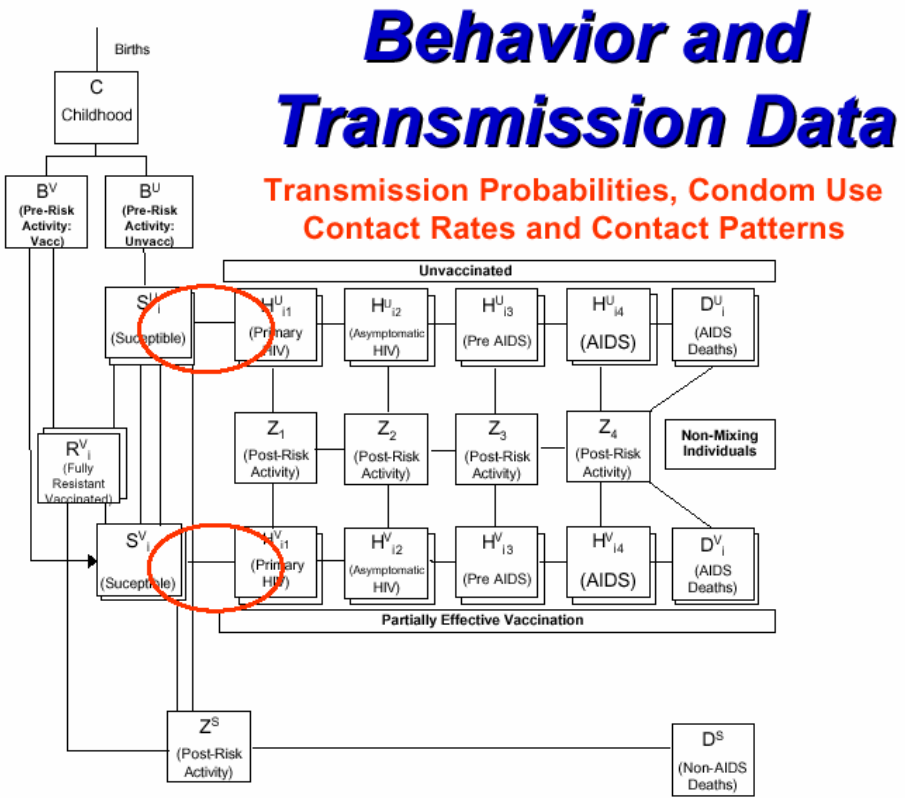
The model includes multiple risk groups stratified on different levels of partnership change, main injection drug users, heterosexuals, homosexuals, etc. Within each of the model compartments the mixing process takes place.



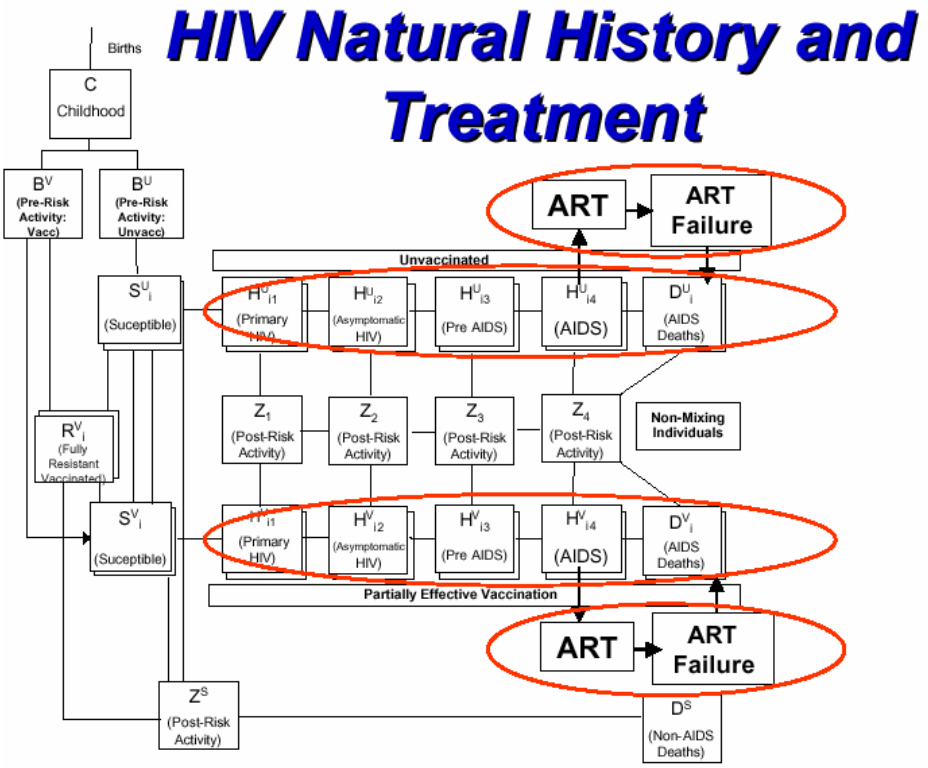




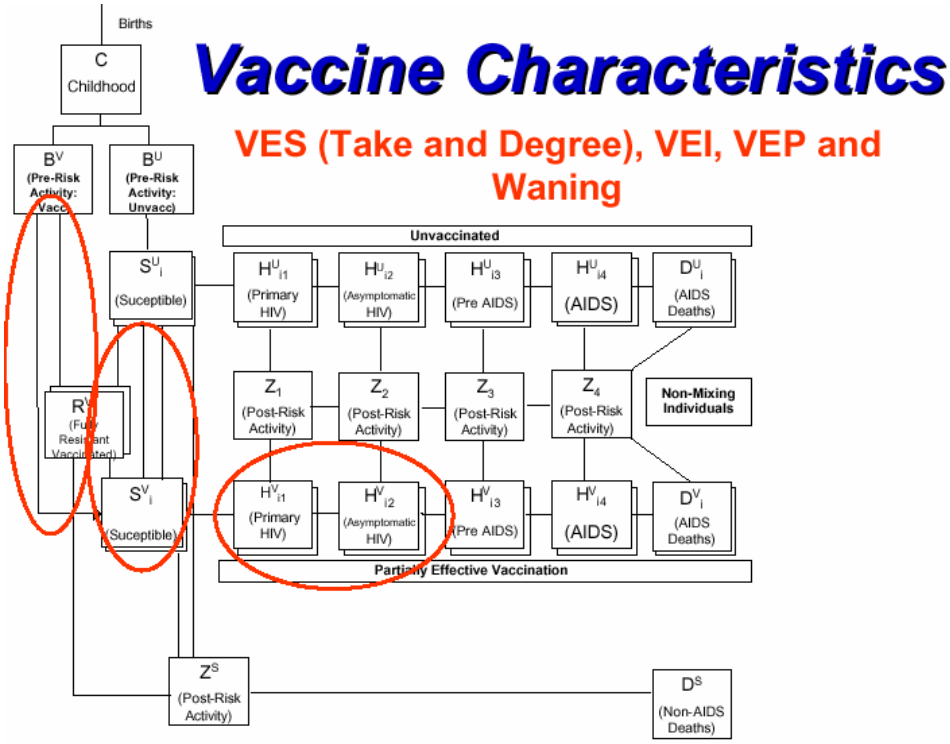
The risk group sizes determine how many people are at risk; the double boxes shown above indicate that in the model there are separate risk strata.



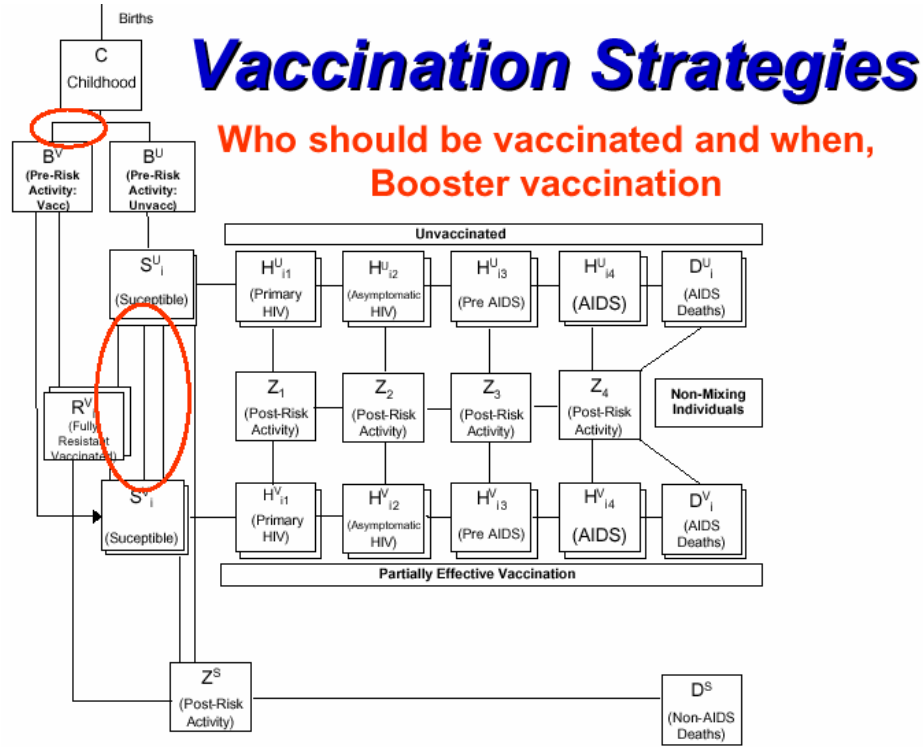
The transmission patterns described above explain the likely infection rates for people who are fully susceptible or partially protected.



There is also a need to have information on the stages of HIV: how long do people remain in the different states; and if they receive treatment, how long does that delay the onset of AIDS?



The above model describes the vaccine characteristics and helps determine the vaccine's effect on susceptibility (VES) in a mixed model with both "take" and "degree"; the vaccine's effect on infectiousness (VEI); the vaccine's effect on progression (VEP) and the waning.

