

Stochastic modelling of Chlamydial infections

Torrington Callan¹ Stephen Woodcock²

May 18, 2020

Abstract

Chlamydia trachomatis is a bacterial pathogen that can cause serious reproductive harm. We describe a class of stochastic branching processes and their application in modelling the growth of an infection by Chlamydia. Using simulations we show that the model can reproduce biological phenomena of interest, and we show the variability in outcomes of infections under the same parameter conditions. We further speculate how this model might be used to explain long-term adverse reproductive sequelae.

Contents

1	Introduction	2
2	Branching process	2
3	Modelling infectious outcomes	3
3.1	Given an initial infectious load, what are the chances the infection will develop?	4
3.2	Given an initial infectious load, what are the chances the infection will last until a given stopping time?	4
3.3	Given an initial infectious load, what are the chances the infection will be above a certain threshold at a given stopping time?	5
3.4	Given an initial infectious load and subcritical/critical reproduction, when will extinction occur?	6
3.5	Given an initial infectious load and subcritical/critical reproduction, what will the total infectious load be?	6
4	Modelling the immune system	8
4.1	Macrophage engulfment	8
4.2	Cell-mediated immune response	8
4.3	Burst size	9
5	Concluding remarks	9

1 Introduction

Chlamydia trachomatis is a bacterial pathogen and the cause of the most common notifiable sexually transmitted infection. There is considerable evidence to indicate that a Chlamydial infection in women may cause Pelvic Inflammatory Disease (PID), and that these women are more likely to experience tubal factor infertility [4]. The development of tubal pathology is hypothesised to be a result of a Chlamydial infection ascending beyond the cervix and into the upper reproductive tract, combined with a pro inflammatory response to the pathogen [7].

The low rate of pathology development in women with infections suggests that a number of host response and pathogen factors moderates ascension. The immune system response is particularly critical in determining the development of pathology [6].

An infection develops by extracellular Elementary Bodies (EBs) entering into a cell. Inclusions may have many bacteria per cell. Once within the host cell, the metabolically inert EB differentiates into the Reticulate Body (RB) form of Chlamydia. The RBs replicate by binary fission for some cycles of division, until some RBs convert back into EBs while others begin to replicate. The cycle of development continues until approximately 40-48 hours post infection, after which EBs continually exit the cell in an extrusion, which leaves the host cell intact, or via the lytic cycle in which the host cell releases a burst of EBs. The burst of EBs released can then go on to infect new cells, thus beginning the cycle again [1].

The micro-population dynamics of a bacterial infection and the associated host cell response have been represented by a system of ordinary differential equations [10]. More realistic behaviour was represented with a spatial dimension (represented with a set of partial differential equations) [5]. Host response scenarios can be modelled by changing parameter values of a particular equations, which in turn yields insights into the long-term behaviour of the system under different conditions. However, this modelling approach does not capture the randomness of the initial phase of an infection.

As a precursor to explaining PID and infertility, we must describe the requirements for an infection to become established and to survive in the mucosal layer of the cervix for a sufficient period of time until ascension can occur and pathology can develop. In the model we describe here, we explicitly keep track of the count of infected cells and account for the stochastic elements present in the development of an infection.

2 Branching process

Let N_t be the number of infected host cells at time $t \geq 0$, with initial conditions $N_t = N_0$. At time t , for the i^{th} infected cell, with $i = 1, \dots, N_t$, let Y_t^i be a t -measurable random variable that represents the number of infected progeny a particular cell generates. We assume that Y_t^i is independent and identically

distributed according to some distribution F that is not a function of time, nor of the current state of the process N_t . We then model the number of infected cells from previous incidences as

$$N_t = \sum_{i=1}^{N_{t-1}} Y_t^i.$$

This defines the basic single-type branching process [2].

