PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

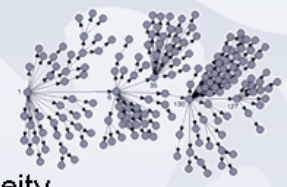
Aspects of homogeneous vs. heterogeneous transmission

Ping Yan, Research Manager

Surveillance and Risk Assessment Division
Centre for Communicable Diseases and Infection Control

Infectious Diseases Prevention and Control Branch

Highlight



1. Phenomenological observation

"A picture is worth 1000 words."

1.1. Implications of phenomenological heterogeneity

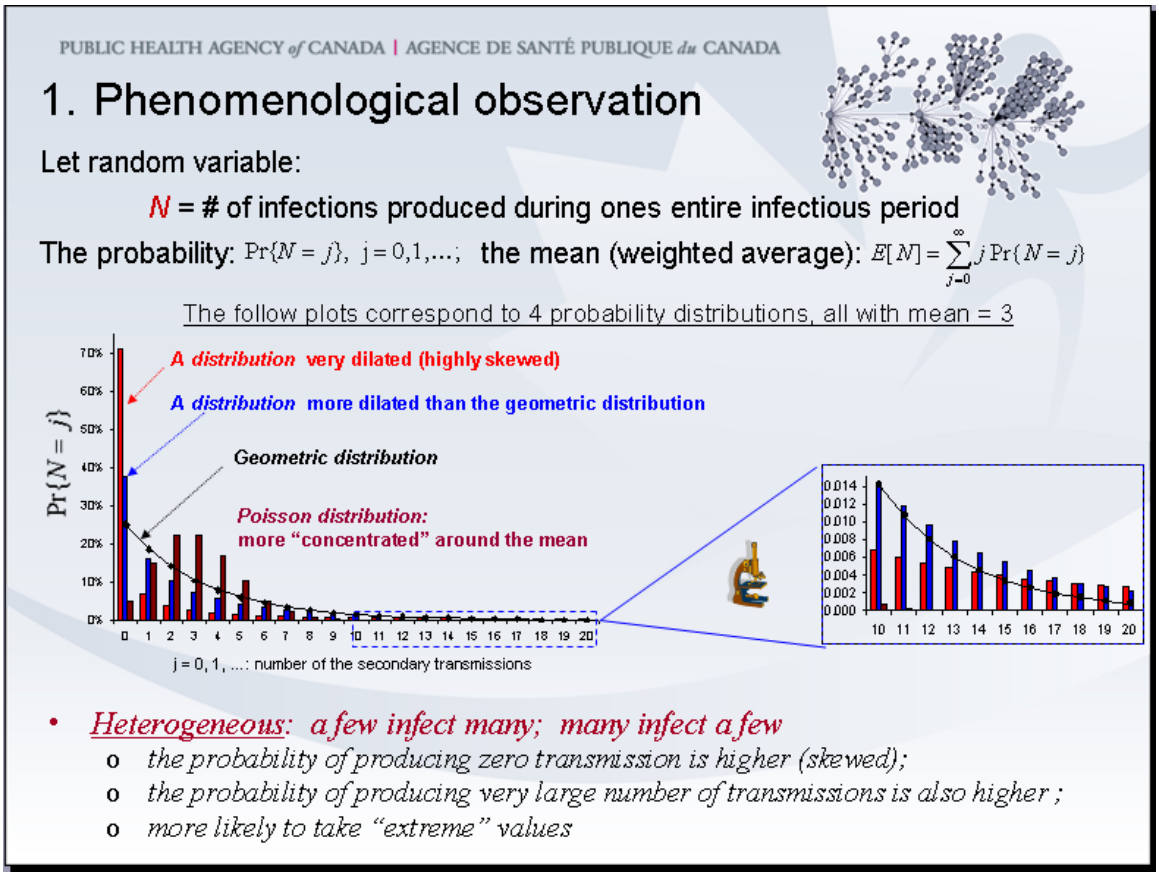
At *the very beginning:* as an infectious individual ***seeded***
in an infinitely large susceptible population … …

2. Heterogeneity in the agent-host-environment interface

- Conceptual assumptions vs. tactical assumptions

2.1. Heterogeneity in tactical assumptions, implications on dynamics

2.2. Heterogeneity in conceptual assumptions, implications on dynamics



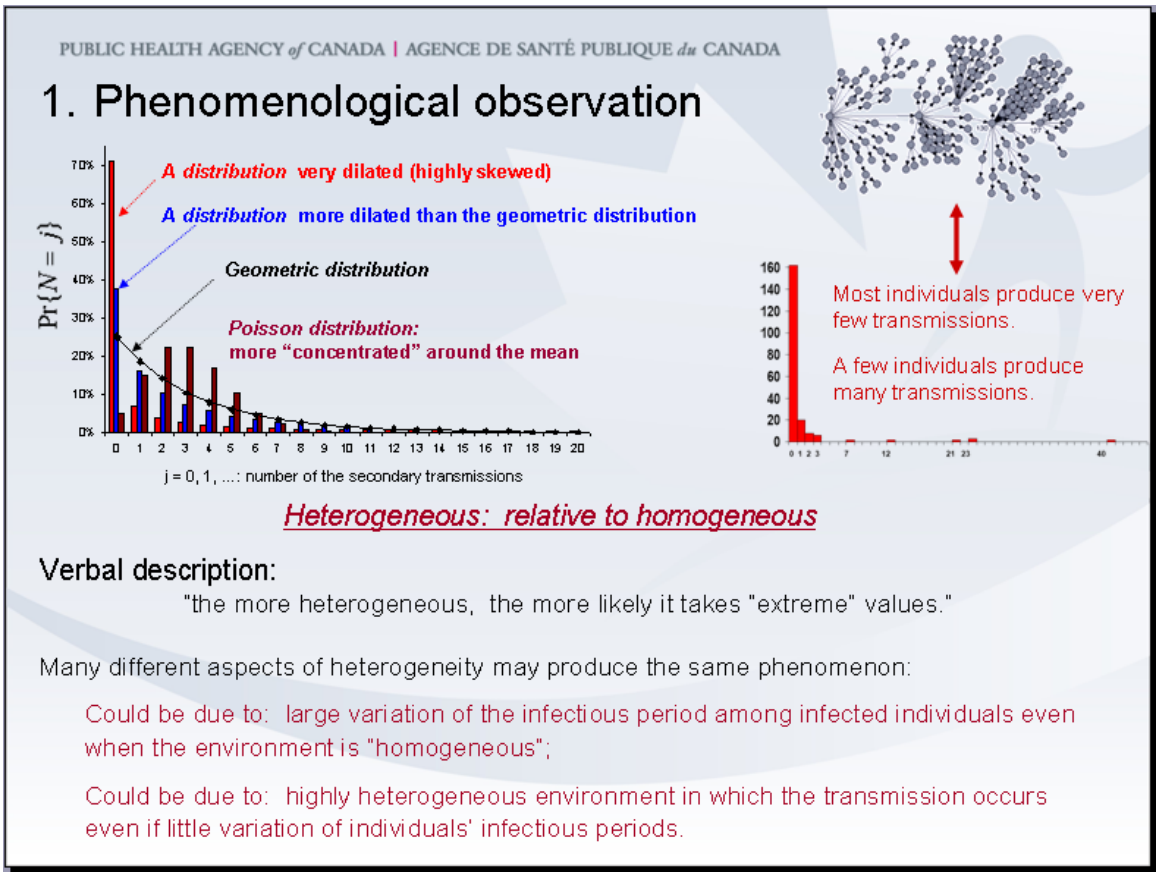
Heterogeneity is a relative term to homogeneity.

The ODE models for SIR and SEIR and the corresponding stochastic compartment models with Markov property generate the geometrically distributed infectious contacts, and those models are often attached with the word "homogeneity".

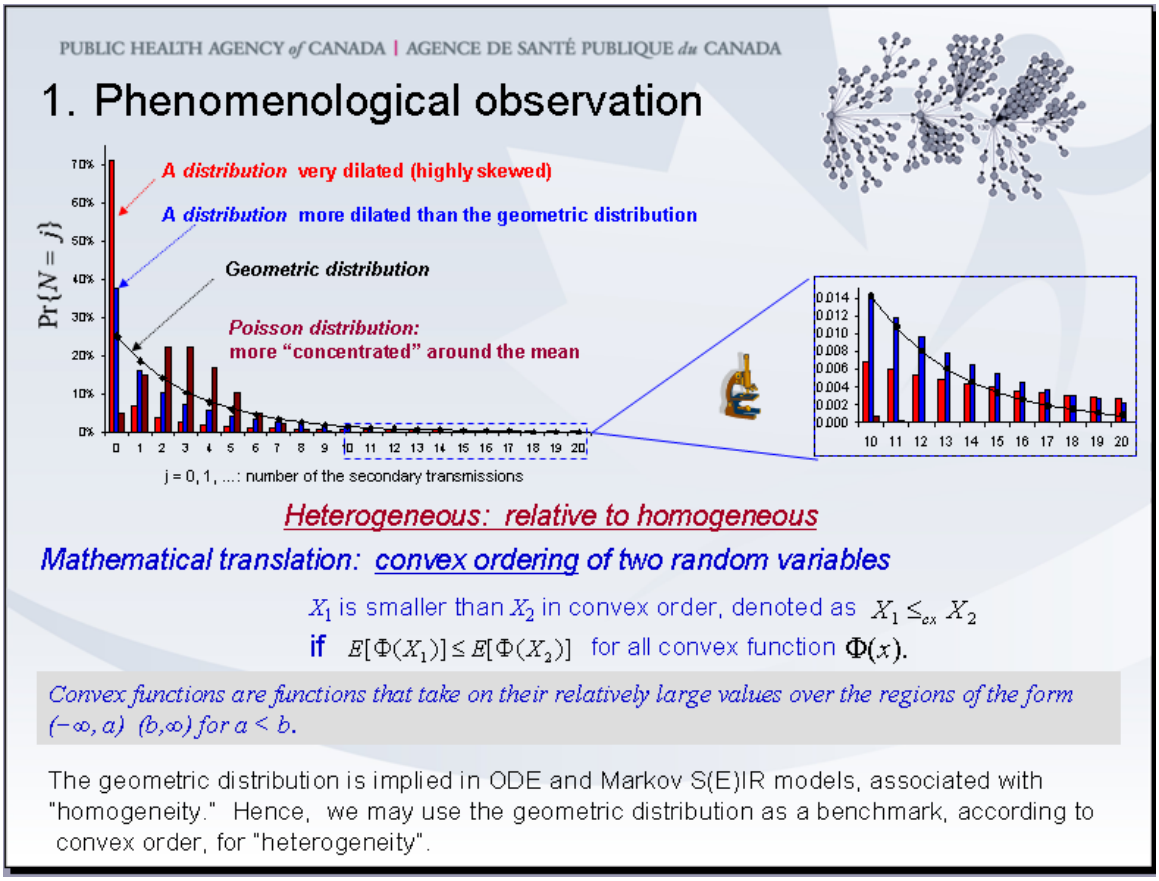
Therefore, we use the geometric distribution as a benchmark. In the plots, the geometric distribution is illustrated as a line. All the distributions have the same mean value.

If more heterogeneous than the geometric distribution, the probability of N=0 is higher and the probability of N taking large values is also higher (see the magnified image). Hence, the more likely it takes extreme values.

On the other hand, when more homogeneous than the geometric distribution, the distribution is more concentrated around the mean. The Poisson distribution is one of such examples.



The two “could be due to” aspects will be Part 2 of this presentation.