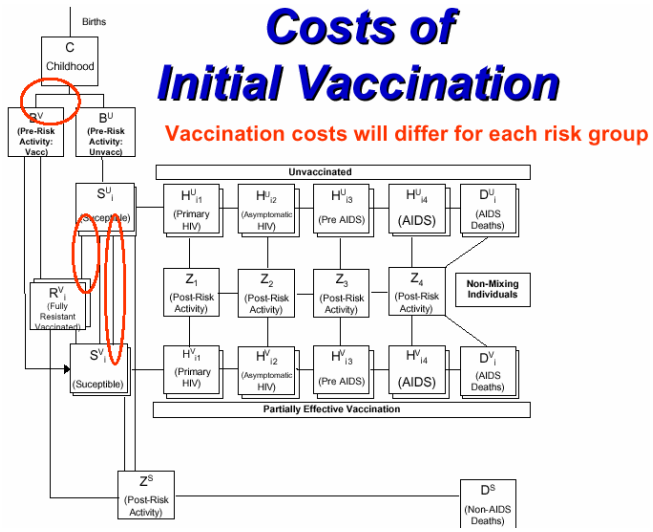


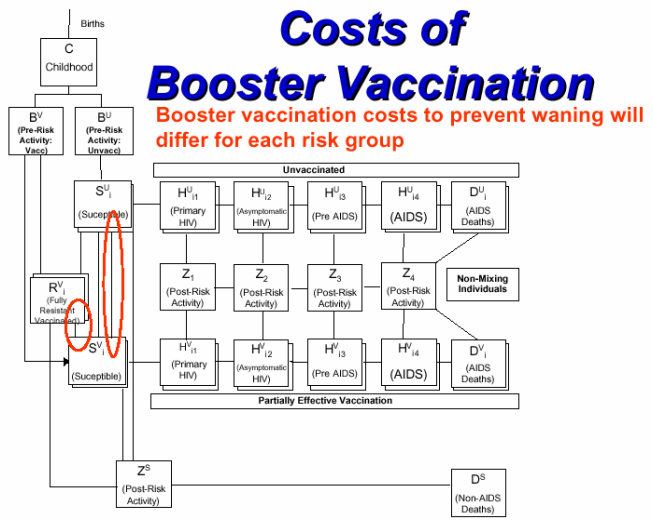
This model helps decide on vaccination strategies: who should be vaccinated and when, as well as when booster vaccinations are needed.

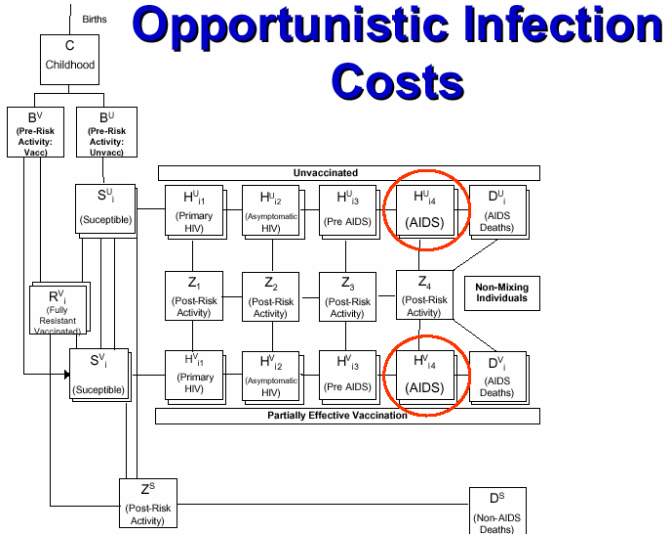
Vaccination process

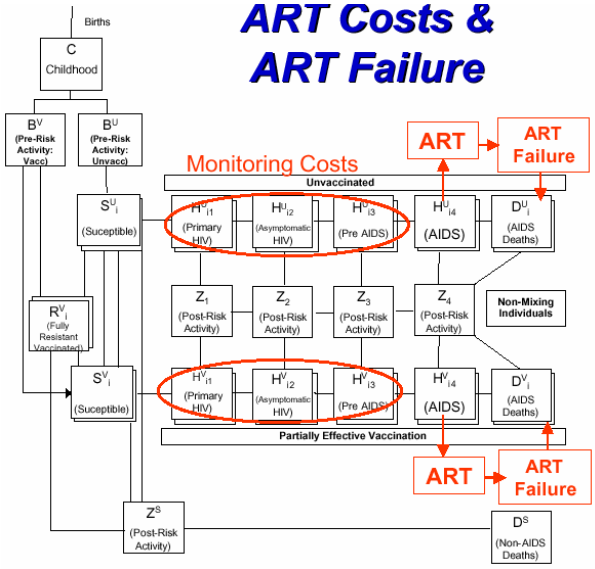
- Annual vaccine allocation: a fixed number of vaccine doses are available per year. This allocation can be modified to cope with changing vaccine availability over time.
- Vaccination priority: high *versus* low priority assigned to each vaccine-eligible group. High priority target proportions are met before low priority vaccination commences.
- Unvaccinated susceptibles in each group are vaccinated so that the target proportions are met and maintained as soon as possible, given the vaccination priorities.
- Determine when the vaccination program must begin.
- Target proportions for vaccination (what are the initial and eventual levels of vaccine acceptance in the vaccine eligible groups and how long does it take for vaccine acceptance to reach maximal levels).
- Determine how much vaccine will be available over time; the annual vaccine supply.

Some costs that are taken into consideration in the model for the HIV VaccSim Development are: vaccine costs, delivery costs, booster vaccination costs, HIV screening costs, CD4 monitoring costs, HIV treatment and care costs.
The following figures show how the model includes these costs:

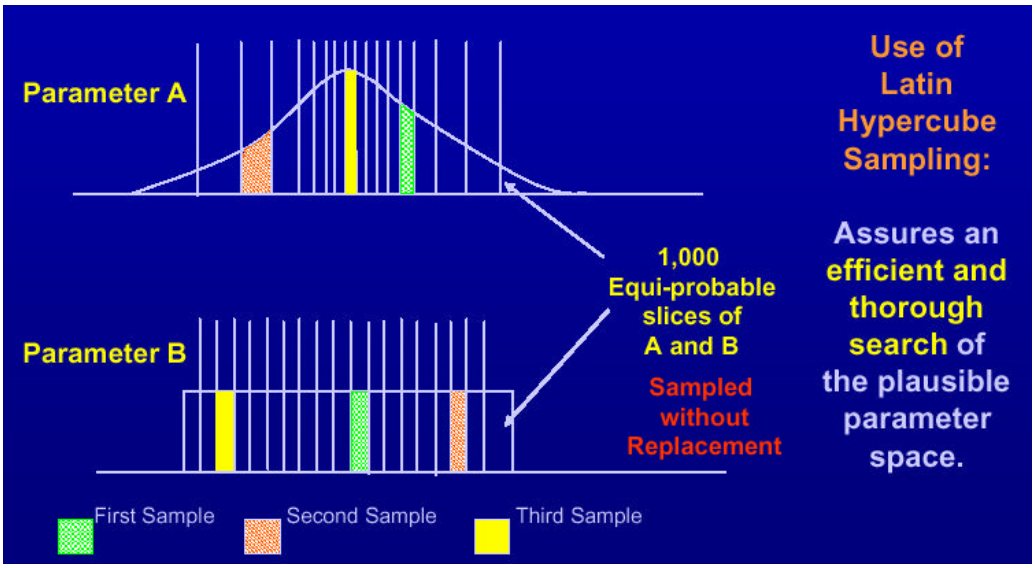




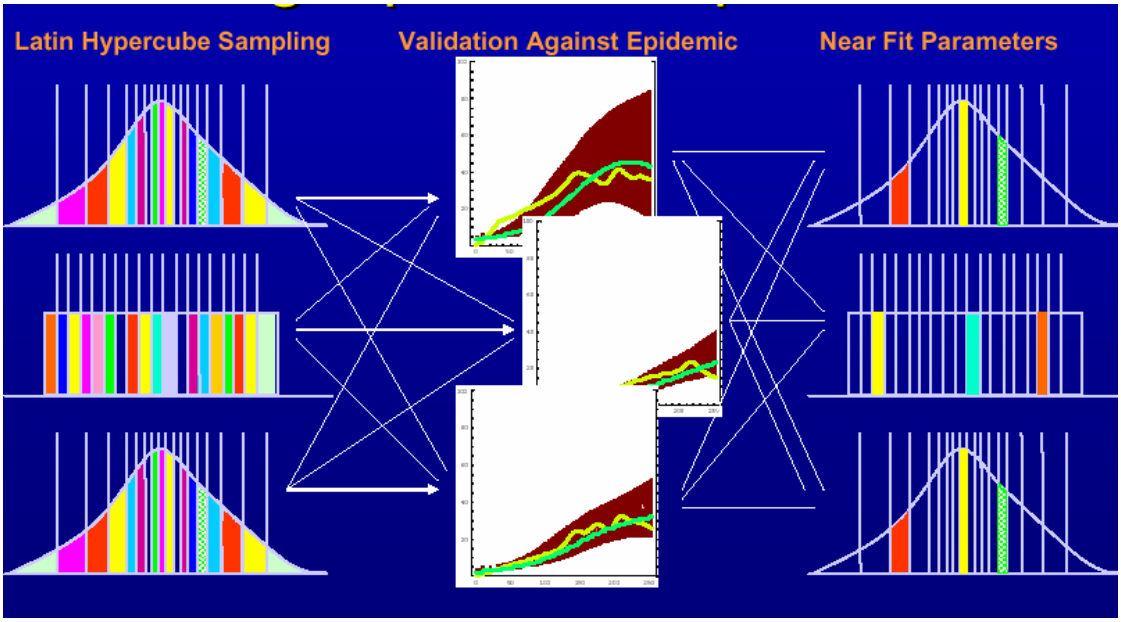




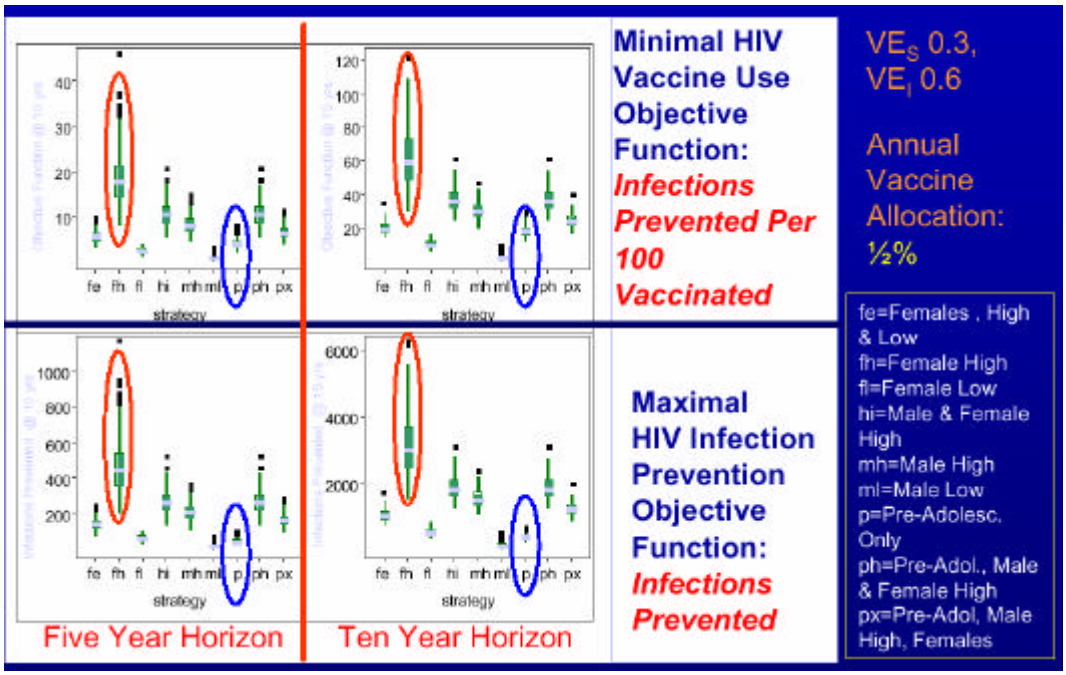
Two simultaneous versions of the simulation were performed, one with the vaccine program and one without it, and then compared. Using latin hypercube sampling in uncertainty analyses with the parameters, distribution was developed based on the literature and available information on the input parameters (see the graph below).



