Stochastic modeling of in vitro bactericidal potency

Anita Bogdanov[*], Péter Kevei[†], Máté Szalai[‡], Dezső Virok[§]

Abstract

We provide a Galton–Watson model for the growth of a bacterial population in the presence of antibiotics. We assume that bacterial cells either die or duplicate, and the corresponding probabilities depend on the concentration of the antibiotic. Assuming that the mean offspring number is given by $m(c) = 2/(1 + \alpha c^\beta)$ for some α, β, where c stands for the antibiotic concentration we obtain weakly consistent, asymptotically normal estimator both for (α, β) and for the minimal inhibitory concentration (MIC), a relevant parameter in pharmacology. We apply our method to real data, where *Chlamydia trachomatis* bacteria was treated by azithromycin and ciprofloxacin. For the measurements of *Chlamydia* growth quantitative PCR technique was used. The 2-parameter model fits remarkably well to the biological data.

Keywords: multitype Galton–Watson process, asymptotically normal estimator, quantitative PCR, *Chlamydia*, MIC.

MSC2020: 60J85, 92C70,

1 Introduction

Since the discovery of penicillin, antibiotics have been used increasingly worldwide to treat bacterial infections. As the overuse of antibiotics may results drug-resistant bacteria, determining the bactericidal potency is of the utmost importance.

---

[*]Department of Medical Microbiology and Immunobiology, University of Szeged, Dóm tér 10, 6720 Szeged, Hungary; e-mail: `varga-bogdanov.anita@med.u-szeged.hu`

[†]Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary; e-mail: `kevei@math.u-szeged.hu`

[‡]Bolyai Institute, University of Szeged, Aradi vértanúk tere 1, 6720 Szeged, Hungary; e-mail: `szalaim@math.u-szeged.hu`

[§]Department of Medical Microbiology and Immunobiology, University of Szeged, Dóm tér 10, 6720 Szeged, Hungary; e-mail: `virok.dezso.peter@med.u-szeged.hu`

In the present paper the bacterial population is modeled by a Galton–Watson branching process. The offspring distribution, in particular the offspring mean $m(c)$ depends on the antibiotic concentration $c > 0$ as

$$m(c) = m_{\alpha,\beta}(c) = \frac{2}{1 + \alpha c^\beta}, \tag{1}$$

where $\alpha > 0$, $\beta > 0$ are unknown parameters. Under this model the minimal inhibitory concentration (MIC), the smallest antibiotic concentration preventing bacterial growth, is the smallest c for which $m(c) = 1$, that is $\alpha^{-1/\beta}$. Based on measurements at different concentrations we obtain weakly consistent asymptotically normal estimator both for (α, β), and for the MIC.

We assume that the bacterial population is homogeneous, all the cells behave similarly. In particular, there is no resistant type. As mutation is rare under normal conditions and in short time, this is a natural assumption for our dataset. Long-term evolution of bacterial populations with both *resistant* and *susceptible* types was investigated in several papers using deterministic models, see Svara and Rankin [11], Paterson et al. [10], and the references therein. Closest to our model is the deterministic model given by Liu et al. [7]. In [7] a deterministic expression for the number of colony forming units is obtained in terms of the antibiotic concentration.

Branching processes are classical tools to model cell proliferation, see the monographs by Haccou et al. [3], Kimmel and Axelrod [5]. However, to the best of our knowledge for estimation of bactericidal potency of antibiotics only deterministic models are used.

In the experiments growth of *Chlamydia trachomatis* bacterial population was a analyzed by quantitative PCR (qPCR) method with 12 different antibiotic concentrations and 2 different antibiotics.