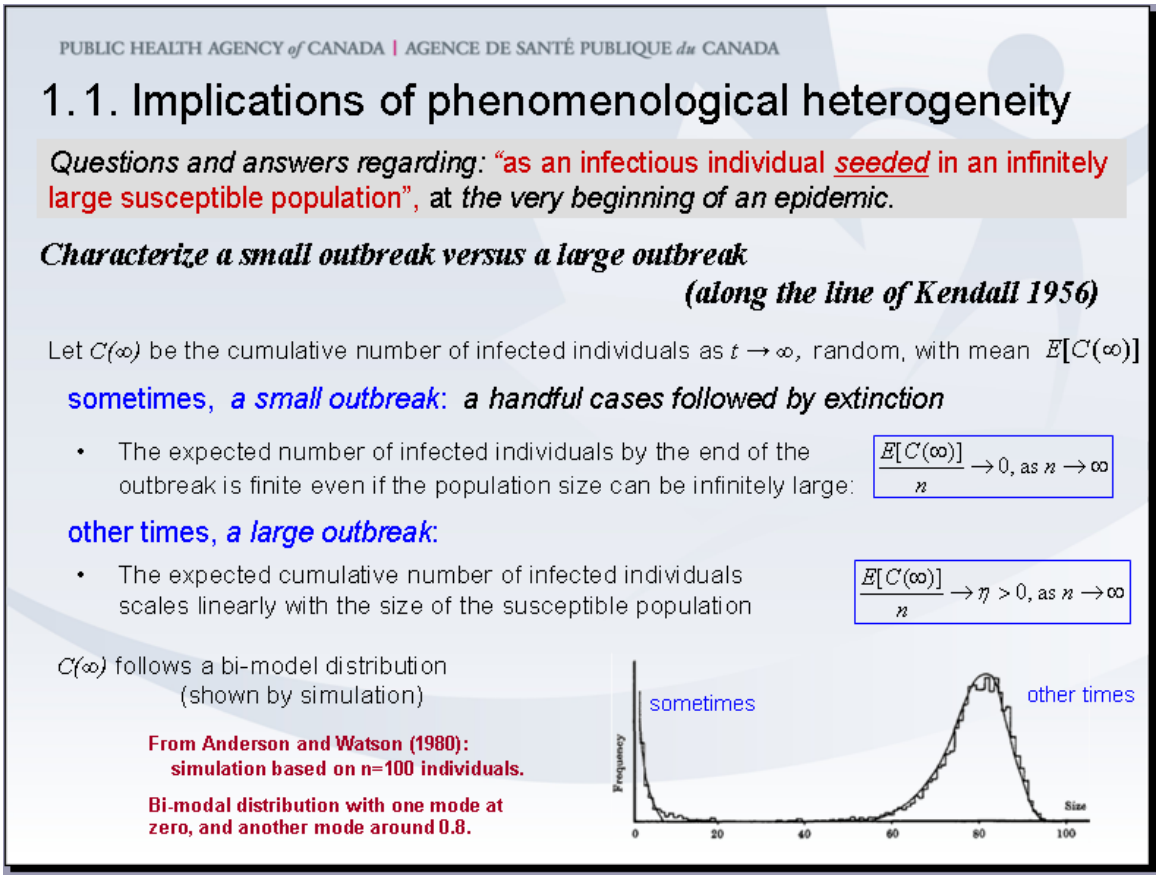


For more on convex order, I recommend

Shaked, M and Shanthikumar, J.G (2007) Stochastic Orders. Springer.

It is the mathematical language for the verbal description:

"The more heterogeneous, the more likely it takes "extreme" values."



The original Kendall paper is hard to find. It is

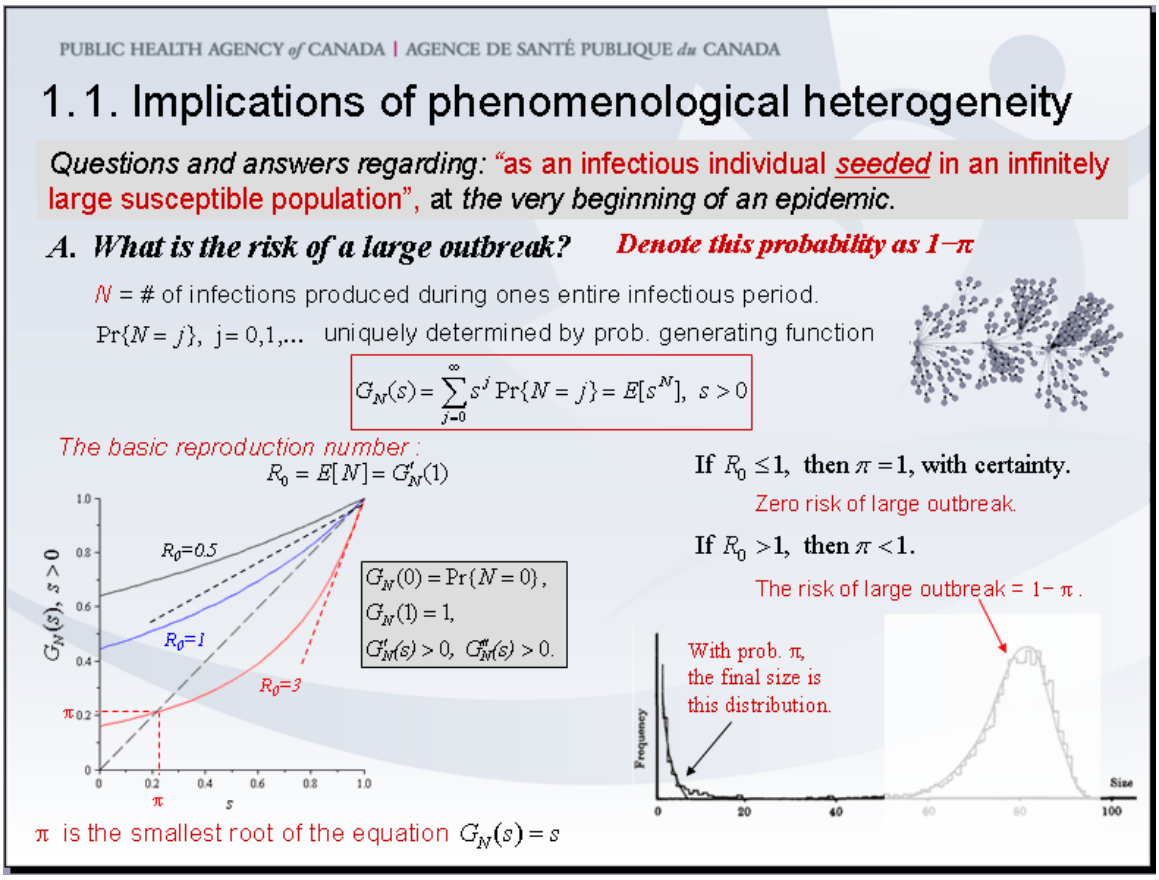
Kendall, D. (1956) Deterministic and stochastic epidemics in closed populations. Proc. Fifth Berkeley Symp. Math. Statist. Probab. 4. University of California Press 149-165.

Alternatively, in Diekmann and Heesterbeek (2000), the authors provide the same description.

Diekmann, O. and Heesterbeek, J.A.P. (2000). Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation. Wiley Series in Mathematical and Computational Biology.

The simulation example is taken from

Anderson, D. and Watson, R. (1980) On the spread of a disease with gamma distributed latent and infectious periods. Biometrika, 67, 1, 191-198.



The probability generating function (p.g.f.) is a very powerful tool to study non-negative integer values discrete random variable. The probabilities can be uniquely defined through the p.g.f.

The use of p.g.f. to study the extinction probability (i.e. the risk a small outbreak) can be found in every textbook on branching processes.

Relating it explicitly to the risk of a small outbreak, as well as the expression of R0 as the slope of the p.g.f. evaluated at s=1, along with all the strictly convex property of the p.g.f., can be found in Diekmann and Heesterbeek (2000).

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

1.1. Implications of phenomenological heterogeneity

B. What does heterogeneity do to the risk of a large outbreak $1-\pi$?

Convex ordering: $X_1 \leq_{cx} X_2$ if $E[\Phi(X_1)] \leq E[\Phi(X_2)]$ for all convex function $\Phi(x)$.

Variance by definition: $\text{var}[X] = E[(X-\mu)^2]$, where $\mu = \text{E[X]}$

$(x-\mu)^2$ is a convex function of x

Prob. generating function: $G_X(s) = E[s^X]$, $s \in [0,1]$

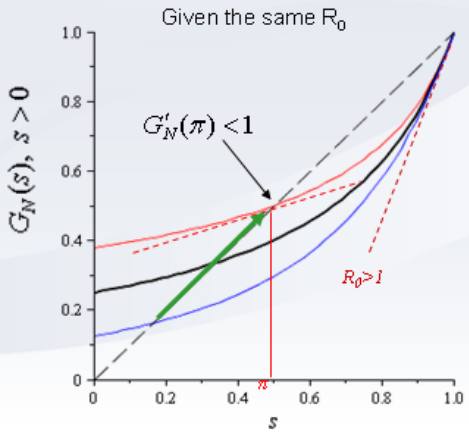
s^x is a convex function of x

Heterogeneity with respect to the random variable N, in the sense that,

"the more heterogeneous, the more likely it takes 'extreme' values"

then the more heterogeneous:

- the larger is the variance var[N];
- the larger is the value $G_N(s)$ for all $s \in [0,1]$
- the larger is the probability, π, of a small outbreak, as the smallest root of $G_N(s) = s$

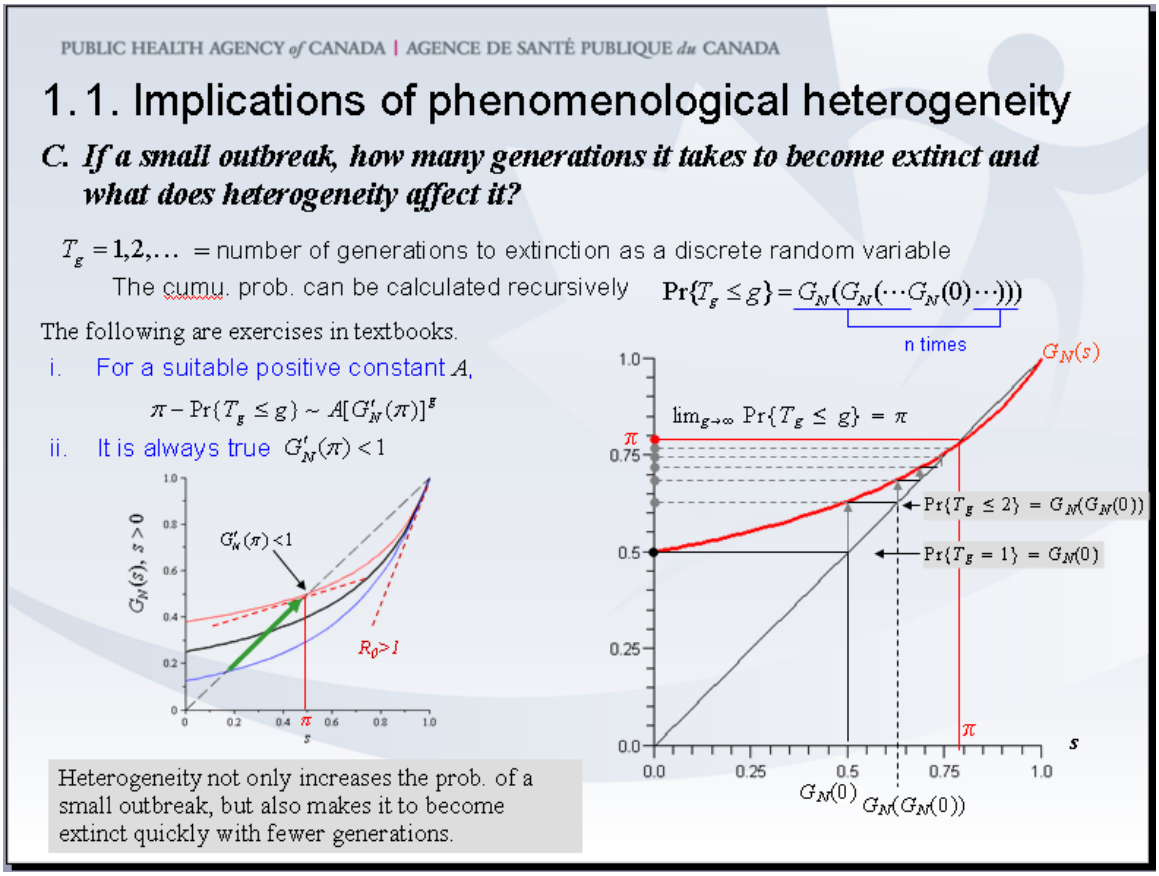


Heterogeneity is described by convex ordering.

Larger in convex order gives larger variance, which is a more intuitive measure.

Larger in convex order gives ordering of the probability generating functions.

Later in the presentation, we shall see a correspondence between the p.g.f. and the Laplace transform function of the infectious period.



This slide is based on Ch. 10 (Yan, P.) of Springer Lecture Notes (Ed. Brauer, van den Driessche and Wu).

Note that if $R_0 > 1$, $\lim_{g\to\infty} \Pr\{T_g \le g\} < 1$, which defines a sub-distribution, rather than a proper distribution function.

The graph showing $\lim_{g\to\infty} \Pr\{T_g \le g\} = \pi$ is the line of thinking that leads to the proof that π is the smallest root of the fixed-point-equation $G_N(s) = s$.

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

1.1. Implications of phenomenological heterogeneity

D. If a small outbreak, what is the distribution of its final size?

$C(\infty) = 1, 2, \ldots$ is a discrete random variable.