For model validation fitting inputs and epidemics, an approximation is done to make sure there is not or little uncertainty in the input and output parameters. For this the "near-fit" method is used to ensure that randomly sampled sets of input parameters fit observed epidemics. And then the set of combined parameters are retained that have an approximate fit to the epidemic.



When performing the uncertainty analysis, entire distributions are evaluated rather than single runs. This is shown below for an uncertainty analysis of a vaccination time of five years.



The aspects that are important in influencing HIV vaccine distribution strategies are:

- The stage of the epidemic.
- The size and contact-patterns of risk groups.
- The vaccine acceptance of risk groups.
- The specific vaccine effects on VE_S, VE_I, VE_P and waning.
- Vaccine related disinhibition.
- Quantity of vaccine available.
- HIV control objectives (prevent infection, AIDS, etc)

Using the HIV VaccSim Model

The HIV vaccine strategy analysis is best undertaken by interdisciplinary teams that include epidemiologists, health economists, statisticians, modelling methodologists, vaccination program and HIV prevention policy-makers.

Conclusions

- Vaccination strategies will be a critical determinant of potential HIV vaccine impacts, particularly when vaccine supplies are limited or vaccine effects are moderate.
- Modelling methods can contribute importantly to the development of robust and effective country-specific HIV vaccination policies.

Discussion

**I am very interested in what you call the disinhibition factor; how do you take this factor into account in your analysis and do you model it explicitly?*
It is presupposed hypothetical, in the phase III trials there has been an effort to measure people's recorded contacts and whether there is a change in response to vaccination. So we put that in the hypothetical and try to account for that by increasing the contact rate.
**Because you have a number of risk categories in some age groups, you have to define the level of contacts between this person and another type of person, so how does the mixing balance of the equations work?*
Right now the equations are proportional mixing with gender balance.

Contact Patterns and the Spread of Infectious Diseases: POLYMOD Project

John Edmunds, Health Protection Agency, Centre for Infections, London, UK

As we know, the transmission of close-contact infectious diseases such as flu, measles, chickenpox, TB, etc. requires individuals to contact others for transmission to take place.
The contact patterns determine the pattern of spread and affect estimates of the effectiveness of control programs (vaccination, antivirals, behavior interventions etc).

The traditional (indirect) approach to the parameterization of these close contact patterns in an epidemic model is basically to assume the contact pattern. There are different approaches that have been used, such as: proportionate, preferred, and determine class & workplace size distribution to fit or calibrate the model to that epidemiological data.

Generally the estimation of the mixing patterns is not done directly, but is inferred indirectly from the fitted epidemiological data. However, we don't know if the mixing patterns that are assumed are the true mixing patterns. There is also an identity problem because the mixing patterns can be quite complex to identify clearly with the available epidemiological data into individual mixing parameters.

The question is: can we do this in a better way?

The direct approach: diary-based estimates do the following:

- Define potential "at risk" event, e.g., conversation, or physical contact or both, or other.
- Ask individuals to record details of contacts.
- Ask individuals the age, sex, setting, frequency & length of contact and touch.
- Record these details on a randomly assigned day.

The POLYMOD Project

- 7,300 participants, 98,000 contacts, 8 countries
- Population based sample, recruited by random digit dialing (4 countries), face to face (2 countries), and population registers (2 countries).
- Training done by phone or face to face.
- Sample size varied between countries between about 267 & 1,328.
- Also collected serological data from 5 countries from which people were not vaccinated against the Varicella-Zoster virus (VZV), and the parvovirus B19.
- Collection of data on childcare attendance, household sizes etc.

- Other data sources such as time-use surveys

No details on the results of the project will be given in this lecture as they will be published in the near future. However, in general we can say that there are indeed differences between countries, particularly in the number of contacts that an individual will incur. For example the Germans recorded the least number of contacts and the Italians the largest number of contacts in this survey.

With regards to age, contacts are higher particularly in the young age groups, but get fewer with older age groups. The 0 to 19 year olds tend to have more contacts than other age groups. When measuring the intensity of contacts in different ways, in terms of duration and settings, results showed that longer duration of contacts occur in settings such as home, and tend to occur on a daily basis.

We also determined where these contacts occur. If within a social context, what is the social distancing, such as on public transport, or at work, home, school, etc? Once again the multiple contacts generally occurred at home.

Serological data were collected from 5 countries on 2 SIR infections not vaccinated against. A transmission coefficient was estimated through the contact pattern data, using maximum likelihood. A similar level of the coefficient was obtained across the different countries.

The results obtained for the 5 countries showed that VZV is more transmissible than B19. When splitting the results between physical and non physical contacts, the results showed clearly that physical contacts are more important for spreading the infection.

The coefficient obtained for non-physical contacts for B19 and VZV for each of the countries was not significantly different from zero; physical contact appears to have more explanatory power.

Time-use (TU) surveys

Many of the countries conduct regular TU surveys to record activities throughout the day (eg. whether alone or with someone else…). This information can be used to construct contact matrices. The following assumptions have been taken into consideration: within a location and small time interval, individuals (of different age groups) divide time proportionately to others in that setting at that same time. For example, on public transport at a given time, when looking at the age distribution, it is observed that people are mixed proportionally according to age distribution.

The results obtained can help to build an age-specific time exposure matrix, which can be divided by setting and fit to serological data.

Summary

- The diary approach allows models to be informed by relevant data, much of which is not routinely available. Moreover, it is flexible and simple as the model structure dictates data correlation.
- Allow similar patterns of mixing between countries, and they are diverse with regards to age and probably other variables.

- The diary approach can quantify varying risk in different settings, such as correlating measurements of intimacy. Some settings are likely to be lower risk for multiple reasons.
- The results suggest that more intimate contacts may be more important for the spread of common infections.

The POLYMOD project was limited in that it did not collect distance information, and it is anonymous so it is not possible to elucidate networks. It is currently being adapted for use in a number of other countries. A summary paper is under review, and the individual-level data are to be made publicly available in September.

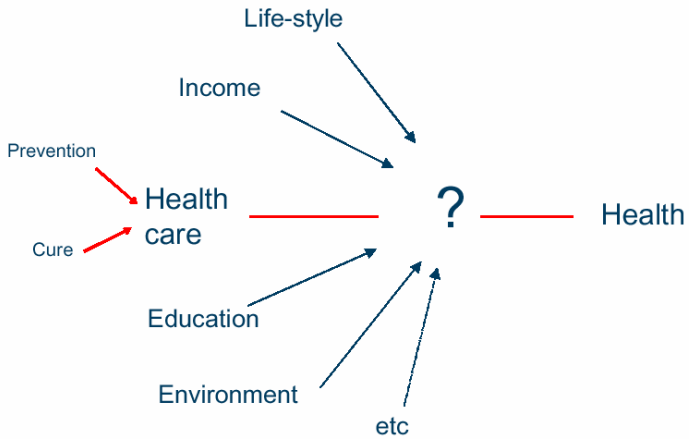
Discussion

**I am curious to know about the data from the Netherlands you mentioned on serological surveys done to people, is that going to be available next year?*
I think data will come out gradually, not sure all by next year.
**Do you plan to work on the fact that the disease by itself is modified by the contact?*
In the survey for flu that I mentioned earlier looking at incident patterns, actually we are getting individuals to record contact patterns when they are real.

Modelling Options for Economic Analysis: Realism vs. Pragmatism and Fiction

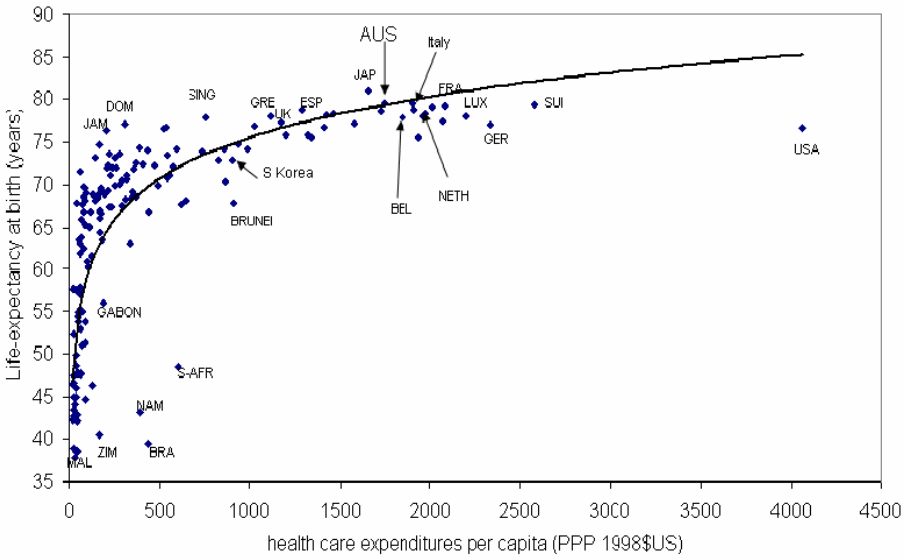
Philippe Beutels, Health Economics & Modelling Infectious Diseases, Centre for the Evaluation of Vaccination, Vaccine & Infectious Disease Institute, Antwerp, Belgium.

In the context of the economy, and in relation to health, the basic view is of health as "production". In order to see how to produce as much health in the population as possible, we have to look at the various inputs, as shown in the following diagram:



However, when looking at health care expenditure in each different country alongside life expectancy, we see that there is not necessarily a correlation between expenditure and the quality of health care provided by a country, especially within rich countries. The more a

country spends does not necessarily mean that more and better health care are provided, see graph below.



Beutels P, 2002, based on WHO and WB

For instance, the USA spends more than twice as much as the healthiest country in terms of life expectancy (Japan).

In implementing a program the following should be known preferably in advance:
1. Efficacy and safety (does it work, is it safe?).
2. Effectiveness (how well does it work in the real world?). This is where modelling comes in.
3. Efficiency (how do the costs relate to the effectiveness?).
4. Equity (does it (dis)advantage subgroups of the population?).

Efficiency and equity are the main focus of health economics/welfare economics.
As noted, efficiency is an integral part of the effectiveness question; thus in order to know how efficient things are, you need to know in advance how effective they are.

In drug regulation there is a fourth hurdle: quality , safety, efficacy, which correspond to the Phase I, II and II clinical trials respectively, and efficiency (cost-effectiveness), which is usually combined with trial results which are mostly model-based.

This hurdle is for drugs as well as vaccines, and in countries were there is a mandatory fourth hurdle, they tend to treat vaccines as they treat all drugs. This is important to bear in mind.