Chlamydiae are obligate intracellular bacteria that primarily infect epithelial cells of the conjunctiva, respiratory tract and urogenital tract. They have a unique developmental cycle, with two phenotypic bacterial forms, the elementary body (EB) and the reticulate body (RB). The EB is the infectious form that can be found outside of the host cells and it is not capable to multiply. After infection of the host cell, the EB differentiates to RB. The RB multiplies in the host cell by binary fission in a specific area of the infected host cell, the inclusion. After a certain period of time, depending on the chlamydial species, the RB redifferentiates to EB. The EB is then released from the host cell ready to infect new host cells. This unique life-cycle triggered lot of mathematical work to model the growth of the population. Wilson [14] worked out a deterministic model taking into account the infected and uninfected host cells and the extracellular *Chlamy-*

2

dia concentration. Wan and Enciso [13] formulated a deterministic model for the quantities of RB's and EB's, and solved an optimal control problem to maximize the quantity of EB's when the host cell dies. The same problem in a stochastic framework was investigated by Enciso et al. [1] and Lee et al. [6]. In these papers population growth is modeled without the presence of antibiotic.

There is a third form of the bacterium, the aberrant body or persistent body. This form is induced by various adverse environmental stimuli, such as the lack of nutrients and the presence of antibiotics, see Panzetta et al. [9]. The persistent body is not capable to multiply. After elimination of the stress stimuli, the persistent body may reenter the normal developmental cycle, differentiates to RB, multiplies and redifferentiates to EB. If there is an excess of antibiotics reaching the so-called bactericide concentration, the bacterium is killed, and no multiplication can be observed. A lower antibiotic concentration does not kill all of the bacterium, but leads to the formation of non-multiplying aberrant bodies. Further lowering the antibiotic concentration more RB can be observed, while the formation of aberrant body decreases. At very low antibiotic concentration, the antibiotic has no effect on the bacterial growth and all the bacteria enter the normal developmental cycle. Azithromycin and doxycycline are the most commonly used antibiotics in *Chlamydia* infections (Miller [8]), but *Chlamydiae* are also sensitive to quinolone type antibiotics (Vu et al. [12]). In our study *Chlamydia trachomatis* infected cells were treated with azithromycin and the quinolone ciprofloxacin. The dose response curves, the concentration dependent impacts of these antibiotics on chlamydial growth were measured 48 hours post infection. A major challenge is the accurate measurement of chlamydial growth. The golden standard is the immunofluorescent labeling and manual counting of the chlamydial inclusions, which has several disadvantages, including that the concentration of the individual bacteria cannot be counted. Instead of counting the bacterial cells, the quantity of bacterial genomes (which is a constant times the number of bacteria) can also be measured. Chlamydial genome concentration in the infected host cells can be measured by a quantitative polymerase chain reaction (qPCR). This method is accurate and theoretically measures the genome of all individual bacteria. Eszik et al. [2] developed a version of the qPCR, the so-called direct qPCR method for chlamydial growth monitoring. Direct qPCR is capable to perform qPCR measurements without the labor-intensive DNA purification. The qPCR method gives a so-called cycle threshold (Ct) value to each bacterial sample. If the effectivity of the qPCR is 100% then the

theoretical Ct value equals $a - \log_2 Z^{(i)}_{n;c,x_0}$, where $a \in \mathbb{R}$ is an unknown constant and $Z^{(i)}_{n;c,x_0}$ stands for the total number of dead and alive bacterial cells at antibiotic concentration $c > 0$, after n generations starting with x_0 bacterias, in experiment i. Adding a measurement error, the measurements have the form

$$C_i(c, x_0) = a - \log_2 Z^{(i)}_{n;c,x_0} + \varepsilon_{i;c}, \qquad i = 1, \ldots, N, \qquad (2)$$

where measurement error $\varepsilon_{i;c}$ is assumed to be Gaussian with mean zero, and variance σ^2_ε. This simple linear model is suggested by Yuan et al. [15]. Due to the measurement method lower Ct value means higher genome concentration. The dose response curves measured by a direct qPCR method are given in Figures 3 and 4.

The rest of the paper is organized as follows. The model and some basic properties are given in Section 2. The estimator of $m(c)$ for c fixed is provided in Section 3, while in Section 4 we consider different antibiotic concentrations together. Section 5 contains a small simulation study, and real data is analyzed in Section 6. The proofs are gathered together in the Appendix.