The distribution of $C(\infty) = 1, 2, \ldots$ **can be also precisely calculated** using a recursive formula based on $G_N(s),\ s > 0$

$$G_{C(\infty)}(s) = sG_N(G_{C(\infty)}(s))$$

where $G_{C(\infty)}(s)$ stands for the prob. generating function of $C(\infty)$, which can be used to calculate

$$\Pr\{C(\infty) = j\}, \mathrm{j} = 1,2,\ldots$$

The mean and variance of the final size:

Evaluate $G'_N(\pi)$ and $G''_N(\pi)$

$$E[C(\infty) \mid \text{small outbreak}] = \frac{1}{1 - G'_N(\pi)}$$

$$\text{var}[C(\infty) \mid \text{small outbreak}] = \frac{G''_N(\pi) + G'_N(\pi) - G'_N(\pi)^2}{(1 - G'_N(\pi))^3}$$

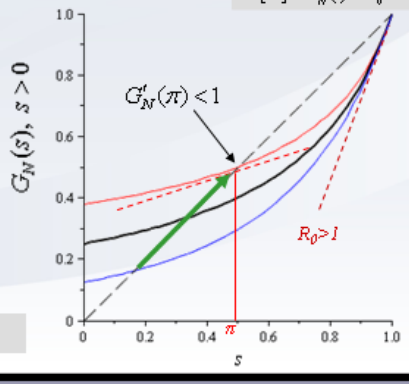
If $R_0 \le 1,\ \pi = 1, G'_N(1) = R_0$:

$$E[C(\infty)] = \frac{1}{1 - R_0},\quad \text{var}[C(\infty)] = \frac{\text{var}[N]}{(1 - R_0)^3}$$

Heterogeneity → large var[*N*], hard-to-predict final size

Given the same R_0:

$\text{var}[N] = G''_N(1) + R_0 - R_0^2$

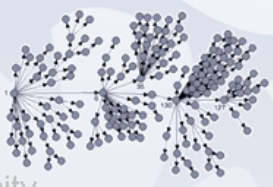


This slide is also based on Ch. 10 (Yan, P.) of Springer Lecture Notes (Ed. Brauer, van den Driessche and Wu).

The case R0≤1 leads to large variance.

In the case R0>1, $G'_N(\pi)$ replaces R0, and $G''_N(\pi)$ replaces $G''_N(1)$. Hence $G''_N(\pi) + G'_N(\pi) - G'_N(\pi)^2$ replaces var[N]. Some theory that leads to the statement of large variance of the final size needs to be fixed.

Highlight



1. Phenomenological observation

   "A picture is worth 1000 words."

   1.1. Implications of phenomenological heterogeneity

   At *the very beginning:* as an infectious individual *seeded* in an infinitely large susceptible population … …

2. Heterogeneity in the agent-host-environment interface

   - Conceptual assumptions vs. tactical assumptions

   2.1. Heterogeneity in tactical assumptions, implications on dynamics

   2.2. Heterogeneity in conceptual assumptions, implications on dynamics

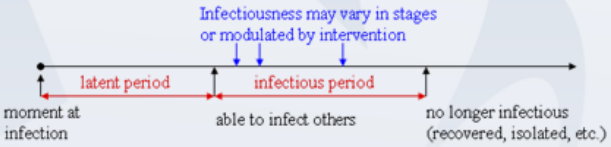
PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

2. Heterogeneity in agent-host-environment interface

- Tactical assumptions vs conceptual assumptions

Example of tactical assumptions

within an infected individuals



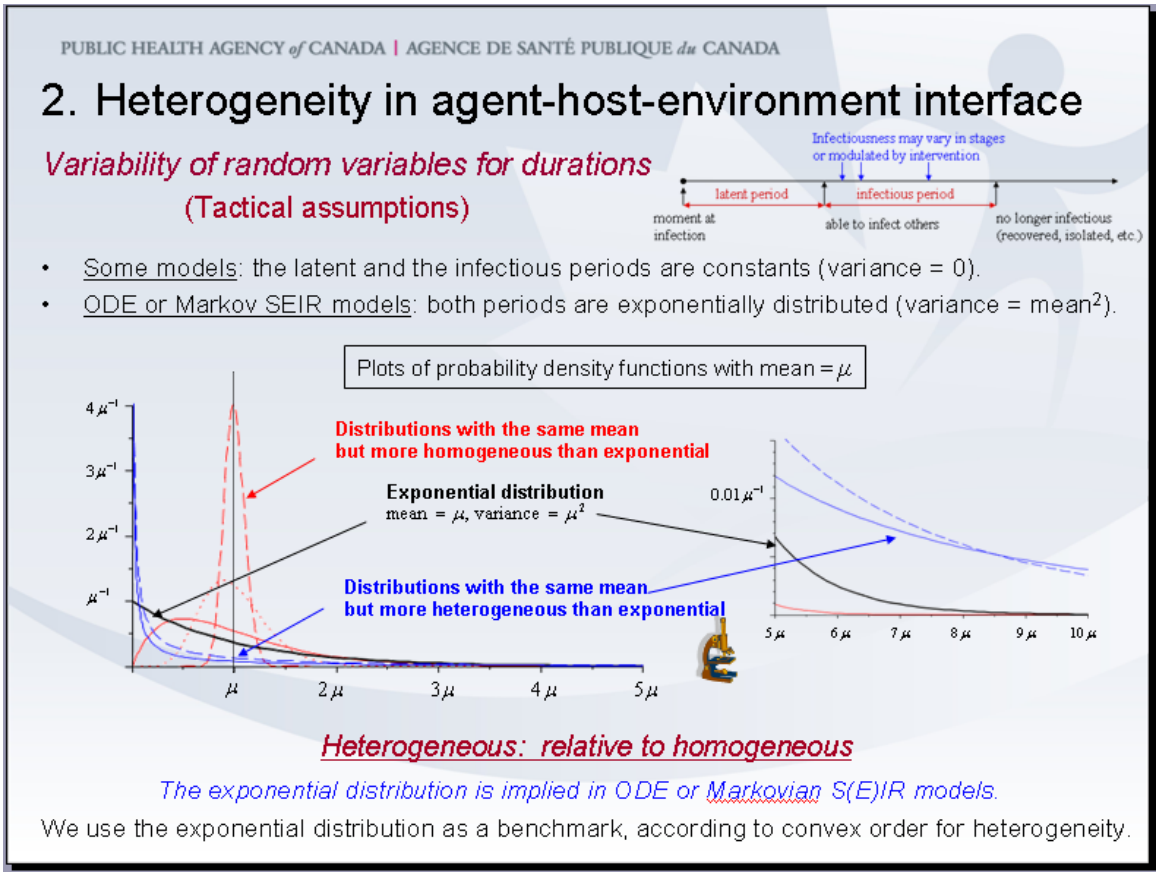
1. *Is there a latent period? If yes, how long on average? How variable?*
2. *Besides the average infectious period* μ_I *, how variable is it distributed?*

Example of conceptual assumptions: *environment, population, among individuals,… .*

1. The force of infection onto *a susceptible individual* $\propto \frac{I(t)}{n(t)}$: % infectious individuals
   $\rightarrow \beta \frac{I(t)}{n(t)} S(t).$
2. The instantaneous rate of passing the infection from *a typical infectious individual* to another $\propto \frac{S(t)}{n(t)}$: % of susceptible individuals $\rightarrow \beta^* \frac{S(t)}{n(t)} I(t).$

One of the (hidden) conceptual assumptions about homogeneity: $\beta^* = \beta$ (bilinearity)

The tactical assumptions are made with an infected individuals, otherwise known as the "natural history". The conceptual assumptions are made about how individuals interact.

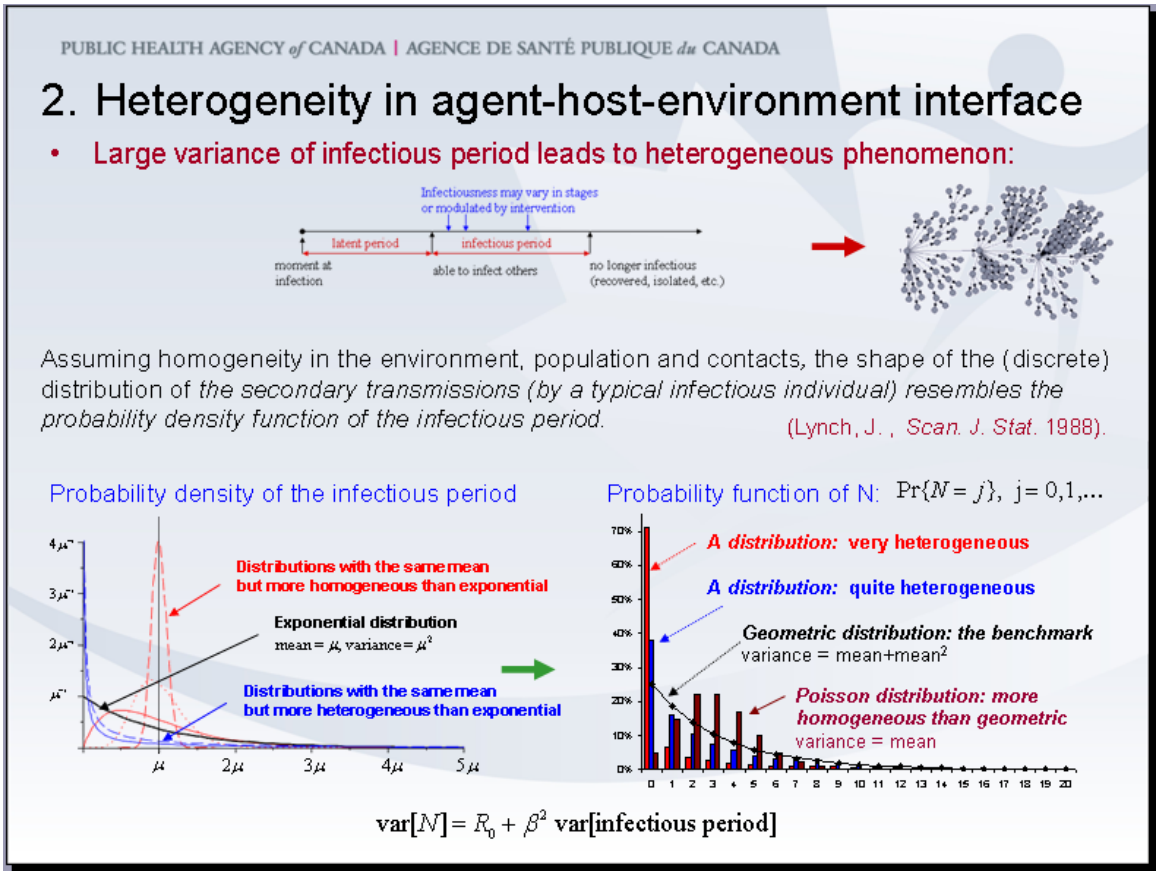


These are continuous time distributions, used to describe time, which is a non-negative continuous random variable. For continuous distributions, the probabilities are described by probability density functions.

In Part 1, we have seen plots of probabilities, used to describe the distribution of a discrete random variable N.

The exponential distribution presented here share many similar properties of the geometric distribution for discrete random variables.

We shall see more similarities in the next slide.



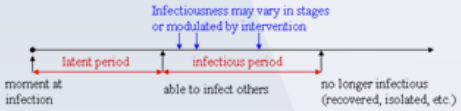
The shapes of the probability density functions of the infectious period (as a continuous random variable) correspond to similar shapes of the probability mass functions for the discrete random number N (such that R0 = E[N]). There is a probability theory for this, but beyond the scope of this presentation. For reference,

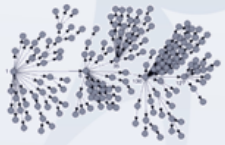
Lynch, J. (1988) Mixtures, generalized convexity and balayages. Scandinavian Journal of Statistics. 15, 203-210.

The geometric distribution for N corresponds to the exponential distribution for the infectious period.