The ICER: Incremental Cost-Effectiveness Ratio

$$ICER = \frac{COST}{QALY}$$

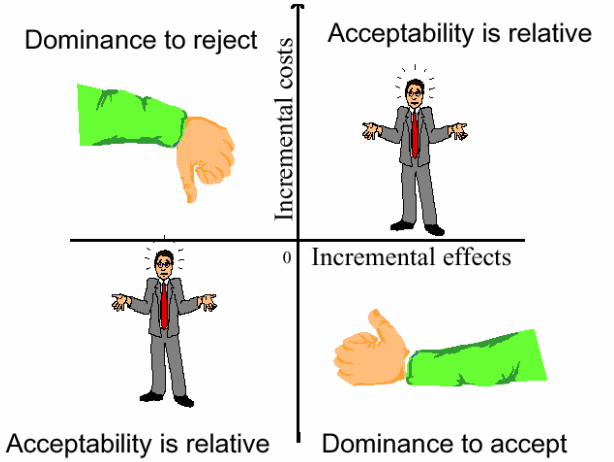
QALY stands for quality of life years, and is the most widely used measure in health economics today.

The main point to note in this equation is that the difference you make will be highly dependant on the choice of comparison. For example, if you compare vaccinating to not

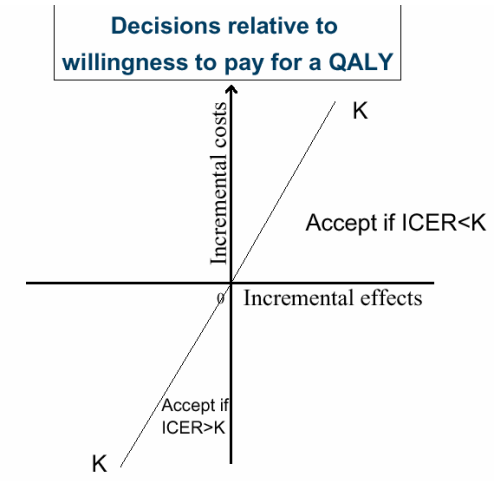
vaccinating, etc…this will have a strong impact on your ICER results, so it is important to be aware of the comparator you are using.

All of the countries that have a mandatory fourth hurdle required transparency in the model. All parameters in the model are based on data and distributions derived from the data are an integral part of the model. For this it is necessary to have all distributions, all costs inputs, and all health outcome inputs to get a distribution on the cost/effectiveness ratio.

If a new intervention (drug, vaccine, etc) is plotted against the comparator on the cost-effectiveness plane, you can position your intervention somewhere as follows:



Then the decision to accept an intervention will depend on the willingness to pay for an additional increment of effectiveness.



According to the above graph the willingness to pay for a QALY is measured by K with regards to the ICER. So the question is how much is the "K"? This changes by country with regards to a given intervention.

Since determining these measurements can be a complex task, if you do multivariable sensibility analysis, taking part of the uncertainties into account, and determine the probabilities of ending up in one of the four quadrants of the cost-effectiveness plane, then the

cost-effectiveness acceptability curve is used. This is becoming a standard approach in CEA/CUA.

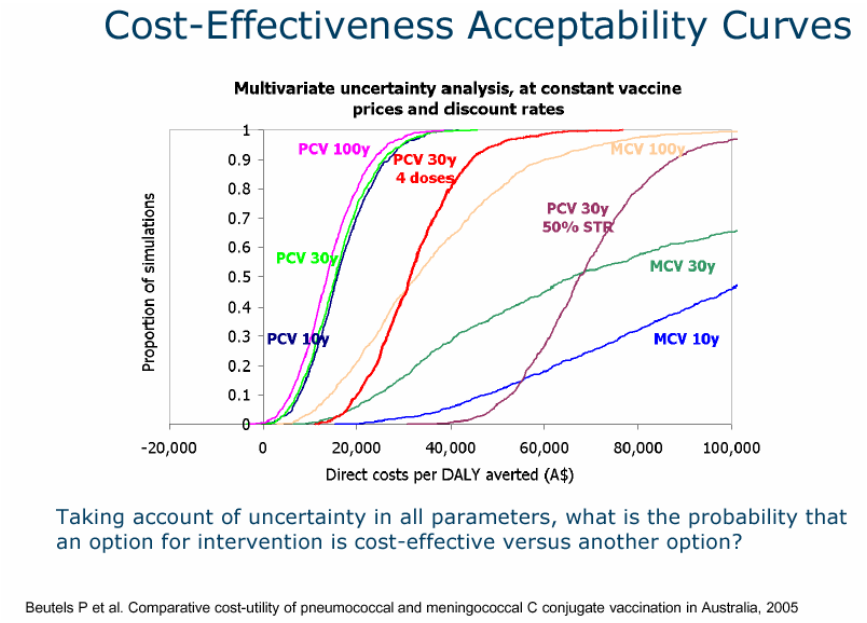
The cost-effectiveness acceptability curve function is as follows:

CEacc(K)

$=Pr(ICER < K \mid \Delta E>0) \cup Pr(ICER > K \mid \Delta E<0)$

$=Pr(NB(K) > 0)$

Example of a cost-effectiveness acceptability curve for pneumococcal and meningococcal "C" conjugate vaccination in Australia in 2005:



Beutels P et al. Comparative cost-utility of pneumococcal and meningococcal C conjugate vaccination in Australia, 2005

Usefulness of economic evaluation of vaccines

The first generation of vaccines (before 1975) were cheap and prevented common and serious illnesses, thus the effectiveness of their use was obvious and there was no real need for economic analysis.

The current generation of vaccines is expensive due to technical and regulatory complexity, and they are for diseases that are not so common (e.g., meningococcal C) and/or not so serious (e.g., chickenpox).

The idea of economic evaluation is thus to support prioritisation of vaccines and vaccination strategies (schedule, age group, etc) vis-à-vis other health care interventions, and price negotiations. It can also help to understand uncertainties regarding vaccine decisions.

Specific issues for the economic evaluation of vaccines

- Herd immunity
- Very sensitive to analytical time span and assumptions regarding time preference (discounting).

• Often short-lived illness (often in very young children), which causes extra familial care and work loss, for which valuation methods lack credibility and acceptability.
-quality of life assessment
-indirect time cost estimates
• Some infections are eradicable.
• Some emerging infections (eg, SARS, pandemic influenza) would have a major macroeconomic impact that goes beyond lost productivity of the sick people and their families.

The epidemiological consequences of childhood vaccination:
*The force of infection declines.
*The average age at infection increases.
*The inter-epidemic period increases.
This has consequences for the occurrence and severity of the illness, especially with the average age of infection, as some illnesses get more severe with age.

Different modelling options
*Deterministic – Stochastic
*Individual based – grouped
*Discrete – continuous (age and/or time).
*Open – closed
*Type 1 – type 2 mortality.
*Spatial – non-spatial
*Static – dynamic.

In the model there is a question of whether or not you take into account the effect of herd immunity; this has implications for effectiveness, efficiency and equity.
In addition to this pattern of uncertainty there is model uncertainty which is often not tested. For example:

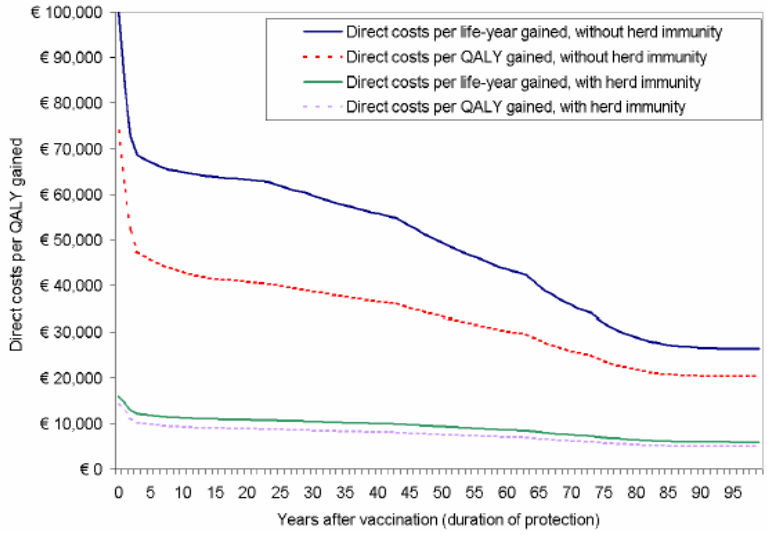
The static model: typically a deterministic Markov model for a single ageing cohort. The force of infection is independent of the proportion infectious at each point in time. Herd immunity can only be introduced into the model based on observations from a similar setting. However, static models are easy to develop and belong to the traditional toolbox of health economists and epidemiologists.

The dynamic model: typically a deterministic population-based model, with a constant total population size over time. The force of infection is recalculated as a function of the proportion of infectious people at each time point, and herd immunity impact is a built-in part of the model. In these models the underlying infectious disease transmission process is modeled. They need data or assumptions on mixing patterns and duration of infectivity and they are not part of a traditional health economist or epidemiologist's toolbox.

In the modeling practices for economic evaluations people tend to use the static more than the dynamic models for evaluating both vaccines and drugs. And for a number of vaccination programs this is wrong, for this there is an increasing tendency to use more dynamic models for more recent analysis. For example for Varicella-Zoster virus (VZV) from 2000 until now, there have been nine analyses, six of which have used the dynamic model.

Huge differences may occur by including herd immunity, the quality of life years, and a longer time frame

Childhood pneumococcal conjugate vaccination in Belgium
(Beutels et al, 2006)



Statistical modelling of past decisions to fund

Probability of recommending funding (YES/NO recommendation) explained by:

- Incremental cost per QALY
- Annual additional cost per patient
- Average QALY gain
- Annual predicted cost to government
- Clinical uncertainty
- Economic uncertainty
- Toxicity
- Highest cost per QALY in sensitivity analysis

The trouble with recent vaccines in affluent countries

• Pneumococcal conjugate: €150-200 per person, there is a need to reduce the schedule or herd immunity to be cost-effective, and has a limited production capacity (2000-2004). (See Beutels et al, Vaccine 2007)

• HPV: €250-350 per person, the long term effectiveness is uncertain, and there is an effective alternative: screening. (See Newall et al, Lancet ID 2007)

• VZV: €15-50 per person, is not cost-effective for health care system, has indirect effects which are uncertain (more zoster). There is a possible quality of life impact on young children, and risk of post herpetic neuralgia? (See Brisson & Edmunds, MDM 2006)

• Rotavirus: €60-120 per person, there is a possible quality of life impact in very young children and their parents. Don't know how much gastro is due to rotavirus. (See Newall et al, Vaccine 2007)

Conclusions

• Economic evaluation and modelling are not exact science, but they help policy-making.

• For vaccines, economic analysis is more difficult, and often with more uncertainties than for curative drugs.

• Realism vs. pragmatism vs. fiction: Inform decisions in the interest of speed, when cost /effectiveness decision rule is independent of simpler/more realistic model structure.

- *"Make everything as simple as possible, but not simpler" (Einstein?)*
- *"The future, according to some scientists, will be exactly like the past, only more expensive." (John Sladek)*
- *"Wise men make proverbs, but fools repeat them." (Samuel Palmer)*

General references

Beutels P. Economic evaluation of vaccination programmes in humans: a methodological exploration with applications to hepatitis B, varicella-zoster, measles, pertussis, hepatitis A and pneumococcal vaccination. January 2002, Antwerp: University of Antwerp (ISBN 90-9015528-7), 523 pp.