The progeny distribution f has mean and variance μ and σ^2 respectively. These two moments characterise the main properties of the process. The trajectories of the process either die out or explode, such that

$$\mathbb{P}(\{\lim_{t \rightarrow \infty} N_t = \infty\} \cup \{\lim_{t \rightarrow \infty} N_t = 0\}) = 1.$$

The mean number of progeny μ acts as the bifurcation parameter of the process, in that $\mu \leq 1$ implies the almost sure extinction of the process. As a point of interest, the deterministic model $N_t = \mu N_{t-1}$ has the same bifurcation parameter. However, in the deterministic model the population persists to $N_t = N_0$. The stochastic model with $\mu = 1$ goes extinct almost surely.

In the supercritical case $\mu > 1$, the probability of the extinction of the process is given as the fixed point $q = f(q)$, where $f(z) := \mathbb{E}(z^{Y_t^i})$. That is, the probability generating function of the infectious progeny distribution gives an expression for the probability of extinction of an infection.

For the critical and sub-critical cases, define T to be the extinction time of the process. Then $N_T = \sum_{k=0}^T N_k$ follows a power series distribution when Y_t^i follows a power series distribution. That is, $\mathbb{P}(Y_t^i = k) = \alpha_k \lambda^k$. In particular, if the progeny random variable follows a Poisson distribution, with parameter λ , then the total infectious load follows a Borel-Tanner distribution, with parameters λ and N_0 [3].

3 Modelling infectious outcomes

The branching process model tracks the randomness in the size of each burst, along with the chances these bacteria that go on to infect another cell. The primary entities in this setup are the number of infected host cells. Changes in the amount of extracellular infectious material are represented in the model to support a mechanistic interpretation of the results.

The model is generated with events occurring in time according to a Poisson process with parameter $\lambda = 0.25$. When a lysis or extrusion event occurs, a random draw is made from a Binomial distribution with parameters $n = 200, p = 0.004$, to represent the burst size and chance of each bacteria infecting a new cell respectively. The random draw quantity is then added to the total number of infected cells, with one subtracted to represent the original cell's demise, so that there is a significant chance of no infectious progeny being produced and the population total decreasing in size. The value of parameters in

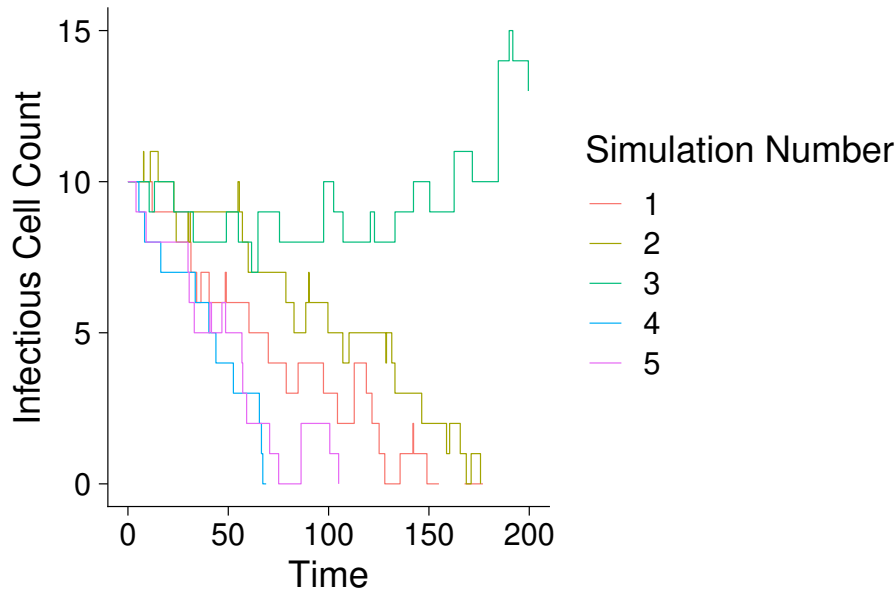


Figure 1: Simulation of a branching process with event times according to a Poisson Process and Binomially distributed offspring distribution

these simulations were chosen for illustrative purposes, but are mostly consistent with other models in the literature [10] [5].

3.1 Given an initial infectious load, what are the chances the infection will develop?

In Figure 1, we can observe the qualitative difference between realisations of a branching process under the same parameter conditions. Simulation four dies out almost immediately, whereas simulation three has a similar infected cell counts to the initial starting population. It is interesting to note that while the survival time in simulation one and two is approximately similar, the total cell count over time is significantly greater in two than in one. This represents a scenario in which the infectious burden is greater for one realisation compared to another, conditional on survival until a particular time.