2. Heterogeneity in agent-host-environment interface

- Large variance of infectious period leads to heterogeneous phenomenon:



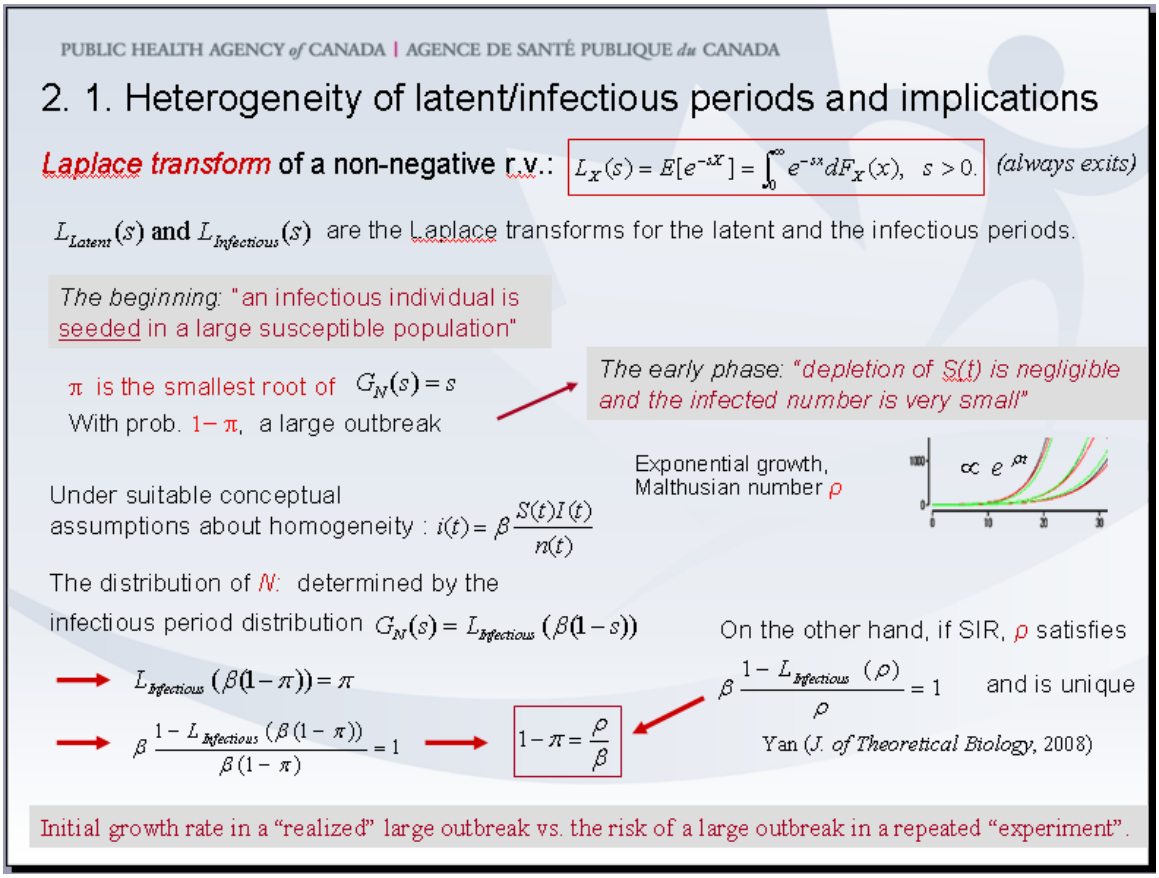


- But if the infectious period is constant, environmental factors (e.g. social network) can produce infectious contacts that lead to exactly the same phenomenon (same distribution).

Regarding small outbreaks when an infectious individual is seeded at the beginning, it is the phenomenon itself that leads to: ***the more heterogeneous***,

- the smaller is the risk of a large outbreak;
- the smaller is the mean final size of a small outbreak with larger variance;
- the smaller is generation-to-extinction in stochastic order for a small outbreak.

Regarding large outbreaks, their dynamics over time (e.g. growth, peak, duration), their final outcomes, and effectiveness of control measures to mitigate the outcomes, different aspects of heterogeneity at tactical level and conceptual level, have different impacts.



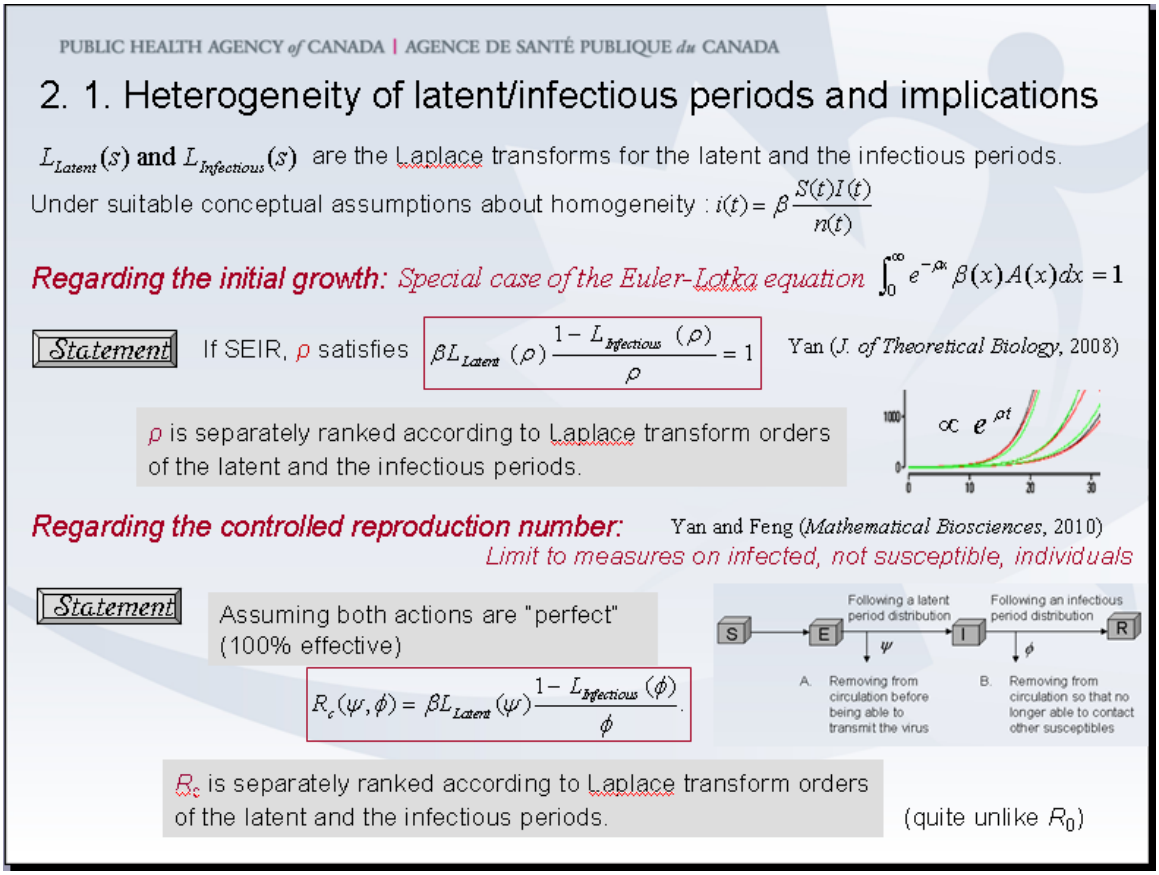
The correspondence between the p.g.f. and the Laplace transform function of the infectious period is given here. This correspondence is only true under "suitable conceptual assumptions of homogeneity" which will be discussed in detail later. This correspondence implies that the distribution of N, the risk of a large outbreak, as well as R0 , do not depend on whether there is a latent period.

On the right side of the slide, the formula is a special case of a more general formula that depends on the latent period as well. It is found in

Yan, P. (2008) Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks. Journal of Theoretical Biology 251, 238-252.

The key result in the slide only works if the underlying model is SIR. The implication is that one can use the observed initial growth rate in a large outbreak to assess the risk of a large outbreak in a similar community, under similar initial conditions (regarding each outbreak as a "random experiment" by nature).

It also implies to "patch models" with an "infected patch" defined as the one with observed large outbreak. Movement of individuals to a susceptible patch do not necessarily result in another infected patch, but with a probability of it. This probability can be modelled, if both "beta" and "rho" are estimated.

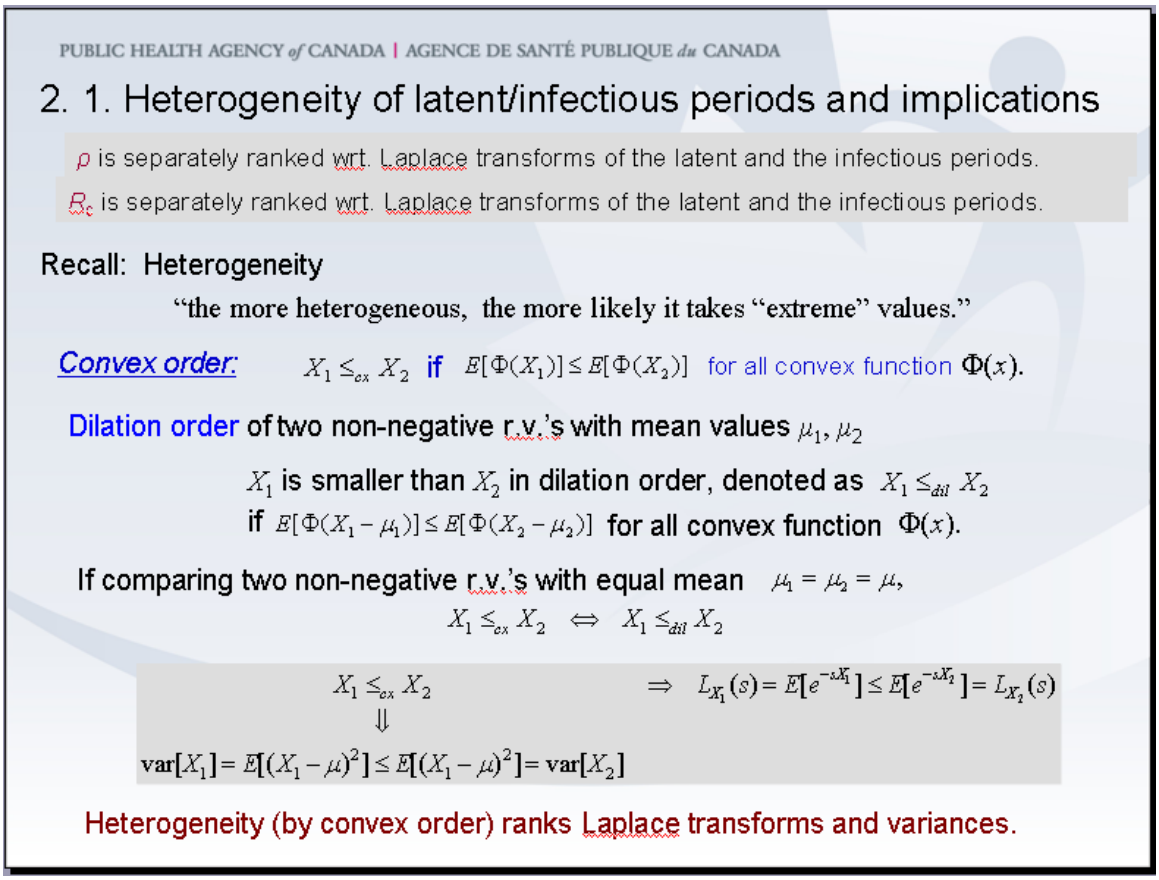


The general formula on the initial growth is from

Yan, P. (2008) Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks. Journal of Theoretical Biology 251, 238-252.

The formula on the controlled reproduction number is from

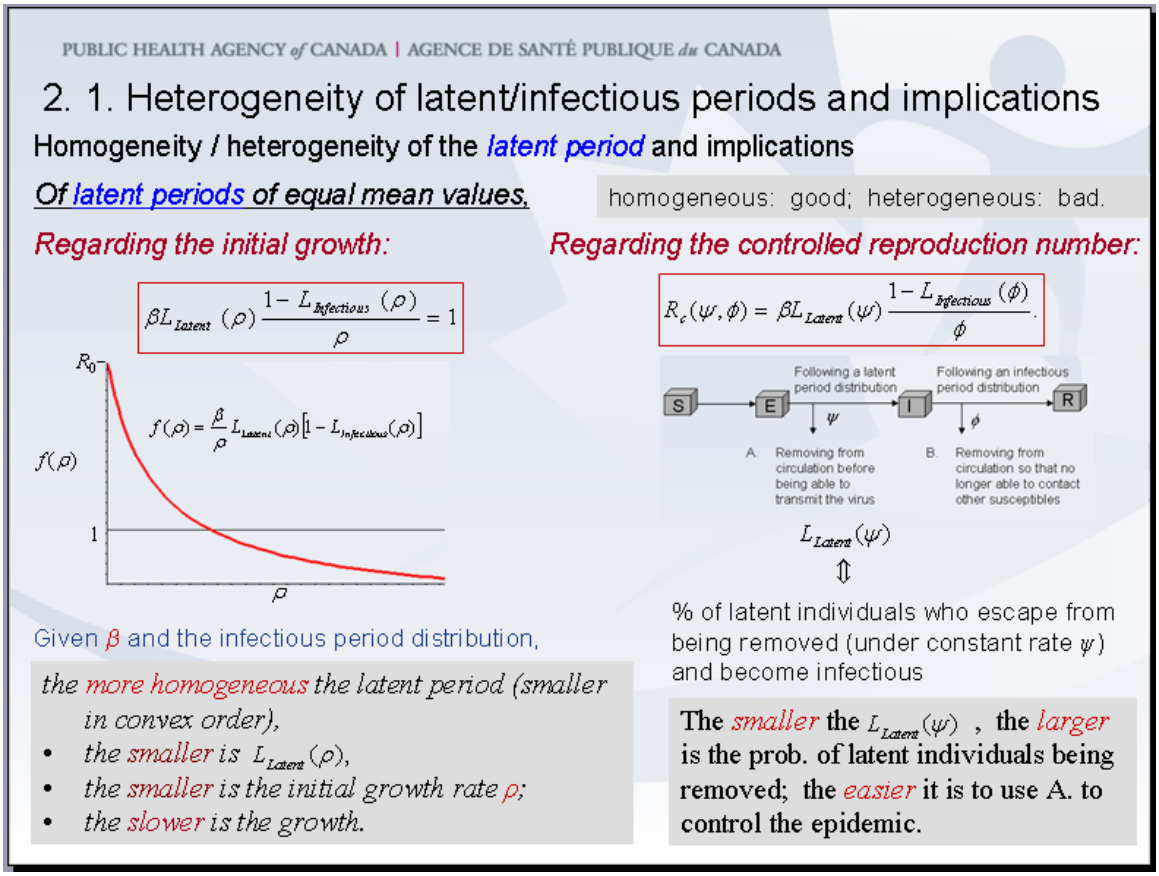
Yan, P. and Feng, Z. (2010) Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness. Mathematical Biosciences 224, 43-52.