Beutels P, et al. Methodological issues and new developments in the economic evaluation of vaccines. Expert Reviews of Vaccines 2003; 2(5): 89-100.

Beutels P. Economic aspects of vaccines and vaccination: global perspective, p. 3-20, In: Advances in Infectious Diseases : The grand challenge for the future: Vaccines for poverty-related diseases from bench to field. Stefan H.E. Kaufmann & Paul-Henri Lambert (eds), Birkhäuser Publishers, Basel, 2005.

Beutels P, et al. Economic evaluation of vaccination programmes: a consensus statement focusing on viral hepatitis. Pharmacoeconomics 2002

Brisson, M & Edmunds WJ. Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs. Med Decis Making 2006
Edmunds WJ, et al. Evaluating the cost-effectiveness of vaccination programmes: a dynamic perspective. Stat Med 1999

4. Session IV: The Future of Infectious Diseases Modelling

Public Health Authorities Point of View: Data Needs and Analysis Requirements

Lara Wolfson, WHO, Geneva, Switzerland (First paragraphs taken exactly from Speaker's Abstract)

As decisions have to be made between competing priorities within the health and development sector, it is becoming increasingly important that these decisions are based on the best available evidence, including consideration of likely health benefits that will accrue, and sound assessments of the associated costs.

In this talk, an overview is given of the types of modelling and cost-effectiveness data required by decision-makers as they consider the future development of immunization programs from a public health perspective. Particular focus will be on the type of epidemiological and economic evidence that is required to consider the introduction of a new vaccine from both a global and country perspective.

Current tools, data repositories and challenges will be discussed, and examples given of different analyses used for advocacy in financing immunization. The use of evidence for priority settling, and the role of priority setting in global public health, will also be discussed.

It is a challenge for policy-makers to use models, but it is clear that modelling is needed. One of the reasons is that models help measure death and destruction, and this 'catchy phrase' is appealing to policy-makers as it gives them sound bites. Models try to give us educated guesses for measuring death and destruction because in the developing world very few deaths are actually registered. When reading the statistics about any disease, these are not based largely on primary data, they are mainly based on modelling and understanding the nature of the disease. However, this information can get us only so far; there is a need for further information such as the population of the country, vaccination coverage rates, etc. These data are inputs to the models, so policy-makers ask: What is the minimum required data to do modeling that might give a more specific result? And how flexible are models for

extrapolating the information from one country to another, and what kind of data do we need to do this?
The WHO uses routine surveillance systems. Although we know they are weak, we ask whether we could supplement them with target studies to make these models better?

There are different degrees of complexity in models; the aim is to know what model to use and where to apply the appropriate model to obtain the outcomes we need.

Policy-makers want to know how much they have to invest in getting the data to feed the models, and they have a hard time dealing with the uncertainty rate that exists in modeling.
The communication flow between modelers and policy-makers is key to reducing gaps in understanding modeling and its impact on public health.

Policy-makers also want to know how they can use these models to raise money? How to prioritize a disease? Thus, models are increasingly becoming the basis for funding allocation in public health.

There is a need for better models that can provide increasingly accurate predictions and that provide credibility to policy-maker's activities. There is also a need for policy-makers to be educated in the science behind modelling to understand that a degree of uncertainty is an intrinsic part of modelling

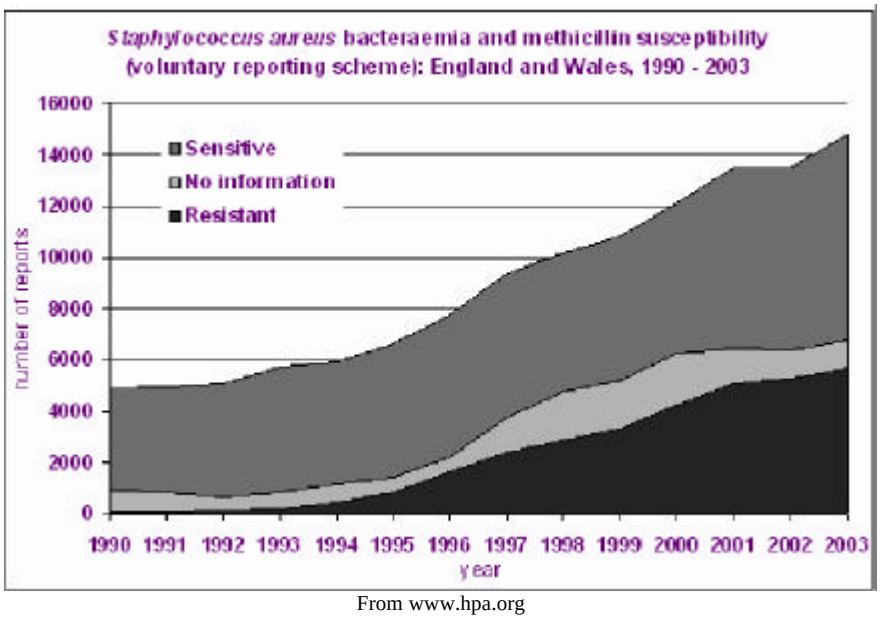
Modelling the Transmission of Hospital Infections

Ben Cooper, Statistics, Modelling & Bioinformatics Department, Centre for Infections, Health Protection Agency, UK.

This talk will address how to model the transmission of nosocomial infections; these are likely caused by pathogen bacteria that is normally carried asymptomatically, and by the resistant forms of the pathogens causing hospital infections.

Staphylococcus aureus is a typical pathogen of nosocomial infections; about 30% of the population are nasal carriers, and in vulnerable people it can cause serious life threatening diseases (pneumonia, bacteraemia, endocarditis, etc).

The Methicillin-resistant *S. aureus* (MRSA) is common in health care settings in many countries; though rarely found in the community. Prevalence is low (1-1%).



From www.hpa.org

The big increase in bacteraemia infections is almost entirely due to the increase in MRSA infections.

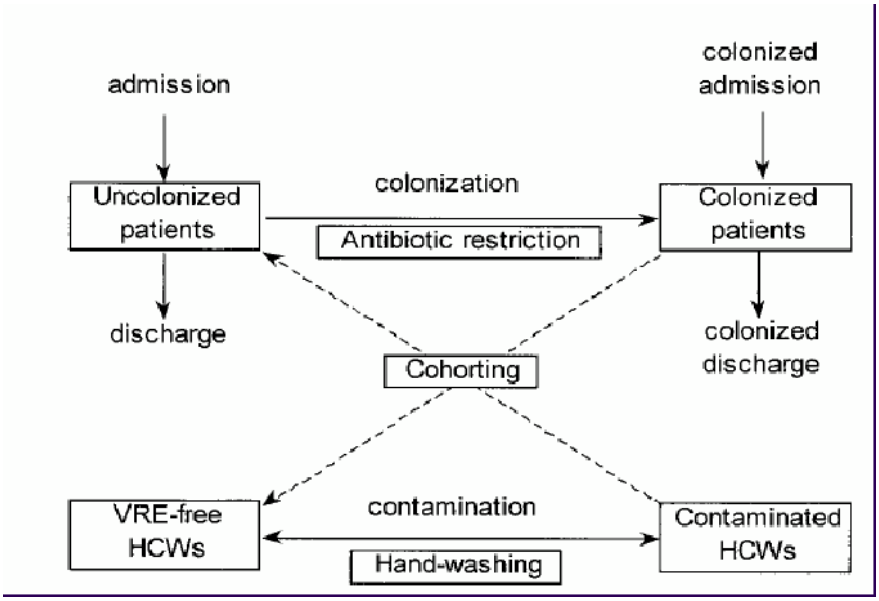
Another pathogen, among many others responsible for nosocomial infections, is VRE (Vacomycin Resistant Enterococci). This pathogen is asymptomatically carried in the gut but can cause wound infections and urinary tract infections UTIs. These types of infections caused by VRE are potentially untreatable and have in recent years led to a rapid increase in hospital populations.

The use of mathematical models for nosocomial infections serve:

- To gain insight into the consequences of interactions between simple processes (transmission, recovery, discharge, immunity, etc), and identify key factors affecting behavior.
- To suggest interventions most likely to be effective.
- To choose between competing hypotheses (represented by the models) by comparing fits of different models to data.
- To integrate evidence and generalize results of trial data to different populations.
- To aid decision-making, including economic decision-making.
- To aid forecasting.

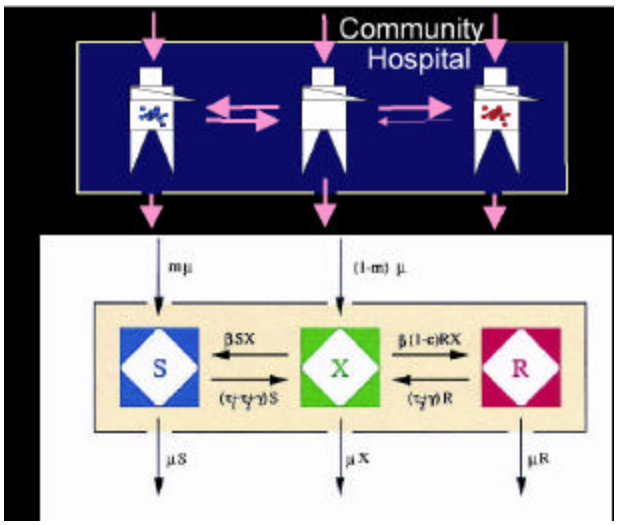
Many models have been built to address nosocomial infections. Transmission is mainly assumed to occur through the constant contact between patients and health care workers which allows the pathogen to spread in the hospital setting.

The following flow chart model was one of the first to come out regarding nosocomial infections:



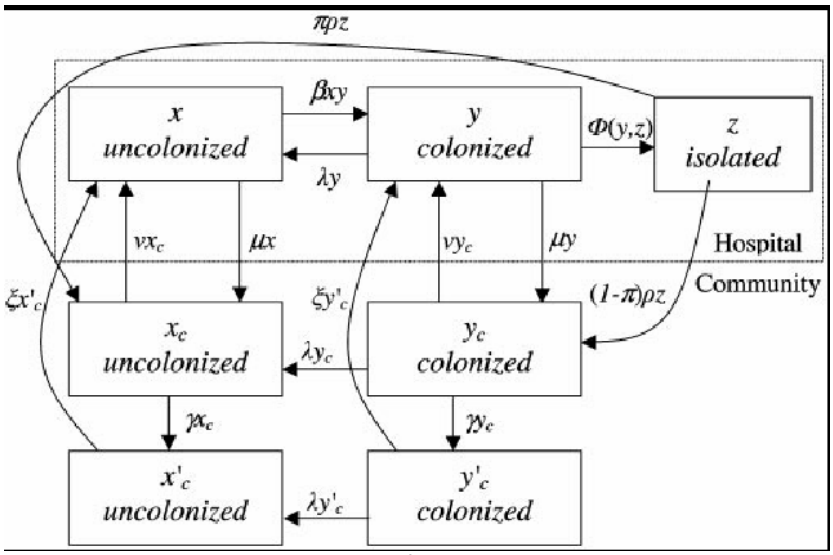
Austin et al. PNAS 1999

Then the model was further developed taking into account the interactions of the health care workers contacts, the sensibility of the pathogen resistant forms, and the antibiotics.