2   The theoretical model

We consider a simple Galton–Watson branching process where the offspring distribution depends on the antibiotic concentration $c \geq 0$. Each bacteria either dies (leaves no offspring), survives (leaves 1 offspring), or divides (leaves 2 offsprings) with respective probabilities $p_0 = p_0(c)$, $p_1 = p_1(c)$, and $p_2 = p_2(c)$. Let $f(s) = f_c(s)$ denote the offspring generating function and $m = m(c)$ the offspring mean if the antibiotic concentration is c, i.e.

$$f(s) = f_c(s) = \mathbf{E}s^{\xi_c} = \sum_{i=0}^{2} p_i(c)s^i, \quad s \in [0, 1],$$

$$m = m(c) = f'_c(1) = \mathbf{E}\xi_c,$$

where ξ_c is the number of offsprings. The process starts with $X_0 = x_0$ initial individuals, and

$$X_{n+1;c} = \sum_{i=1}^{X_{n;c}} \xi^{(n)}_{i;c},$$

where $\{\xi_c, \xi^{(n)}_{i;c} : i \geq 1, n \geq 1\}$ are independent and identically distributed (iid) random variables with generating function f_c. Note that the offspring

4

distribution does depend on the antibiotic concentration c, but here and in the next section we suppress this dependence from the notation.

Using the qPCR method the observed quantity is the genom of all individual bacteria, which is a constant times the *total* number of bacteria, that is live and dead cells together. Therefore, we have to keep track of the dead bacterias too. In order to do this we consider a two-type Galton–Watson branching process $\mathbf{X}_n = (X_n, Y_n)$, $n \geq 0$, where X_n, Y_n stands for the number of alive, dead bacterias respectively, in generation n. Then the total number of bacteria at generation n is $Z_n = X_n + Y_n$. We also write Z_{n,x_0} to emphasize that $X_0 = x_0$. The process evolves as

$$X_{n+1} = \sum_{i=1}^{X_n} \xi_i^{(n)}$$

$$Y_{n+1} = Y_n + \sum_{i=1}^{X_n} \eta_i^{(n)}, \quad n \geq 0,$$

$(X_0, Y_0) = (x_0, 0)$, where (ξ, η), $(\xi_i^{(n)}, \eta_i^{(n)})$, $n = 1, 2, \ldots$, $i = 1, 2, \ldots$ are iid random vectors such that $\mathbf{P}((\xi, \eta) = (0,1)) = p_0$, $\mathbf{P}((\xi, \eta) = (1,0)) = p_1$, $\mathbf{P}((\xi, \eta) = (2,0)) = p_2$. The offspring mean matrix $\mathbf{M}$ has the form

$$\mathbf{M} = \begin{pmatrix} \mathbf{E}\xi & \mathbf{E}\eta \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} m & p_0 \\ 0 & 1 \end{pmatrix}.$$

Next we determine the mean vector of $\mathbf{X}_n$.

Lemma 1. *If $x_0 = 1$ then for the mean we have $\mathbf{E}X_n = m^n$, and $\mathbf{E}Y_n = p_0(1 + m + \ldots + m^{n-1})$, thus*

$$\mu_n := \mathbf{E}Z_{n,1} = \begin{cases} m^n \left(1 + \frac{p_0}{m-1}\right) - \frac{p_0}{m-1}, & m \neq 1, \\ 1 + p_0 n, & m = 1. \end{cases}$$

We note that the covariance matrix of $\mathbf{X}_n$ can be determined explicitly. The computation is straightforward but rather lengthy. Since we only need the explicit form of the mean and the finiteness of the second moments, we skip the computation.

The strong law of large numbers and the central limit theorem imply that for each fixed n as $x_0 \to \infty$

$$\frac{Z_{n,x_0}}{x_0} \longrightarrow \mu_n \quad \text{a.s.}$$

and
$$\frac{Z_{n,x_0} - x_0\mu_n}{\sqrt{x_0}} \xrightarrow{\mathcal{D}} N(0, \sigma_n^2), \tag{3}$$

where $\xrightarrow{\mathcal{D}}$ stands for convergence in distribution, and
$$\sigma_n^2 = \mathbf{Var}(Z_n).$$