There are two excellent textbooks about different kinds of variability orders. The results displayed here are only a small part of a broader theory, that are directly related to the current subject.

Marshall, A.W. and Olkin, I. (2007) Life Distributions, Structure of Nonparametric, Semiparametric and Parametric Families. Springer.

Shaked, M and Shanthikumar, J.G (2007) Stochastic Orders. Springer.



This is a synthesis among the two papers:

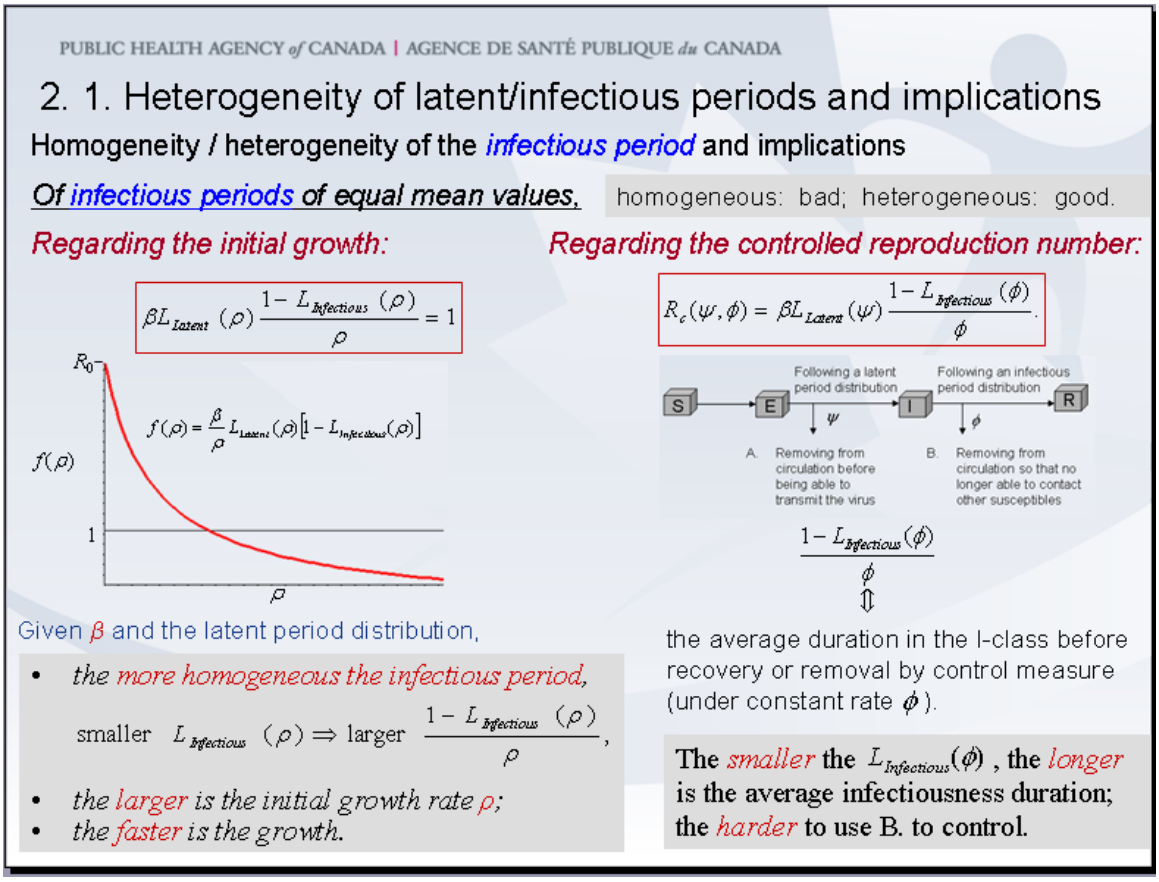
Yan, P. (2008) Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks. Journal of Theoretical Biology 251, 238-252.

Yan, P. and Feng, Z. (2010) Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness. Mathematical Biosciences 224, 43-52.

and theories from

Marshall, A.W. and Olkin, I. (2007) Life Distributions, Structure of Nonparametric, Semiparametric and Parametric Families. Springer.

Shaked, M and Shanthikumar, J.G (2007) Stochastic Orders. Springer.



Ditto: This is a synthesis among the two papers:

Yan, P. (2008) Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks. Journal of Theoretical Biology 251, 238-252.

Yan, P. and Feng, Z. (2010) Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness. Mathematical Biosciences 224, 43-52.

and theories from

Marshall, A.W. and Olkin, I. (2007) Life Distributions, Structure of Nonparametric, Semiparametric and Parametric Families. Springer.

Shaked, M and Shanthikumar, J.G (2007) Stochastic Orders. Springer.

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

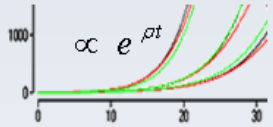
2. 1. Heterogeneity of latent/infectious periods and implications

Some further implications

Under suitable conceptual assumptions about homogeneity so that $i(t) = \beta \frac{S(t)I(t)}{n(t)}$

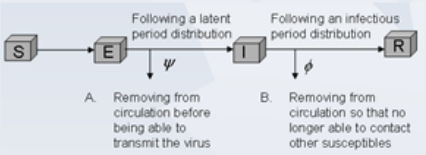
Regarding the initial growth:

$$\beta L_{Latent}(\rho) \frac{1 - L_{Infectious}(\rho)}{\rho} = 1$$



Regarding the controlled reproduction number:

$$R_c(\psi, \phi) = \beta L_{Latent}(\psi) \frac{1 - L_{Infectious}(\phi)}{\phi}.$$



If one can set control objectives $\psi \geq \psi_c$ and $\phi \geq \phi_c$ in order to achieve $R_c(\psi, \phi) \leq 1$,
then one can successfully prevent a large outbreak from taking place.

Ideally, it is achievable if $\psi_c = \phi_c = \rho$, assuming both actions are "perfect" (100% effective).

Lessons can be learned from observed initial growth rate in large outbreak that have happened elsewhere to set control targets to prevent a large outbreak from happening in a similar community, under similar initial conditions.

This is a synthesis between the two papers:

Yan, P. (2008) Separate roles of the latent and infectious periods in shaping the relation between the basic reproduction number and the intrinsic growth rate of infectious disease outbreaks. Journal of Theoretical Biology 251, 238-252.

Yan, P. and Feng, Z. (2010) Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness. Mathematical Biosciences 224, 43-52.

Clearly, if "psi" and "phi" both take the value "rho", then *R*c = 1.

In other words, control measures are race against time.

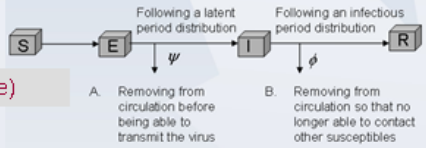
PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

2. 1. Heterogeneity of latent/infectious periods and implications

Extension for: **the controlled reproduction number**

Under suitable conceptual assumptions about homogeneity : $i(t) = \beta \frac{S(t)I(t)}{n(t)}$

Before: $R_c(\psi, \phi) = \beta L_{Latent}(\psi) \frac{1 - L_{Infectious}(\phi)}{\phi}.$



assuming both actions are "perfect" (100% effective)

Extension: "leaky situations"

1. For action B, infected individuals may be put into "leaky isolation", with reduced transmissibility $(1-\sigma_I)\beta$.
2. For action A, latent individuals may be put into "leaky isolation", and when they become infectious, they have reduced transmissibility $(1-\sigma_L)\beta$.

Quantitative: $R_c(\psi, \phi | \sigma_L, \sigma_I) = (1-\sigma_L)R_0 + (\sigma_L - \sigma_I)L_{Latent}(\psi)R_0 + \sigma_I \beta L_{Latent}(\psi) \frac{1 - L_{Infectious}(\phi)}{\phi}.$

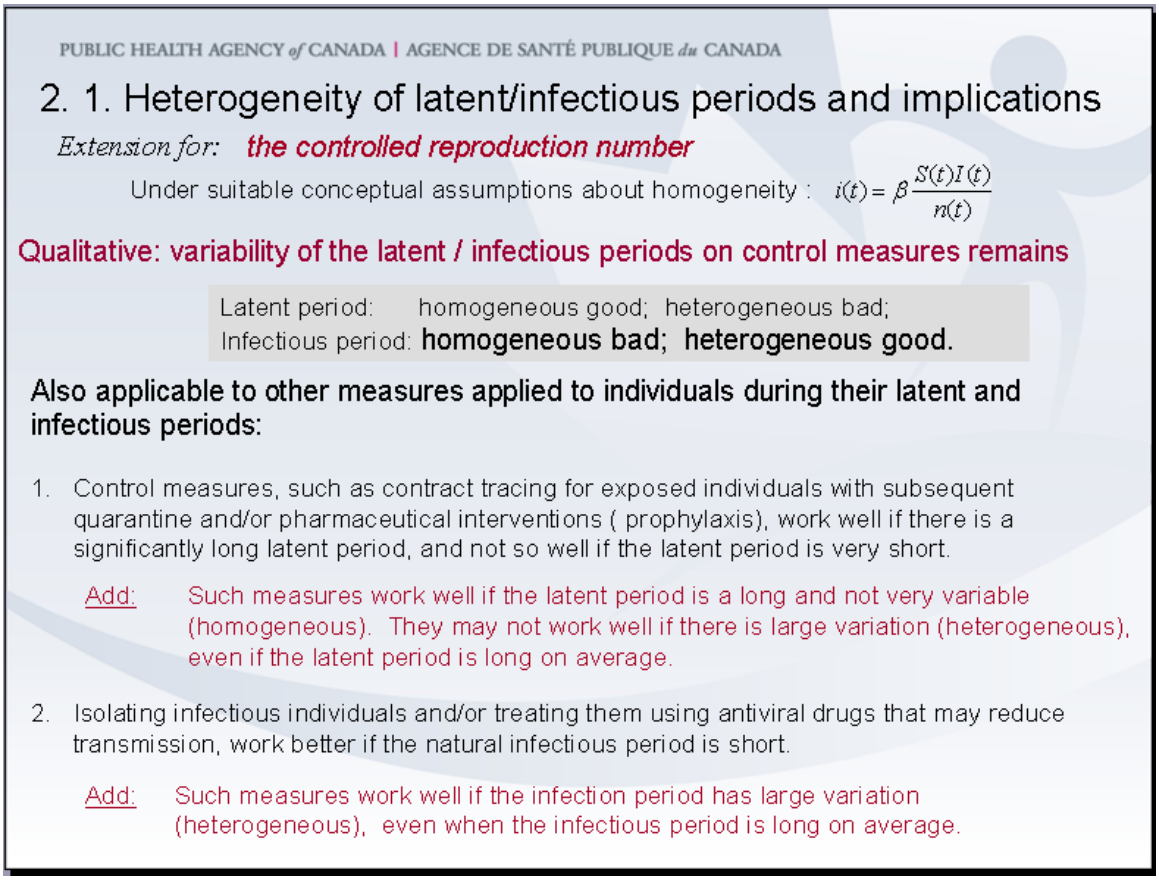
Yan and Feng (*Mathematical Biosciences*, 2010)