Lipsitch et al. PNAS 2000

But in order to take into account the gradual increases of MRSA, the model was revised to include a community reservoir. Even if there is no significant transmission in the community, there is a significant percentage of carriers which can last for a long time. The revised model is as follows:



Cooper et al. PNAS 2004

Key results from the "cartoon" models

- SIS dynamics obtained, reaching equilibrium after a few days.
- The stochastic effects are dominant because the population is very small.
- In contrast to the community in hospitals the resistance levels change quickly in response to changes in drug use and disappear quickly after a drug is discontinued or in response to other interventions.
- Non-specific control (e.g. handwashing) disproportionately reduces resistant infection.
- However, it is important to take into account that the long term dynamics are driven by changes in the community reservoir.

Translating research into clinical practices

The past:

• All models shown so far have been (to varying degrees) "cartoon" models, based on plausible model structures and guesses for sensible parameter values.
• They aim to provide insight into and reasonable guesses about impacts of policies. They tell us what the world would be like if our assumptions were true, not whether our assumptions are true.

The future:

• To use models to test our assumptions we need to use mechanistic models statistically & choose between competing hypotheses (models): this requires developments in model calibration, model assessment and model choice.
• To use models to directly inform clinical practice (through decision-analytical models) we also need models to more accurately capture patient heterogeneities, and to more fully characterise uncertainty in parameter values.
• Better models also need better epidemiological data.

The following is a simple hospital infectious model that shows how relatively easily with the likelihood method one can come out with estimates of the key parameters determining the model. The limitations are that it requires cross-sectional carriage data for which patients have to be screened at fixed time points.

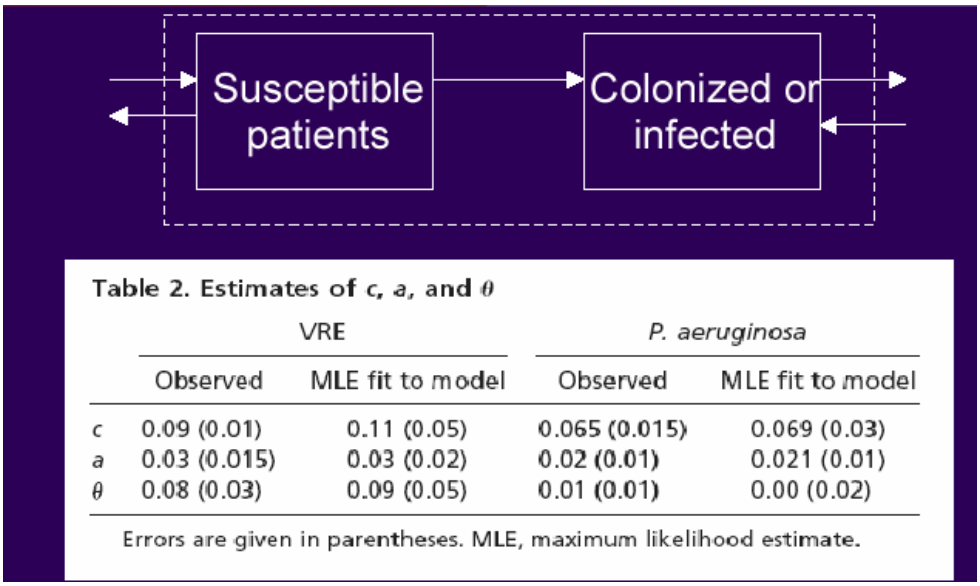


Table 2. Estimates of c, a, and θ

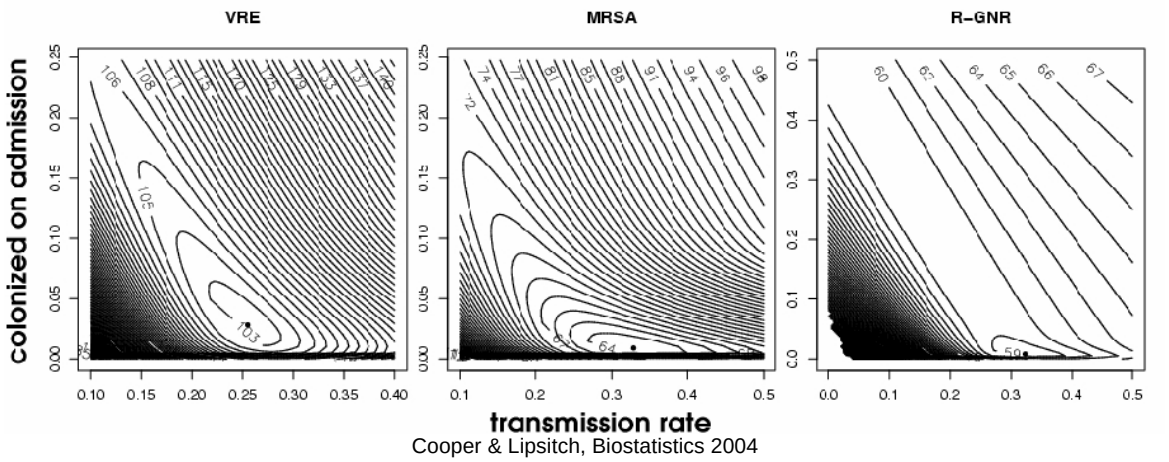
	VRE		*P. aeruginosa*	
	Observed	MLE fit to model	Observed	MLE fit to model
c	0.09 (0.01)	0.11 (0.05)	0.065 (0.015)	0.069 (0.03)
a	0.03 (0.015)	0.03 (0.02)	0.02 (0.01)	0.021 (0.01)
θ	0.08 (0.03)	0.09 (0.05)	0.01 (0.01)	0.00 (0.02)

Errors are given in parentheses. MLE, maximum likelihood estimate.

Pelupessy *et al.* PNAS 2002

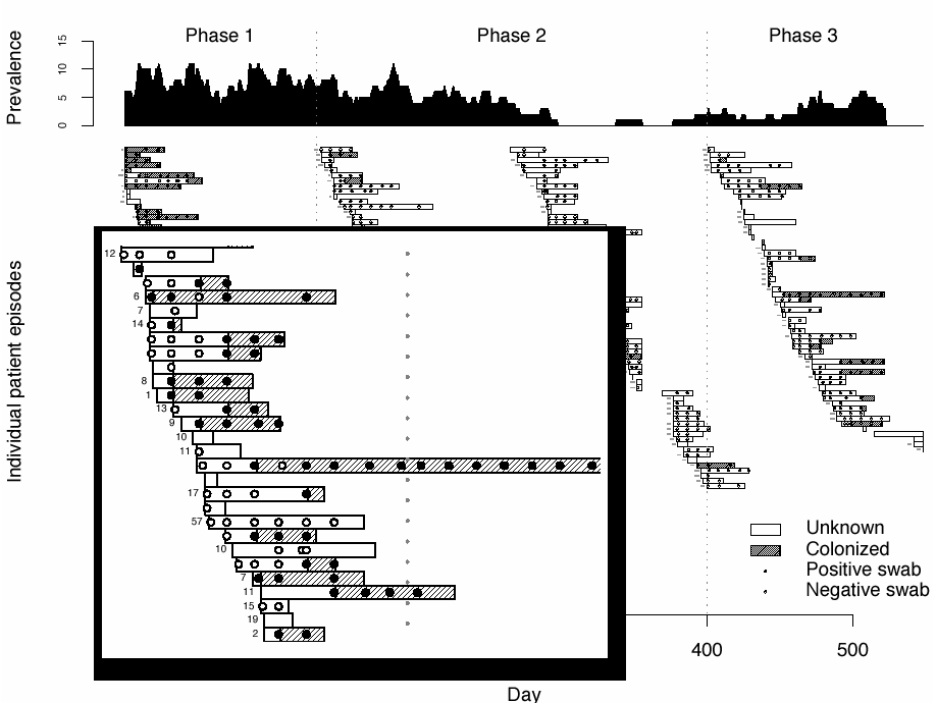
c=rate of replacement colonized by non-colonized patients
a=spontaneous colonization rate (including replacement of non-colonized by colonized patients)
? =transmission rate

More typically the only data you have are the infected patients from the specific pathogen, which is the tip of iceberg. Thus we need to extend the Pelupessy approach to a structured hidden Markov model to take into consideration how much transmission is going on in the hospital, and what proportion of patients are colonized on admission.



Cooper & Lipsitch, Biostatistics 2004

However, with these models there is also a certain degree of uncertainties such as false negatives, etc. (for example, someone can be colonized but still have a negative swab, etc).



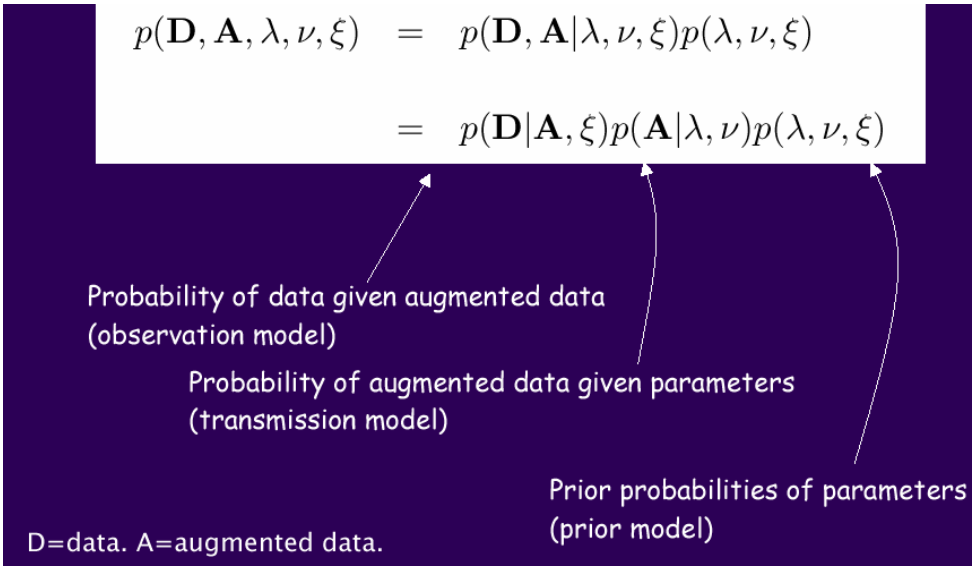
Fitting the model

If we know the exact date when people acquired the infection and who was colonized on admission, it would be easier to work out the likelihood. However, precise data on this are not always easy to obtain or available, and maximum likelihood estimation is therefore not possible. In response to this limitation, Auranen, JASA addressed this through the augmented data approach in 2000.