3.2 Given an initial infectious load, what are the chances the infection will last until a given stopping time?

In Figure 2, we simulate the branching process under the conditions as in Figure 1. We expect that the proportion of zeroes is a function of the stopping time, since the zero is an absorbing state of the process. At time 72, less than

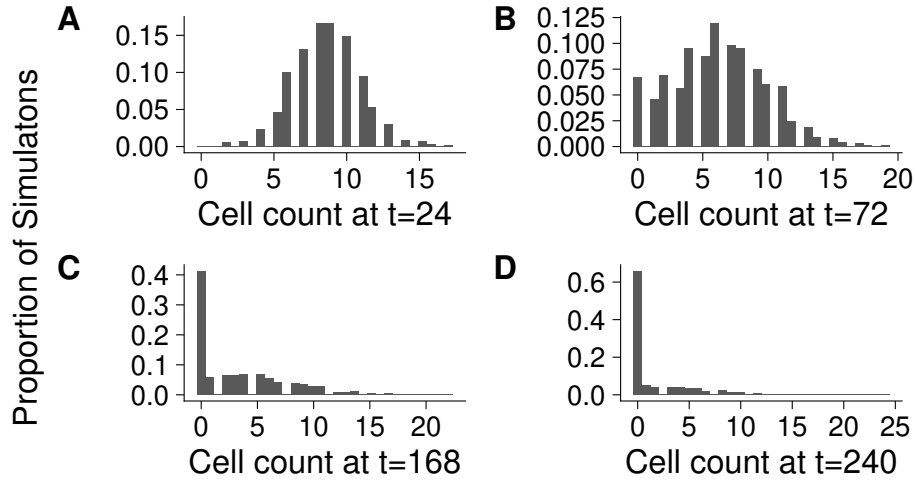


Figure 2: A comparison of cell count distributions for stopping times of A) 24, B) 72, C) 168 and D) 240 hours respectively

a tenth of processes have a cell count of zero, whereas a majority have a cell count of zero at time 240. Alternatively, the closer the stopping time to the initial time, the more processes we expect to observe with a positive population count. Figure 2 shows the change in the distribution of cell counts for varying stopping times.

We count the number of cells at a stopping time $t = 240$, as this represents sufficient time for an infection to clear in most cases. The simulation is repeated 1000 times, to give an empirical density of the population count at a stopping time. The histogram in Figure 2 D) shows the normalised density.

The histogram shows that the distribution of cell counts is dominated by zeroes. In this simulation approximately two-thirds of the simulations had gone extinct by the stopping time.

The other implication of this example is that the distribution of positive cell counts have a long right-tail. Although we expect a significant proportion of processes to have a cell count of zero by a defined stopping time, there is still some significant probability of observing a large population count at the same stopping time. These unlikely but possible may explain in part why some infections cause no observable harm, whilst others persist.

3.3 Given an initial infectious load, what are the chances the infection will be above a certain threshold at a given stopping time?

Whilst the previous question is concerned with the quantity $\mathbb{P}(N_T | t = T)$, this question asks $\mathbb{P}(N_T > N^* | t = T)$ for any arbitrary threshold $N^* \geq 0$.