Qualitative: variability of the latent / infectious periods on control measures remains

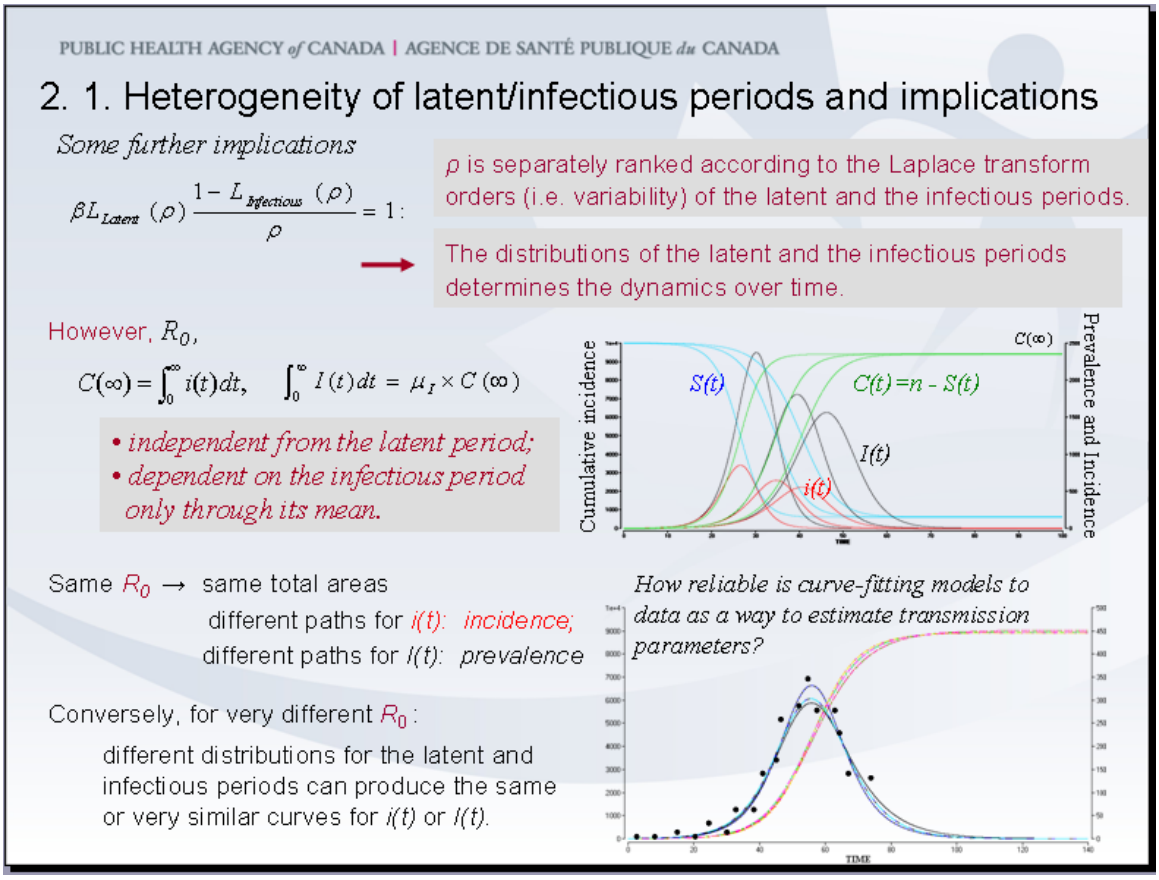
Latent period: homogeneous good; heterogeneous bad;
Infectious period: **homogeneous bad; heterogeneous good.**

For details:

Yan, P. and Feng, Z. (2010) Variability order of the latent and the infectious periods in a deterministic SEIR epidemic model and evaluation of control effectiveness. Mathematical Biosciences 224, 43-52.



These are ad hoc arguments, without rigorous proof.

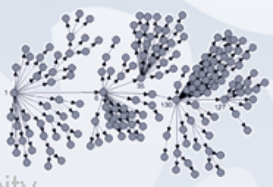


This is my long standing opinion about estimating key epidemic parameters such as R_0, based on observed exponential growth rate.

There are various formulae in the literature, but each of them is crucially dependent on the underlying (and mostly hidden) tactical assumptions about the natural history of an infected individual.

However, I have not been able to provide a better method than those widely used in the literature and understand the importance of estimating R_0 at the very early stage of an epidemic.

Highlight



1. Phenomenological observation

   "A picture is worth 1000 words."

   1.1. Implications of phenomenological heterogeneity

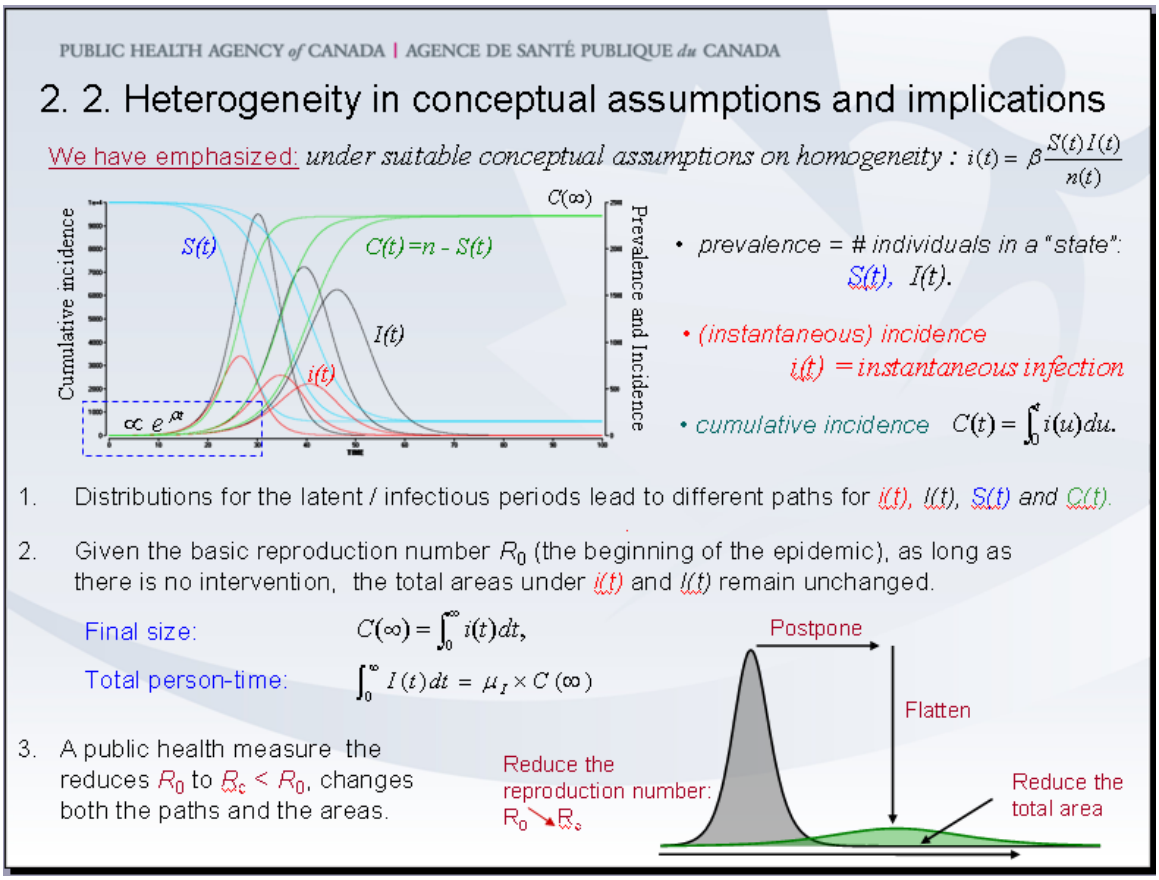
   At *the very beginning:* as an infectious individual *seeded* in an infinitely large susceptible population … …

2. Heterogeneity in the agent-host-environment interface

   - Conceptual assumptions vs. tactical assumptions

   2.1. Heterogeneity in tactical assumptions, implications on dynamics

   2.2. Heterogeneity in conceptual assumptions, implications on dynamics



The very key discussion point from this page onwards is the transcendental relationship between R_0 and the final size.

This relationship is under "the suitable conceptual assumptions on homogeneity".

We shall tease these assumptions apart and show to what extent these assumptions can be relaxed while the transcendental relationship is still valid.

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

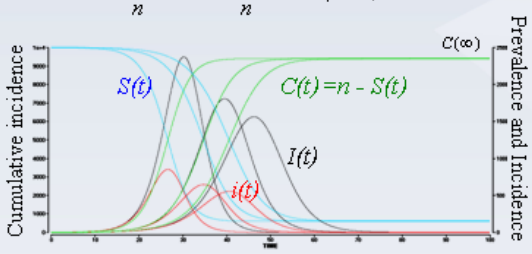
2. 2. Heterogeneity in conceptual assumptions and implications

We have emphasized: *under suitable conceptual assumptions on homogeneity :* $i(t) = \beta \frac{S(t)I(t)}{n(t)}$

This assumption yields a simple expression $R_0 = \beta\mu_I$, and a transcendental relationship

$$1-\eta = (1-\varepsilon)\exp(-R_0\eta)$$

where $\varepsilon = \frac{I(0)}{n}$ and $\frac{E[C(\infty)]}{n} \to \eta > 0$, as $n \to \infty$.



Without intervention,

Final size: $C(\infty) = \int_0^\infty i(t)dt,$

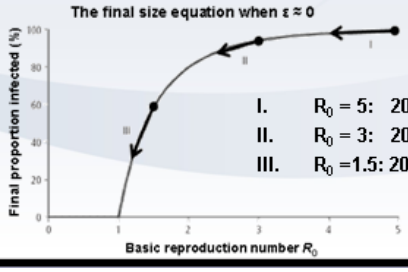
Total person-time: $\int_0^\infty I(t)dt = \mu_I \times C(\infty)$

Both are determined at the very beginning of the epidemic given R_0.

The transcendental relationship

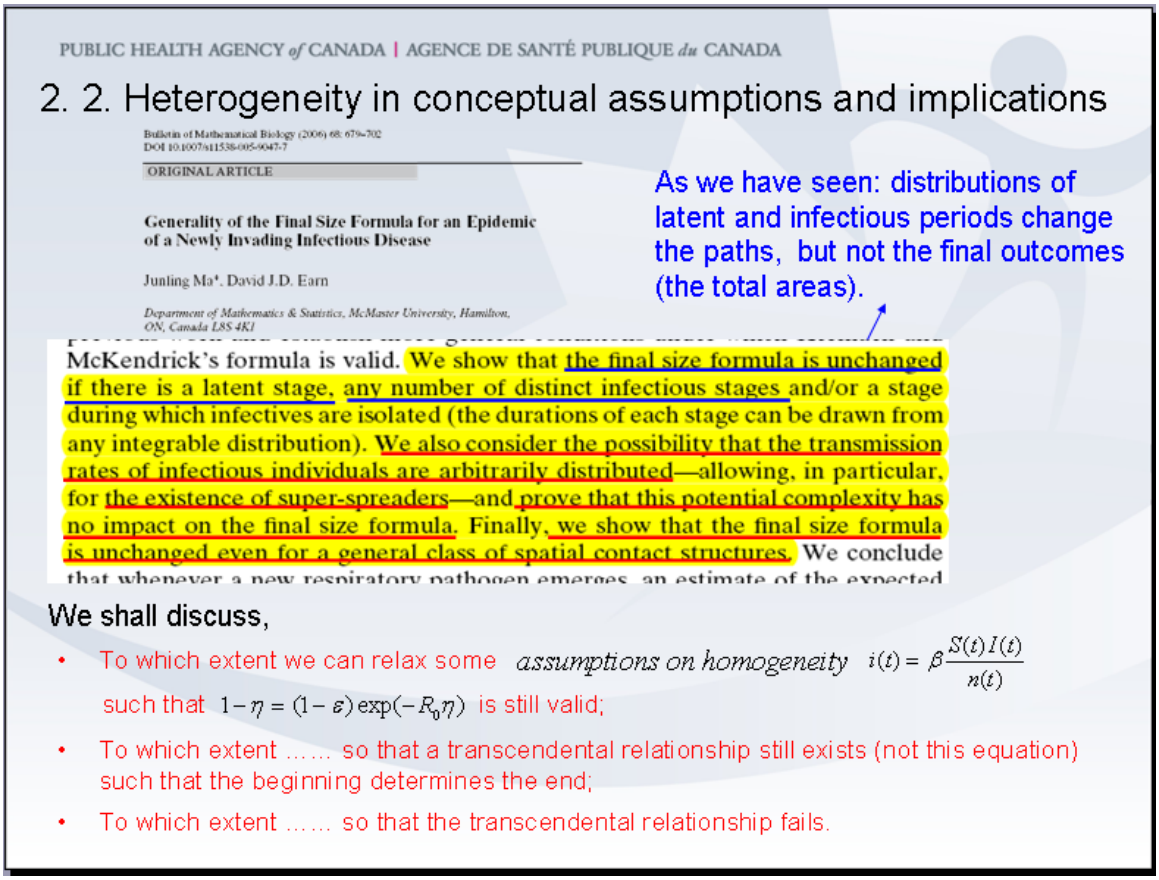
$1-\eta = (1-\varepsilon)\exp(-R_0\eta)$

is extremely useful



I. $R_0 = 5$: 20% reduction of R_0, 1% reduction of final size
II. $R_0 = 3$: 20% reduction of R_0, 6% reduction of final size
III. $R_0 = 1.5$: 20% reduction of R_0, 27% reduction of final size