It is clear that the geometric growth rate of $\mathbf{E}Z_n$ is the offspring mean m, while the precise distribution determines only the constant factor. Simple analysis shows that if $m = p_1 + 2p_2 > 1$ then
$$m^n \leq \mu_n = \frac{p_2 m^n - p_0}{m - 1} \leq \frac{m(m^n - 1)}{2(m - 1)} + 1, \tag{4}$$

if $m = 1$ then
$$1 \leq \mu_n = 1 + p_0 n \leq 1 + \frac{n}{2}, \tag{5}$$

while for $m < 1$
$$1 \leq \mu_n = \frac{p_0 - p_2 m^n}{1 - m} \leq \frac{m(1 - m^n)}{2(1 - m)} + 1. \tag{6}$$

The upper bound is attained at $(p_0, p_1, p_2) = (1 - m/2, 0, m/2)$, while the lower bound is attained at $(p_0, p_1, p_2) = (0, 2 - m, m - 1)$ for $m \geq 1$, and at $(p_0, p_1, p_2) = (1 - m, m, 0)$ for $m \leq 1$.

The process (X_n) is a single type Galton–Watson process with offspring mean $m = p_1 + 2p_2$. If $m \leq 1$ then the process dies out almost surely, while if the process is supercritical, i.e. $m > 1$ then the probability of extinction is the smaller root of $f(q) = q$, which is $q = p_0/p_2$. By the martingale convergence theorem
$$\frac{X_n}{m^n} \to W \quad \text{a.s.}, \tag{7}$$

where W is a nonnegative random variable. For $m \leq 1$ clearly $W \equiv 0$, while if $m > 1$ then $\mathbf{P}(W = 0) = q$, and the distribution of W is absolutely continuous on $(0, \infty)$.

The process $\mathbf{X}_n = (X_n, Y_n)$ is decomposable, because $(\mathbf{M}^n)_{2,1} = 0$ for any n. Limit theorems for supercritical decomposable processes were obtained by Kesten and Stigum [4]. The eigenvalues of $\mathbf{M}$ are m and 1, therefore the process is supercritical if and only if $m > 1$. Applying Theorem 2.1 by Kesten and Stigum [4] we obtain for $m > 1$ that
$$\lim_{n \to \infty} \frac{1}{m^n}(X_n, Y_n) = W\left(1, \frac{p_0}{m - 1}\right),$$

where W is the nonnegative random variable from (7).

3 Estimation of the offspring mean

Recall that the measurements are given in the form (2), where $Z_{n;c,x_0}^{(i)}$ stands for the total number of dead and alive bacteria at generation n, starting with x_0 bacteria under antibiotic concentration c at experiment i, $i = 1, 2, \ldots, N$. We assume that the sequence $\{\varepsilon_{i;c} : i \geq 1, c \geq 0\}$ are iid, independent of the process $\mathbf{X}_n$, and is Gaussian with mean 0 and variance σ_ε^2.

By (3)

$$\log_2 Z_{n;c,x_0}^{(i)} = \log_2(x_0 \mu_n) + \log_2 \left(1 + \frac{Z_{n;c,x_0}^{(i)} - x_0 \mu_n}{x_0 \mu_n} \right)$$

$$= \log_2(x_0 \mu_n) + \frac{1}{\sqrt{x_0} \log 2} \frac{\sigma_n}{\mu_n} \zeta_i + o_{\mathbf{P}}(x_0^{-1/2}),$$

where $(\zeta_i)_{i=1,\ldots,N}$ is a sequence of iid $N(0,1)$ random variables. This implies as $x_0 \to \infty$

$$C_i(c, x_0) = a - \log_2(x_0 \mu_n) + \varepsilon_{i;c} - \frac{1}{\log 2} \frac{\sigma_n}{\sqrt{x_0}} \frac{\sigma_n}{\mu_n} \zeta_i + o_{\mathbf{P}}(x_0^{-1/2}).$$

Put

$$\log_2 \widehat{\mu}_n = a - \log_2 x_0 - \frac{\sum_{i=1}^N C_i(c, x_0)}{N}.$$

In what follows $\xrightarrow{\mathbf{P}}$ stands for convergence in probability. By the law of large numbers and the central limit theorem, we have the following.