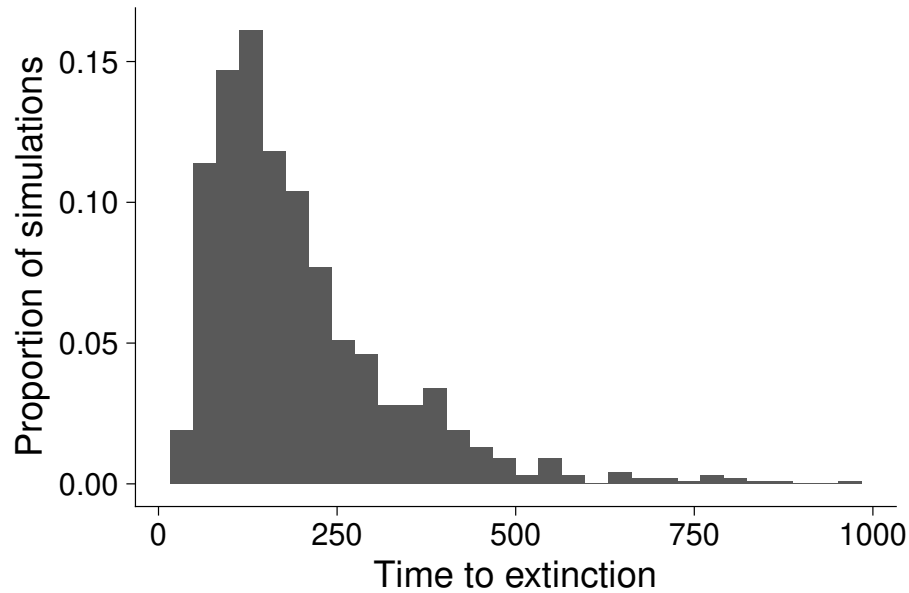


Figure 3: Histogram of time to extinction

We would like to know not only if an infection will be present at a certain point in time, but that it will be present in large enough quantities to cause inflammatory damage. This is crudely addressed with each of the subfigures in Figure 2. For a particular stopping time, the proportion of cell counts above a particular threshold corresponds to a region under the empirical distribution. The figures show that the smaller stopping times are seen to correspond to a greater region of probability density compared to larger stopping times.

3.4 Given an initial infectious load and subcritical/critical reproduction, when will extinction occur?

We are interested in the time to extinction of a branching process with critical or subcritical reproduction, as this is a factor for total infectious load. A longer time to extinction will imply a greater infectious load. It is a straightforward consequence of the results above that most processes will go extinct in a short amount of time. We observe the existence of a long right-tail in the distribution of extinction times in Figure 3.

3.5 Given an initial infectious load and subcritical/critical reproduction, what will the total infectious load be?

The total infectious load is defined as the total cell population count multiplied by the length of each cell's lifetime. It is the metric that gives the best sense of

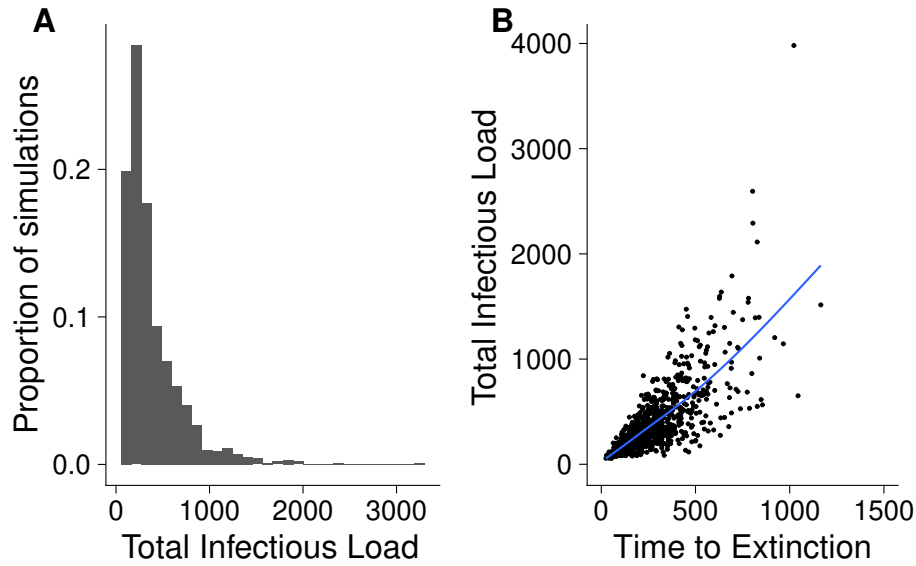


Figure 4: A) Histogram of total infectious load B) The relationship between the time to extinction and total infectious load, where the plotted line is least squares best fit

how severe a particular infection may be, as inflammation due to the immune response is a function of total infectious load. Figure 4 A) shows a similar density type to that of the extinction time, with a majority of processes causing a small amount of total load and a long right-tail.