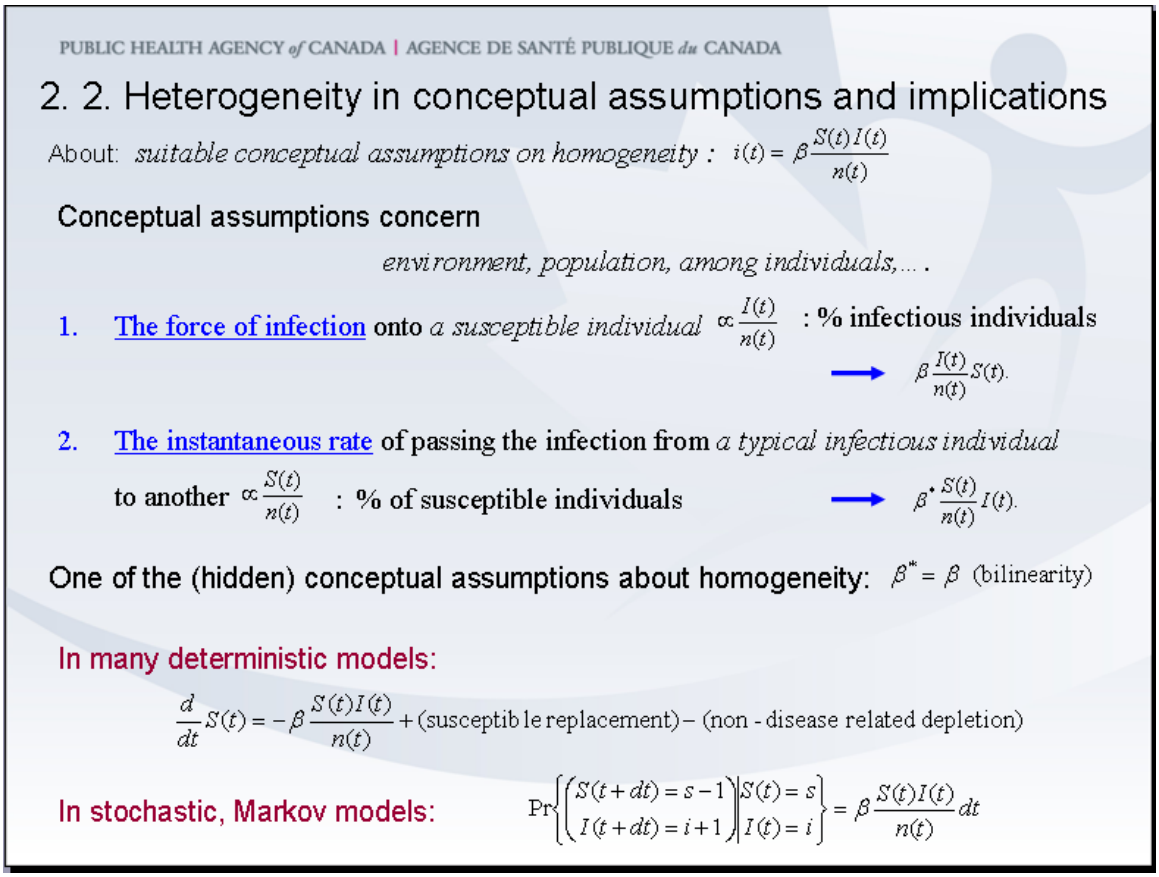
This shows why it is so important to estimate R_0 at the very early stage of an epidemic.



Ma, J. and Earn, D. (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease. Bulletin of Mathematical Biology. 68. 679-702.

The blue underlined texts have been discussed in previous slides.

The red underlined texts will be examined.



Another way to describe this conceptual assumption on homogeneity is that the numbers of new infection contacts generated by a typical infectious individual through its infectious period follow a Poisson process.

2. 2. Heterogeneity in conceptual assumptions and implications

What does *homogeneity* mean regarding bilinearity $\beta\frac{S(t)I(t)}{n(t)}$?

1. The force of infection onto *a susceptible individual:* $\beta\frac{I(t)}{n(t)}$
2. The instantaneous rate of infecting another: $\beta\frac{S(t)}{n(t)}$

Agent: same infectiousness during the study period.

Host:

- All susceptible individuals are the same: equally susceptible.
- All infectious individuals are the same: equally infectious when infectious period starts.
- An infected individual remains equally infectious throughout its infectious period.

Environment (homogeneous mixing): an individual contacts with all other individuals in the population with equal probability. In an infinitely large population, the number of contacts made by a typical individual follows a stationary Poisson process.

$$\beta = \lambda p$$

$$= \text{contact freq.} \times \text{prob. of infection per contact}$$

- independent of time
- Independent of which contact pair

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

2. 2. Heterogeneity in conceptual assumptions and implications

Agent: same infectiousness (√Yes / No)

Host:

All susceptible individuals are the same (√Yes / No)

All infectious individuals are the same (√Yes / No)

Equally infectiousness during infectious period (√Yes / No)

Environment: homogeneous mixing (√Yes / No)

If all the answers are "Yes", then we have the bilinear relationship $\beta\frac{S(t)I(t)}{n(t)}$.

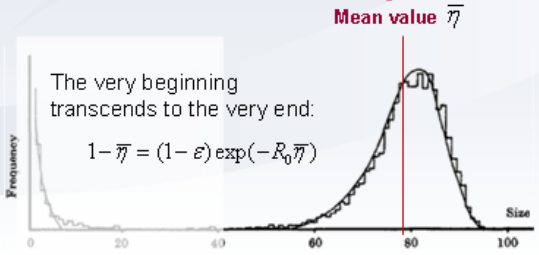
We also have the expression: $R_0 = \beta\mu_I$, where μ_I = average infectious period.

If the population is closed with size n, without replacement of susceptible individuals, the *final size* of the epidemic is meaningful:

$$\eta = \frac{C(\infty)}{n}, \text{ with mean } \overline{\eta} > 0.$$

where $C(\infty)$ is the cumulative number of infected individuals as $t \to \infty$.

The final size distribution of a large outbreak



Not explicitly mentioned here is that when the population size is very large (much larger than n =100 as displayed in the simulation), the final size distribution is asymptotically Gaussian. There is a central limit theorem to prove this.

2. 2. Heterogeneity in conceptual assumptions and implications

Agent: same infectiousness (√Yes / No)

Host:

All susceptible individuals are the same (√Yes / No)

All infectious individuals are the same (√Yes / No)

Equal infectiousness during infectious period (Yes / **X** No)

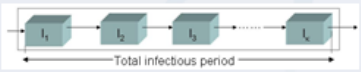
Environment: homogeneous mixing (√ Yes / No)

Staging

Staged infectious period with different infectiousness

$\beta_j : j = 1, \cdots, k$

assigned to each stage.



What have changed:

1. $\beta \frac{S(t)I(t)}{n(t)}$ becomes $(\beta_1 I_1(t) + \cdots + \beta_k I_k(t)) \frac{S(t)}{n(t)}$

   where $I_j(t)$ represents the numbers of infectious individuals in stage $j = 1, \ldots, k$.

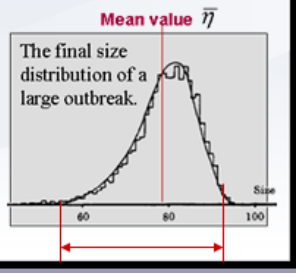
2. $R_0 = \beta \mu_I$ becomes $R_0 = \beta_1 \mu_I^{(1)} + \cdots \beta_k \mu_I^{(k)}$

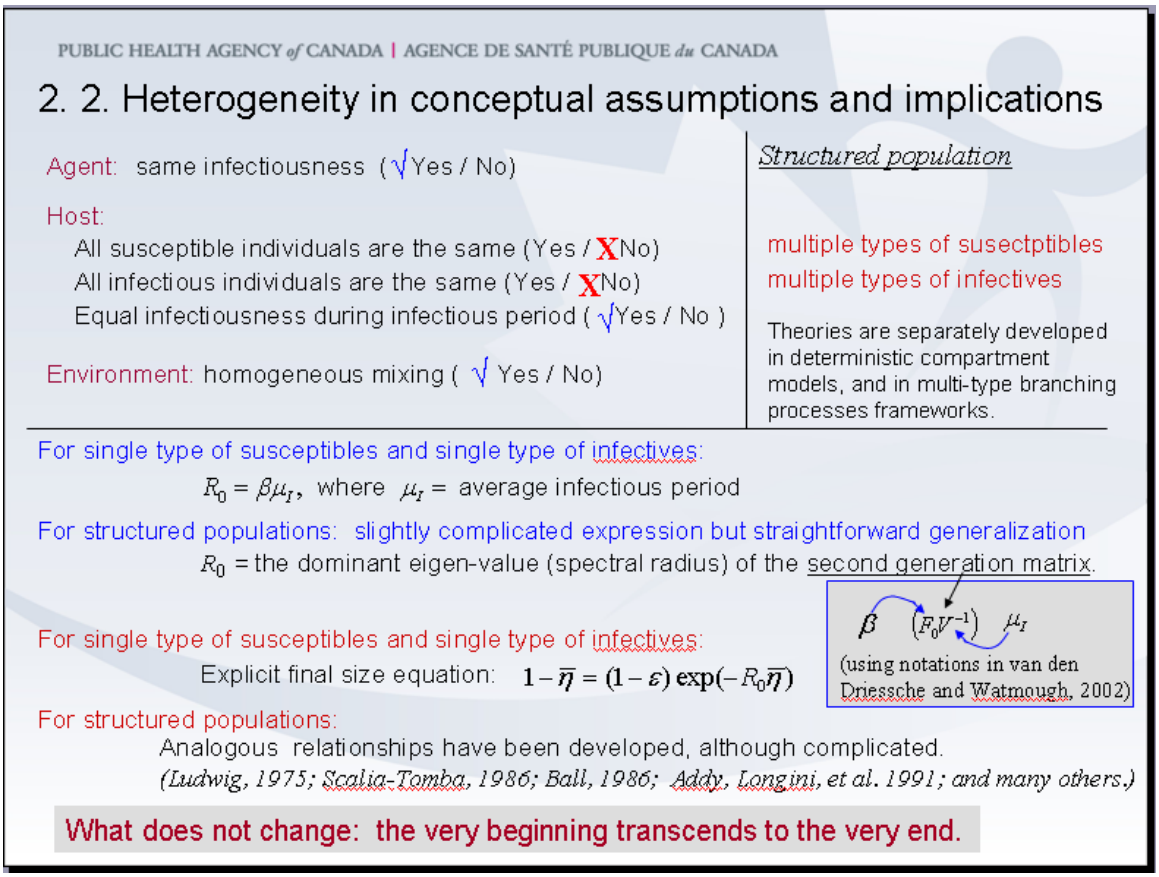
   $\mu_I^{(j)}$ = average time of the j^{th} stage of the infectious period

What does not change: $1 - \overline{\eta} = (1 - \varepsilon) \exp(-R_0 \overline{\eta})$

- no effect to the mean final size *(Ma and Earn, 2006)*

But: - if the variation among β_j's is small, the more the staging, the smaller is the variance for the final size η.