The aim is to assume possible scenarios in the augmented data and integrate them into the model. The solution is to *augment* the data with possible transmission times and events, and consider all possible processes consistent with the data.

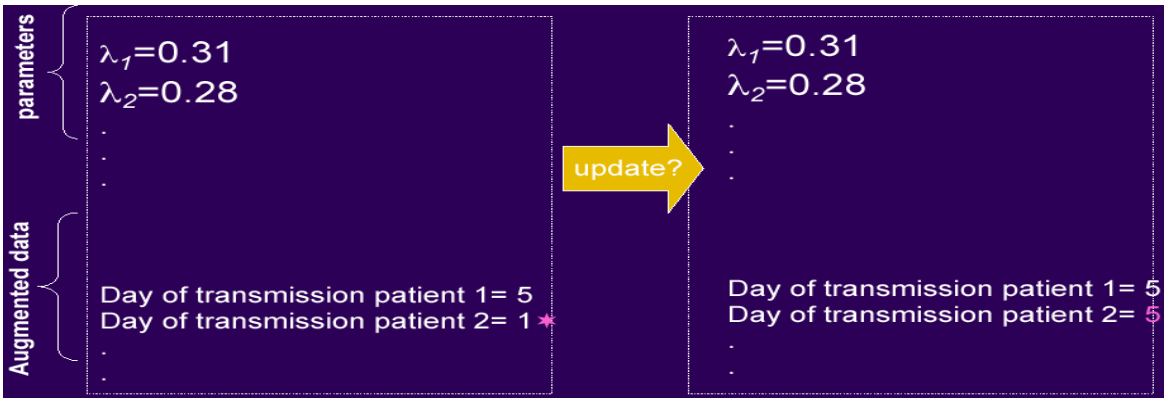
Estimation can be done using a Reversible Jump Markov Chain Monte Carlo (RJMCMC) algorithm. Below is an example of augmented data:

The inferences are done based on a Bayesian Theorem



Once you have an assumed value for the augmented data it is easy to write down likelihoods for each component of the model.

Basically the MCMC algorithm works by constructing a Markov chain with the required distribution as equilibrium distribution.



How the MCMC works
You start off with certain values for the parameters and for the augmented data and then you choose the parameters to change in the augmented data. Then you propose the new value and then accept the proposal with some specified probability chosen to ensure the chain has the required equilibrium distribution.

Although results with modelling are good, reality sometimes bypasses the results as every patient is different and some patients tend to be re-admitted. This needs to be taken into consideration.

The model can be extended. For example:
i) Assume patients colonized when discharged have a probability of being colonized when readmitted of exp(-? *t*), where *t* is time (in days) between discharge and readmission.

Mean duration of carriage following discharge is 1/?
ii) As i) but also allowing patient-level covariates to influence transmission parameters ($\log(?_i(t)) = ?_0 + ?_1 x_{1,i} + ?_2 x_{2,i} + ...$).
The correlations that can be done with the model data can be very important for the outcome results.
Major challenges with this model are to get a detailed understanding of the effects of antibiotics on transmission.

Conclusions

Augmented data MCMC methods:

- Can fit complex models "easily", eg. to account for swab sensitivity, post-antibiotic effect, different strain types.
- Prior information can be incorporated.
- Bayesian model choice can be implemented using algorithms that jump between competing models (ongoing work with Theodore Kypraios and Phil O'Neill).
- A big disadvantage is that this modelling is time consuming (likelihood methods are typically much faster to run and easier to implement, but less flexible).

Discussion

**I wonder if you took into account in your model any spatial information such as number of beds?*
Yes, we took some information data from the ICU that will be relevant spatial information to the MRSA infection and did not find any strong correlation there. Probably with more detailed specific data it would be possible to build stronger spatial correlations.

Epidemiological Modelling and Public Health Decision Making

Alain-jacques Valleron, Université P. et M. Curie & Inserm, St Antoine Hospital, Paris, France

Looking at the evolution of epidemiology from 1950-1990, we can say that it was a time of success in "statistical" epidemiology because:

- Risk factors of major chronic diseases were discovered.
- Attributable parts were measured.
- Prevention was based on epidemiological discoveries.
- Evidence based medicine became the reference in health agencies and for medical practice.
- It was clear that epidemiology was the science for public health.
- It was mostly based on statistics.

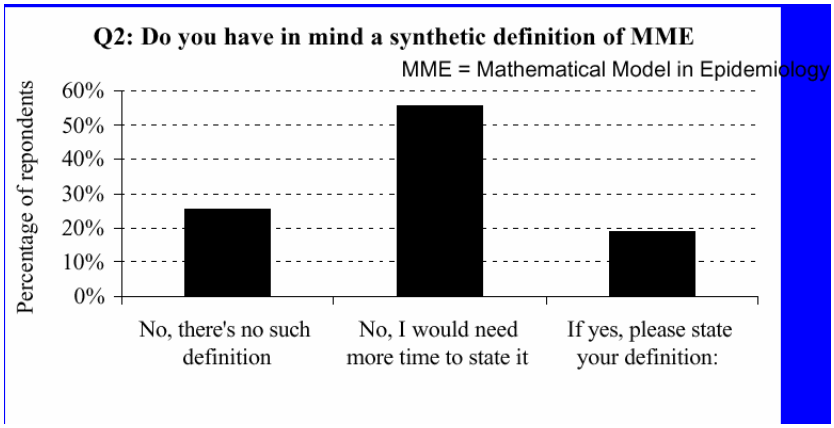
After 1990 we saw the rise of modelling and mathematics. There are several reasons for this:

- General development of informatics (new information systems, powerful computers (x100,000) and more mathematically trained scientists in life sciences).
- New public health issues that statistical epidemiology could not address:
  -The "low dose" problems (low risk x large population)
  -Emerging diseases such as HIV

-Bioterrorism after 9/11.

- Research done:
  -Assessment of the risks that cannot be observed with standard epidemiological methods.
  -Evaluation and/or discovery of new disease control strategies.
  -Development of innovative new methods and algorithms.

When surveying different factors of mathematical model in epidemiology, a very interesting pattern of responses was obtained in answer to the question "Do you have in mind a synthetic definition of mathematical model in epidemiology?":



Hejblum, INSERM U707, 2007

This revealed that it is not easy to determine sometimes what mathematical epidemiology is.

Some of the questions addressed in surveys were presented in random order. The respondents gave the following responses (ordered from 1 to 8, with 8 the lowest response) to: How important do you consider the following conditions to be if you were to say that an MME is "good"?

1. When its results can be replicated.
2. When it is considered to be good science by mathematicians specialised in the methods.
3. When it can be validated with actual data.
4. When it is grounded in actual data.
5. When a sensitivity analysis is presented.
6. When its results can be extrapolated.
7. When it is considered to be good science by the life science specialists in the domain of application.
8. When it is used by decision-makers.

When observing the prioritization of the responses, is interesting to observe that modellers do not give much importance to the use of a model by decision-makers.

To assess models' impact on decision-making in health care issues, we look at the example of the risk of radiation at a low dose. Models have been used for a long time to assess the hazards of low doses of radiation to human health, such as the impact of diagnostic radiology for cancer (see graph below).

Estimation of the risk of cancer associated with diagnostic X-rays (LNT model)

UK

	All	
	Males	Females
Cancer type (ICD-9 code)		
Oesophagus (150)	5	6
Stomach (151)	15	14
Colon (153)	34	73
Liver (155)	2	2
Lung (162)	21	40
Breast (174)	..	29
Bladder (188)	85	26
Thyroid (193)	1	3
Leukaemia (204–208, excluc	27	26
All cancers listed above†‡	190	219
All cancers‡	341	359
% of total	100%	100%
% of total—both sexes	100%	

International

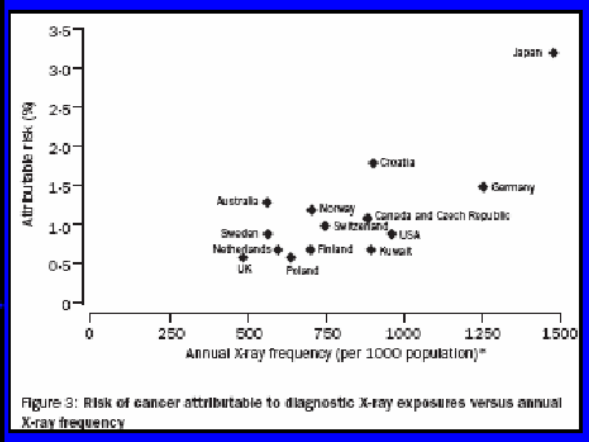


Figure 3: Risk of cancer attributable to diagnostic X-ray exposures versus annual X-ray frequency

Barrington et Darby , Lancet, 2004

The low dose controversy:

• Decision makers need cancer estimations to set regulatory limits of exposure.

• The No Threshold Linear Model (NTLM) is the simplest possible representation of the dose effect curve. But, it is impossible to validate.

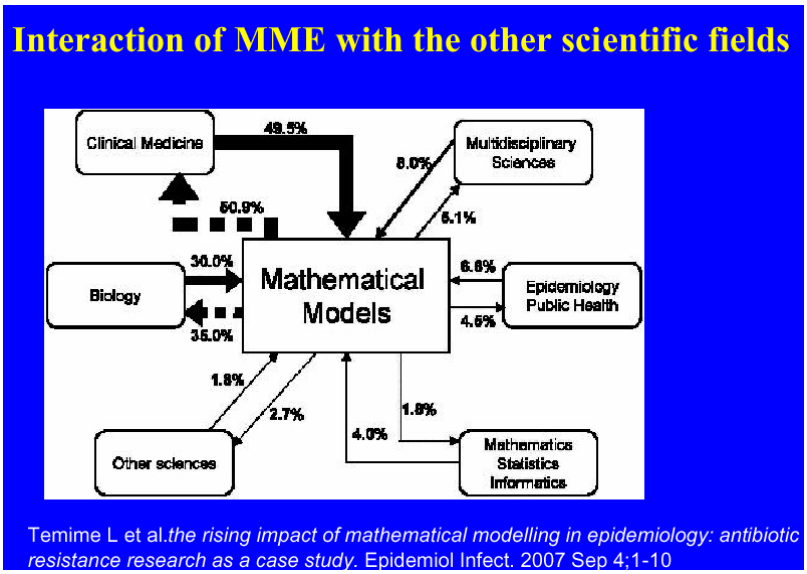
-Initially: a fit. No mechanistic explanation.

-Then biological support to the NTLM is found: the probability that X-rays induce DNA double strand breaks must be proportional to the dose.

-Now, rebuttal of the NTLM based on biology: at low doses cells may trigger defense mechanisms (cell death).

-If true, a threshold must exist.

It is also important to see that, especially in regard to infectious diseases, there are many scientific disciplines that interact, as shown in the following figure:



Epidemiology has faced several storms which have caused concern in this field. Publications and papers came out regarding these storms, such as the paper from Guy Taubes, "Epidemiology Faces its Limits", published in 1995, which highlighted the concerns and contradictions in epidemiology.