Figure 4 B) also shows the relationship between the time to extinction and total infectious load. The two are linearly related, since total infectious load is defined as a the product of extinction times and population count. It is interesting to note that there is a large degree of variability around this mean relationship. There exist some cases where the stopping time of the process is short, but the total infectious load is quite high. Conversely, there are some processes where the extinction time is long but the total infectious load is much lower.

In all of the above examples, it appears that a majority of processes will result in a small impact for the given measure of interest (stopping cell count, extinction time, total infectious load). However, for a number of processes this will not be the case, and the impact (for the given measure of interest) will be large. When considering the chances of a particular infection to progress to further disease, we recognise that most infections do not progress to disease at all, but those that do may have a large impact.

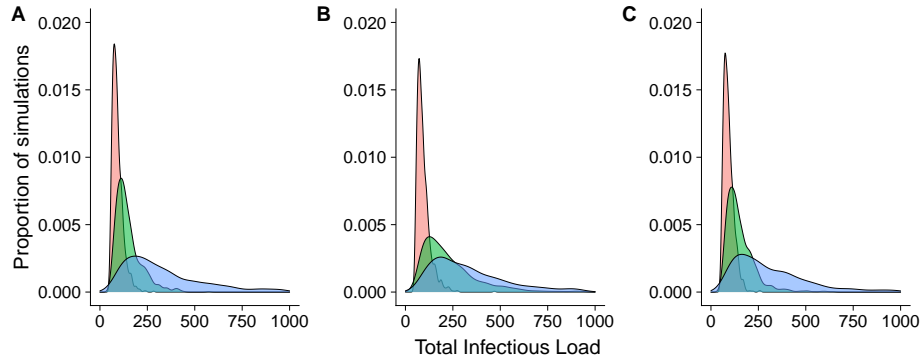


Figure 5: Density of total infectious load for three different immune system responses, where the red densities represents a high immune response, green a medium response and blue representing no immune response. A) Macrophage Engulfment B) Clearance prior to lysis C) Reduction in burst size.

4 Modelling the immune system

We map different responses of the immune system onto parameters of the model. For example, a larger adaptive immune system response may reduce the tail of the progeny distribution. This allows us to use our model to determine how the immune system plays a role in modifying particular outcomes, such as the stopping time of a sub-critical infection or the total infectious load. We describe three specific immune system responses and the outcomes of the model they produce.

4.1 Macrophage engulfment

Macrophages are components of the innate immune system that migrate within tissue and detect the presence of pathogens. Chlamydia are cleared by macrophage engulfment prior to host cell infection [8]. We model this immune system response by varying the success probability of an extra-cellular bacteria infecting a new host cell. Figure 5 plots the total infectious load (as defined above) for three scenarios, where $p = 0.04$ representing no immune response of this type and $p = 0.03, p = 0.02$ representing medium and high responses of this type, respectively.

4.2 Cell-mediated immune response

The T helper cells, particularly TH1 cells (also known as CD4+ cells) are components of the adaptive immune system that are capable of detecting damaged cells of the host, which increases the hosts ability to clear an infection by removing an infected cell prior to lysis or extrusion occurring [9]. This is equivalent to the burst size of the infected cell being equal to zero. We introduce an extra

parameter p_0 to our model, that represents the probability a cell will be cleared prior to lysis or extrusion. To model medium and high immune responses, we simulate the total infectious load for $p_0 = 0.1, 0.5$.

4.3 Burst size

The burst size may also be impacted by a varying immune response [10]. We model this by first assuming that the burst size of an infected cell is drawn from a Poisson distribution with parameter $\lambda = 200$, as opposed to a fixed burst size as above. To represent a medium and high immune response of this type, we then modify the mean of the burst distribution to $\lambda = 150, 100$ respectively.

Figure 5 shows a comparison between no response, medium and high responses of each of the three types described above. This exercise constitutes a sensitivity analysis of the model, however it does give some sense of the impact of the response by each component of the immune system. It should be noted that each response is considered independently of the others, and assumed constant over time, which is a biological oversimplification.