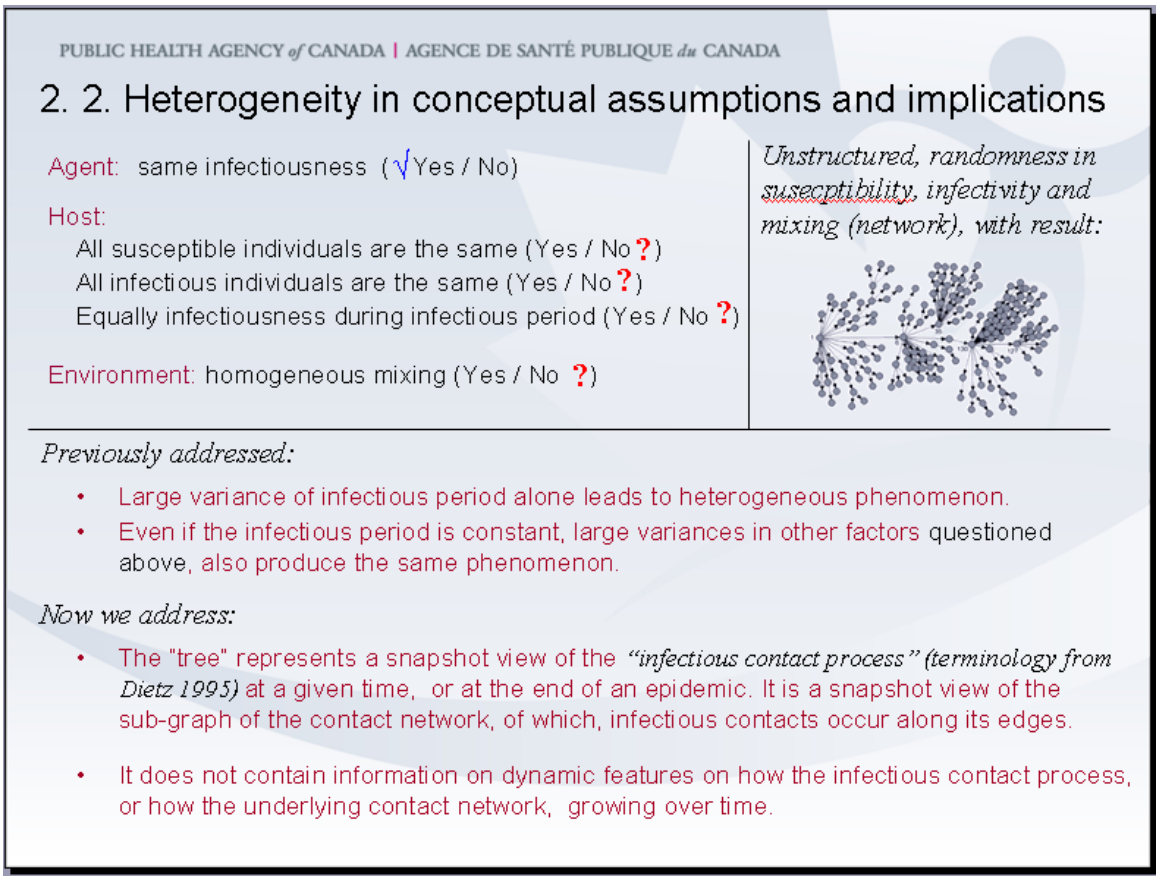


Ludwig, D. (1975) Final size distributions for epidemics. Mathematical Biosciences 23. 33-46.

Scalia-Tomba, G. Asymptotic final size distribution for some chain binomial processes. Advances in Applied Probability. 17, (1985)477-495.

Ball, F. (1986) A unified approach to the distribution of the total size and total area under the trajectory of infectives in epidemic models. Advances in Applied Probability 18, 289-310.

Addy, GL., Longini, IM. and Haber, M. (1991) A generalized stochastic model for the analysis of infectious disease final size data. Biometrics 47, 961-974.



In stochastic processes terms, for a stochastic process {Y(t):t >0 }, a realization of the process at time t , Y(t)=y(t), may be represented by the marginal probability has a mar

Pr {Y(t)=y(t)}

or the conditional probability

Pr {Y(t)=y(t) | given all the past history of Y(s): s<=t ¯ }.

It is the conditional probability presentation defines the path of the stochastic process over time.

The slide shows an analogous argument, by extending these concepts to a “stochastic graph process”.

The snapshot view corresponds to the marginal presentation. It does not contain information of how the graph grows dynamically, which should have been the correspondence of the conditional representation of a stochastic process.

PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

2. 2. Heterogeneity in conceptual assumptions and implications

Agent: same infectiousness (√Yes / No)

Host:

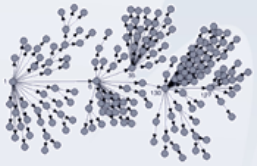
All susceptible individuals are the same (Yes / No ?)

All infectious individuals are the same (Yes / No ?)

Equally infectiousness during infectious period (Yes / No ?)

Environment: homogeneous mixing (Yes / No ?)

Unstructured, randomness in susceptibility, infectivity and mixing (network), with result:



Scenario A: $\beta = \lambda p$ = contact freq. × prob. of infection per contact

λ may be random due to heterogeneous mixing

p may be random due to heterogeneity among individuals susceptibility and/or infectivity

- *β, as a whole, is no longer a constant, but random, with large (but mostly finite) variation*
- *The snapshot view of the tree for infectious contacts has very large variance.*

Dynamic of growth: *stationary increment*

The expected number of infectious contacts by a typical individual is proportional to the length of time during the infectious period.

The two scenarios in the following slides may give the same snapshot views, but with very different stochastic mechanisms on the dynamic features.

2. 2. Heterogeneity in conceptual assumptions and implications

Agent: same infectiousness (√Yes / No)

Host:

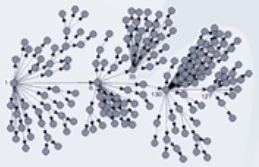
All susceptible individuals are the same (Yes / No ?)

All infectious individuals are the same (Yes / No ?)

Equally infectiousness during infectious period (Yes / No ?)

Environment: homogeneous mixing (Yes / No ?)

Unstructured, randomness in susceptibility, infectivity and mixing (network), with result:



Scenario A: *β is random*

Dynamic of growth: *stationary increment*

The expected number of infectious contacts by a typical individual is proportional to the length of time during the infectious period.

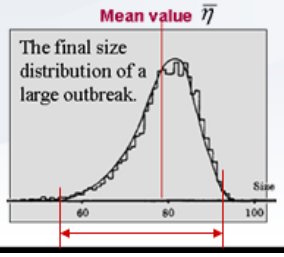
(can be modelled by a mixed-Poisson process)

Results:

$R_0 = \overline{\beta}\mu_I$, where $\overline{\beta}$ is mean of β, μ_I is meaninfectious period

Still true: $1-\overline{\eta} = (1-\varepsilon)\exp(-R_0\overline{\eta})$

- but η will have larger variance.



PUBLIC HEALTH AGENCY *of* CANADA | AGENCE DE SANTÉ PUBLIQUE *du* CANADA

2. 2. Heterogeneity in conceptual assumptions and implications

Agent: same infectiousness (√Yes / No)

Host:

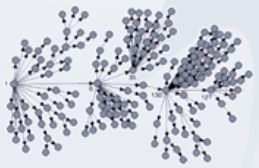
All susceptible individuals are the same (Yes / No ?)

All infectious individuals are the same (Yes / No ?)

Equally infectiousness during infectious period (Yes / No ?)

Environment: homogeneous mixing (Yes / No ?)

Unstructured, randomness in susceptibility, infectivity and mixing (network), with result:



Scenario B: The growth of the tree or the contact network does not maintain *stationary increment*. The expected number of infectious contacts by a typical individual is no longer proportional to the length of time during the infectious period.

Preferential attachment : the more one attracts others, the larger the probability of making more new connections, the # contacts (in any time interval) follows highly skewed distributions (Yule, Waring, power-law, etc.). *("scale-free" networks)*

Variance: can be extremely large, often does not exist (approaches infinity) !

Mean: as defined by mathematical expectation of the form $\int_0^\infty x dF_x(x)$, may not exit.

R_0 is no longer meaningful !

Final size: No well established transcendental relationship between R_0 (if exists) and the mean value of the final size.

There is an ongoing debate whether preferential attachment actually happen in growing networks. Liljeros, et al. are convinced of preferential attachment as a mechanism for sexual networks, as "people become more attractive the more partners they get." However, Jones and Handcock is skeptical and argues that networks with infinitely large variances but dramatically different structures can manifest the same marginal degree distribution, whereas these different network structures produce different epidemic behaviour.

Liljeros, F., Edling, C.R., and Amaral, L.A.N. Sexual networks: implications for the transmission of sexually transmitted infections. Microbes and Infection 3, (2003) 189-196.

Jones, J.H. and Handcock, M.S. An assessment of preferential attachment as a mechanism for human sexual network information. Proceedings: Biological Sciences, The Royal Society. 270. (2003). 1123-1128.

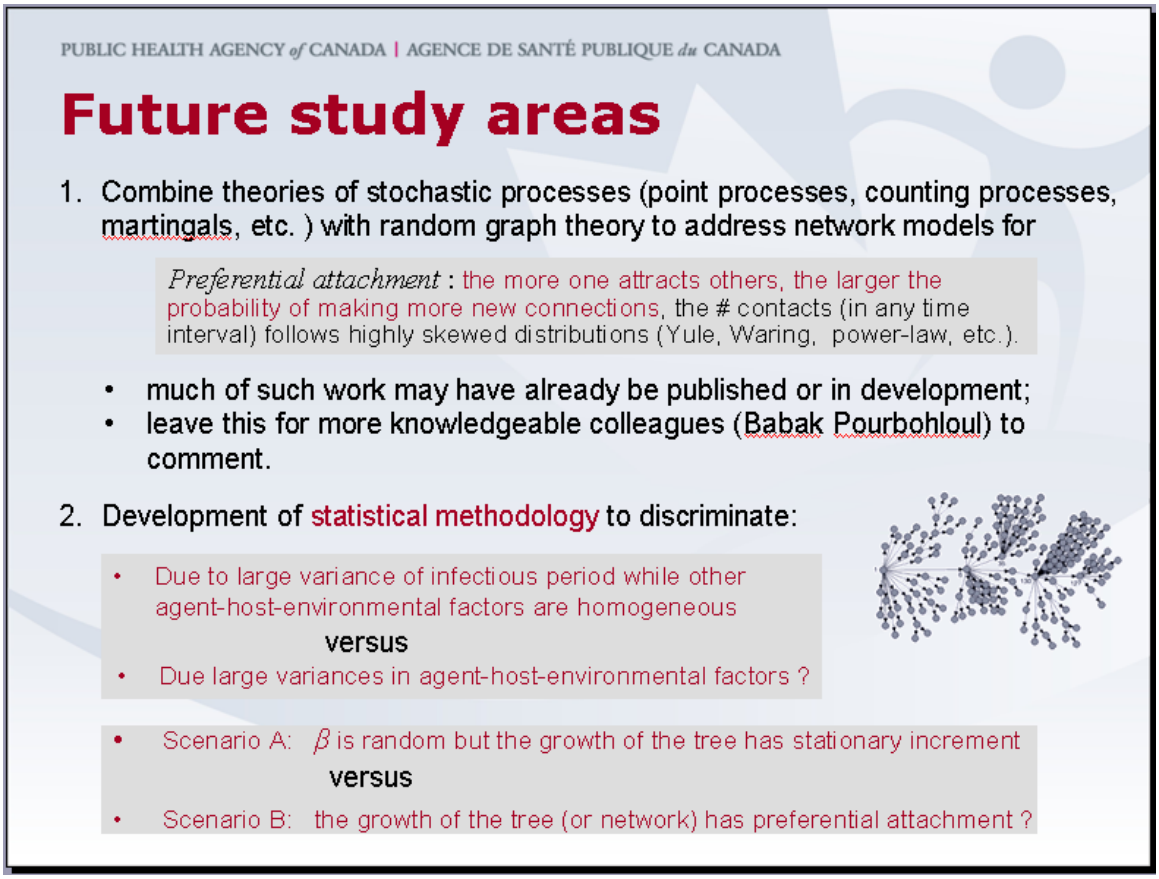
The debate between Scenarios A and B has a longer history . In 1919, Greenwood and Woods put forward three hypotheses into the occurrence of accidents:

1.Pure chance, which gives rise to the Poisson process (corresponding to conceptual assumptions on homogeneity in this presentation);

2.True contagion, i.e. initially all individuals have the same probability of incurring an accident, but this probability is modified by each accident sustained to give rise to the linear pure birth process (corresponding to Scenario B);

3. Apparent contagion (proness), i.e. individuals have constant but unequal probabilities of having an accident and the resulting process being a mixed-Poisson process (corresponding to Scenario A).

Greenwood, M. and Woods, H.M. On the incidence of industrial accidents upon individuals with special reference to multiple accidents. Report of the Industrial Research Board, No.4. London: His Majesty of Stationery Office. (1919).



It is important to develop statistical models and methods to distinguish the two scenarios, as well as the kind of data to seek after. The identical marginal (i.e. snapshot) data may arise from two very different stochastic mechanisms which produce quantitatively very different epidemic behaviour.