Statistical epidemiologists addressed these problems in a document called the "Consort Statement 2001". This document is a set of rules that epidemiologists must enforce to guarantee to decision makers that they can trust their work.

Peng & Dominici in 2006 established a set of criteria necessary for reproducible epidemiologic research:

TABLE 1. Criteria for reproducible epidemiologic research

Research component	Requirement
Data	Analytical data set is available.
Methods	Computer code underlying figures, tables, and other principal results is made available in a human-readable form. In addition, the software environment necessary to execute that code is available.
Documentation	Adequate documentation of the computer code, software environment, and analytical data set is available to enable others to repeat the analyses and to conduct other similar ones.
Distribution	Standard methods of distribution are used for others to access the software, data, and documentation.

In Peng & Dominici, AJE, 2006

Other emerging storms around mathematical epidemiology?

-Other biomedical scientists claim that the models neglect important new knowledge, and over-simplify.
-Public health decision makers find the models "too complicated".

- New models are increasingly complicated, and hard to replicate by independent researchers, even when data and codes are in place.
- No possible validation for most prospective models (e.g. bioterrorism).

To counteract the above it is necessary to:

- Define formally "good practices in modelling"
- Encourage benchmarking, and systematic comparison of models addressing similar issues.
- Encourage contests on selected problems and datasets (see the ISDS initiative about algorithms of outbreak detection).

Discussion

**If you had given this talk about 5 to 10 years ago, I would agree that modelling is really the cherry on the cake; however, today I have seen a big change in how modelling really influences decision-making (comment from Edmunds).*
**If you look closely at the WHO plan for containment of pandemics, they used modeling for containment at source, and the target of antiviral prophylaxis at source (Comment from Longini).*

VI. Closing Remarks

John Edmunds, Modelling and Economics Unit, Health Protection Agency, UK

This meeting has covered several themes in infectious disease modelling: diseases, methods, early warning systems, engagement with policy-makers and economic assessment. In terms of specific diseases, the meeting covered endemic diseases such as Varicella-Zoster virus (VZV), and epidemic diseases such as flu or Chikungunya. There were no presentations on sexually transmitted diseases, although there are quite a number of models for these. The talks compared these diseases in the developed and developing world.

In terms of methods, presenters addressed complexity vs. the simplicity of models, the range of models employed, and models for specific diseases such as flu or HIV. They also addressed the fact that today models can be more spatially explicit than ever before through the use of maps, and highlighted the need to fit data to models and the different techniques available for doing this.

The presentation on early warning systems and real time models addressed environmental vector related diseases such as Rift Valley Fever (RVF) or Chikungunya, close-contact infections where statistical techniques are developed to investigate clusters and estimate key parameters.

We learnt that systems need to be properly tested and evaluated, and that we must be aware of false positives, which are as dangerous as false negatives. Talks also addressed model uncertainty and interaction with policy-makers; they emphasized the need to minimize uncertainty but also the need for policy-makers to understand that a degree of uncertainty cannot be avoided in this field.

There was a clear understanding of the need for policy-makers to engage with public health, and that models need to be integrated with public health practitioners to a safe degree.
It was interesting to see the value of economic health care approaches for decision-making. The degree of uncertainty that could exist in these evaluations, especially with emerging diseases, represents a major challenge given the cost of making the wrong decision .

Some future challenges in epidemiological modelling are: HCAI even for epidemic diseases, statistical analysis and fitting models to data; real-time parameter estimation and model-based forecasts; structural/model uncertainty; models that integrate molecular information, pathogen & host parameters; behavioural change (adaptive behaviour); endemic diseases (update of vaccine); and epidemic diseases.

V. Annexes

Press Release December 10, 2007

Advances in Infection Disease Modelling
December 10-12, 2007

***Lyon, December 10, 2007*- Fondation Mérieux organizes a three day symposium in relation to infectious disease modelling, at "Les Pensieres" Conference Center in Veyrier du Lac, France. The symposium will unfold throughout its sessions the understanding of mathematical modeling and its increase applicability to predict the fate of infectious disease epidemiological strategies, and research design.**

The crafting of mathematical modelling to predict the impact of infectious disease prevention, surveillance and control programs; hence to anticipate the probable outcome in the implementation of pre-design action plans, has evolved over time. From the first model that appeared in 1760 to today's computerize models where scientific research and informatics fields work together.

Most variables that play a role in the fate of infectious disease epidemiology such as; the host, the pathogen, the target population, the transmission patterns, the eco-social enviroment, just to mention few, are considered, analysed and tailored through mathematical predictions.

Infectious disease epidemiology has intrinsic aspects that are not applicable to all diseases, thus in many occasions conventional epidemiological dynamics do not always address the needs of infectious diseases.
The developing of specific modelling methods and measurements to address this type of disease patterns have provided an outstanding and powerful tool, to evaluate and interpret data for critical decision-making and program customization to access infectious diseases.

Various public & private institutions and organizations in relation to infectious disease epidemiology are and have taking advantage of these models for public health strategy-making, and for the optimization in the use of resources; among other applications.

The event will count with the participation of foremost international experts on the subject, their work and findings will be unveiled in four main sessions.

- **What is a model?**
- **What is the expected public health impact of the model approach?**
- **Predicting the impact of interventions.**
- **The future of infection diseases modeling.**

The format of the symposium is intended to generate discussion among participants and to foster the dissemination of new information on this topic. The symposium will provide an opportunity for specialists in infectious disease modelling to exchange knowledge through the sharing of related research studies, innovations and applied methodologies.
The symposium, consistent with the foundation's core mission, contributes to the dissemination of scientific information worldwide and to the epidemiological surveillance of infectious diseases.

About Fondation Mérieux
Fondation Mérieux was created in 1967 by Doctor Charles Mérieux and was granted charity status in 1976. Presided by Alain Mérieux, the Foundation's mission is to fight infectious diseases affecting developing countries. The Foundation works to develop and make available new and affordable approaches based on biotechnologies, in the field of prevention, diagnostics and therapeutics.
To achieve its goal, Fondation Mérieux plays a catalyst role in Research and Development by mobilizing a network of excellence that gathers the foremost international experts working in the scientific world today. The Foundation fosters the dissemination of scientific information and innovation through international seminars and conferences, like the Santiago symposium. The Foundation also provides high-level, practical scientific training for health practitioners in the developing world. Finally, Fondation Mérieux works directly in the field by strengthening and building local health infrastructures to enable long-term sustainable development. It is present in Africa, Asia and in Haïti.

Advances in infectious diseases modelling

Annecy, Les Pensières, December 10-12, 2007

Infectious disease modelling has a long history. The first model was developed for smallpox by Bernouli in 1760. Infectious disease models have first been used to understand the temporal and spatial dynamics of an epidemic and then to estimate treatment or control strategy.

Currently, infectious disease models have been more and more used to predict a variety of different futures, to help and support the knowledge development and the decision process at the scientific, medical and public health level. To achieve these objectives, new methodologies have been developed or adapted from other fields and studies have been performed for model validation, for different infectious diseases or focusing on vaccines.

The aim of this conference is to give an overview of the different questions that modelling approach can resolve, using recent applications examples in different infectious diseases

Monday, December 10, 2007

17h30-18h30	***Registration***	
18h30-18h45	Welcome Address	**B. Miribel**
18h45-19h15	Keynote lecture :	**K.J. Linthicum**
19h45	***Welcome Dinner***	

Tuesday, December 11, 2007

Session I: What is a model?

Models: What are they, what can they do, how do we choose which to use, what data are needed and how do we validate them?

Chairperson : **Odo Diekmann**

08h30-08h50	What is a model and why use one?	**R. Anderson**
08h50-09h05	***Discussion***	
09h05-09h25	Applications of models: roles and approaches	**N. Ferguson**
09h25-09h40	***Discussion***	

09h40-10h00	Simulations: what level of complexity is appropriate?	**S. Eubank**
10h00-10h15	***Discussion***	
10h15-10h45	***Coffee break***	
10h45-11h05	Model parameterisation and validation: Methods and data needs	**A. Ghani**
11h05-11h20	***Discussion***	
11h20-11h40	Existence of a dominant network: From global pandemics to small-scale disease spread	**M. Barthélémy**
11H40-11h55	***Discussion***	
11h55-14h00	***Lunch***	
14h00-14h20	From model to public health decision: Chikungunya story	**A. Flahault**
14h20-14h35	***Discussion***	

Session II: - What is the expected public health impact of the model approach?

Contribution of modelling to the evaluation of possible vaccination strategies

Chairperson : **Daniel Barth-Jones**

14h35-14h55	Impact of a wide vaccination strategy against meningo infection in Africa	**M.P Preziozi**
14h55-15h10	***Discussion***	
15h10-15h40	Impact of combined effect of vaccine and decrease antibiotic use on S. pneumoniae susceptibility to antibiotic	**D.Guillemot**
15h40-16h00	***Coffee break***	
16h00-16h20	HPV vaccination using dynamic and static models	**J. Kim**
6h20-16h40	***Discussion***	
16h40-17h00	Health economic evaluation of vaccine: the example of varicella-Zooster virus	**B. Dervaux**
17h00-17h20	***Discussion***	
17h20-17h40	Evaluation of different vaccination strategies against pertussis in adults	**A. Van Rie**
17h40-18.00	***Discussion***	
18.00	End of the session	
19h00	***Dinner***	

Wednesday, December 12, 2007

Session III: - Predicting impact of interventions

Modelling enables to scientists to estimate interventions impact (i) at the population level with limited resources, comparatively to epidemiological studies;(ii) for the long term;and (iii)in complex interaction system.

Chairperson : **P. Beutels**

08h30-08h50	Strategies for detecting and containing an emerging H5N1 pandemic	**I. Longini**
08h50-09h05	***Discussion***	
09h05-09h25	Long-term impact of potential interventions: Malaria	**M. Eichner**
09h25-09h40	***Discussion***	
09h40-10h00	Modelling HIV vaccines	**D. Bath-Jones**
10h00-10h15	***Discussion***	
10h15-10h45	***Coffee break***	
10h45-11h05	Mixing patterns and the spread of infectious diseases: the results of a large multi-country study	**J. Edmunds**
11h05-11h20	***Discussion***	
11h20-11h40-	Modelling options for economic analysis: realism versus pragmatism and fiction	**P. Beutels**
11h40-12h05	***Discussion***	
12h05-14h00	***Lunch***	

Session IV: - The future of infection diseases modelling

Chairperson : **Martin Eichner, Ira Longini**

14h00-14h20	Public health authorities point of view: needs and requirements	**R. Hutubessy**

14h20-14h35	***Discussion***	
14h35-14h55	Modelling nosocomial diseases	**B. Cooper**
14h55-15h10	***Discussion***	
15h10-15h30	Modelling zoonotic infections and cross-species transfer	
15h30-15h45	***Discussion***	
15h35-16h15	***Coffee break***	
16h15-16h35	Development of the modelling and consequences on the research network organisation	**AJ Valleron**
16h35 16h50	***Discussion***	
16h50	***Closing remarks***	**J. Edmunds**
17h30	End of the meeting	