The comparison above demonstrates that for this model, an effective immune response that neutralises free extra-cellular particles has the greatest impact on the distribution of total infectious load, as the behaviour of the model is most sensitive to changes in the success probability parameter.

5 Concluding remarks

In this article we consider a class of stochastic processes for modelling chlamydial infections, and show that realisations of a branching process that has a mechanistic definition reproduce biological phenomena. We demonstrate how this class of models permit large variability in outcomes. The model is a significant first step in work to best elucidate the mechanism by which an infection will result in severe reproductive sequelae.

Acknowledgements

We acknowledge the ongoing support of Wilhelmina Huston and the Chlamydia research group at UTS, and James Brown for assistance with the first author's thesis. We also thank Alex Badran of Spriggy Labs for his feedback on the manuscript. This research is supported by an Australian Government Research Training Program (RTP) Scholarship.

References

- [1] Yasser M. AbdelRahman and Robert J. Belland. "The chlamydial developmental cycle". In: *FEMS Microbiology Reviews* 29.5 (Nov. 2005), pp. 949–959. ISSN: 0168-6445. DOI: 10.1016/j.femsre.2005.03.002.

- [2] T. E. Harris. “Branching Processes”. In: *Ann. Math. Statist.* 19.4 (Dec. 1948), pp. 474–494. DOI: 10.1214/aoms/1177730146.
- [3] Christine Jacob. In: *International journal of environmental research and public health* 7.3 (Mar. 2019), pp. 1186–1204. DOI: 10.3390/ijerph7031204.
- [4] N Low et al. “Incidence of severe reproductive tract complications associated with diagnosed genital chlamydial infection: the Uppsala Women’s Cohort Study”. In: *Sexually Transmitted Infections* 82.3 (2006), pp. 212–218. ISSN: 1368-4973. DOI: 10.1136/sti.2005.017186.
- [5] Dann Mallet et al. “A Mathematical Model of Chlamydial Infection Incorporating Movement of Chlamydial Particles”. In: *Bulletin of mathematical biology* 75 (Oct. 2013). DOI: 10.1007/s11538-013-9891-9.
- [6] Heather K. Maxion et al. “The Infecting Dose of *Chlamydia muridarum* Modulates the Innate Immune Response and Ascending Infection”. In: *Infection and Immunity* 72.11 (2004), pp. 6330–6340. ISSN: 0019-9567. DOI: 10.1128/IAI.72.11.6330-6340.2004.
- [7] S. Menon et al. “Human and Pathogen Factors Associated with Chlamydia trachomatis-Related Infertility in Women”. In: *Clinical Microbiology Reviews* 28.4 (2015), pp. 969–985. ISSN: 0893-8512. DOI: 10.1128/CMR.00035-15.
- [8] D.P. Wilson and D.L.S. McElwain. “A model of neutralization of Chlamydia trachomatis based on antibody and host cell aggregation on the elementary body surface”. In: *Journal of Theoretical Biology* 226.3 (2004), pp. 321–330. ISSN: 0022-5193. DOI: <https://doi.org/10.1016/j.jtbi.2003.09.010>.
- [9] D.P. Wilson, P. Timms, and D.L.S. McElwain. “A mathematical model for the investigation of the Th1 immune response to Chlamydia trachomatis”. In: *Mathematical Biosciences* 182.1 (2003), pp. 27–44. ISSN: 0025-5564. DOI: [https://doi.org/10.1016/S0025-5564\(02\)00180-3](https://doi.org/10.1016/S0025-5564(02)00180-3).
- [10] David Wilson. “Mathematical Modelling of Chlamydia”. In: *The ANZIAM Journal* 45 (Apr. 2004). DOI: 10.21914/anziamj.v45i0.883.

Author addresses

1. **Torrington Callan**, School of Mathematical and Physical Sciences, University of Technology Sydney, Sydney, Australia
<mailto:torrington.callan@student.uts.edu.au>
2. **Stephen Woodcock**, School of Mathematical and Physical Sciences, University of Technology Sydney, Sydney, Australia
<mailto:stephen.woodcock@uts.edu.au>