Proposition 1. *As $x_0 \to \infty$ and $N \to \infty$*

$$\log_2 \widehat{\mu}_n \xrightarrow{\mathbf{P}} \log_2 \mu_n,$$

which implies that $\widehat{\mu}_n$ is a weakly consistent estimator of μ_n. Furthermore,

$$\frac{1}{\sigma_\varepsilon} \sqrt{N} \left[\log_2 \widehat{\mu}_n - \log_2 \mu_n \right] \xrightarrow{\mathcal{D}} N(0, 1),$$

which implies that

$$\frac{1}{\sigma_\varepsilon \mu_n \log 2} \sqrt{N} \left(\widehat{\mu}_n - \mu_n \right) \xrightarrow{\mathcal{D}} N(0, 1).$$

We see that from the observations $Z_{n;c,x_0}^{(i)}$ we cannot estimate m itself, only μ_n. For $m \in (0, 2]$ fixed we obtained sharp bounds for μ_n in (4), (5),
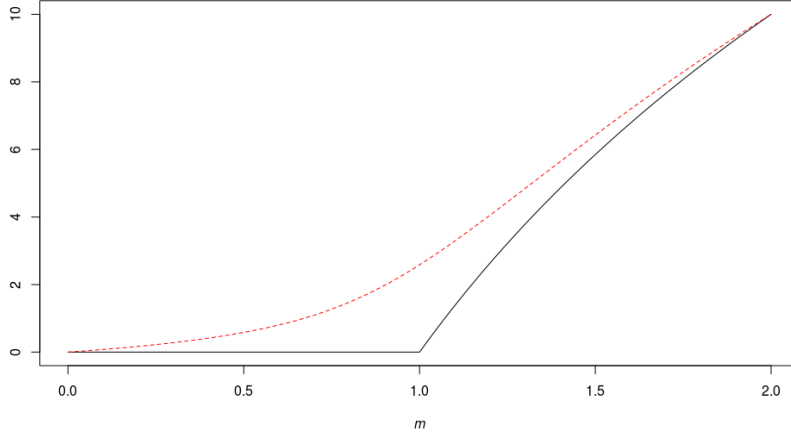
Figure 1: Upper and lower bound for $\log_2 \mu_n$ for $n = 10$.

(6). In Figure 1 we see the corresponding upper and lower bounds for $\log_2 \mu_n$ for $n = 10$. We see that the larger values for μ_n implies more precise bound for m. Furthermore, larger n also implies more precise bound. However, for $m \leq 1$ one cannot determine the value m. This is reasonable, since for both $p_0 = 1$ and $p_1 = 1$ we have $\mu_n = 1$, whereas $m = 0$ in the former and $m = 1$ in the latter case.

However, in real applications we may and do assume that $p_1 \equiv 0$. This is clearly reasonable for *bactericide* antibiotics, which either kill the bacteria, or let it duplicate. While, if a *bacteriostatic* antibiotic blocks the duplication of a single bacteria then it keeps blocking in the later generations as well. Therefore, we can equally count a 'blocked' bacteria as a dead one. Assume now that $p_1 \equiv 0$. Then μ_n is Lemma 1 simplifies to

$$\mu_n(m) = \frac{m}{2}\left(m^{n-1} + \ldots + 1\right) + 1 = \begin{cases} \frac{m(m^n - 1)}{2(m-1)} + 1, & m \neq 1, \\ \frac{n}{2} + 1, & m = 1. \end{cases}$$

Then μ_n is a strictly increasing convex function, $\mu_n(0) = 1$, $\mu_n(2) = 2^n$. Its inverse function $\psi_n : [1, 2^n] \to [0, 2]$ is continuous strictly increasing. Define the estimate

$$\widehat{m} = \psi_n(\widehat{\mu}_n).$$

From Proposition 1 it follows that $\widehat{m}$ is a weakly consistent estimator of m,

8

and

$$\psi_n(\widehat{\mu}_n) = \psi_n\left(\mu_n + \frac{\zeta\sigma_\varepsilon\mu_n\log 2}{\sqrt{N}} + o_\mathbf{P}(N^{-1/2})\right)$$
$$= \psi_n(\mu_n) + \psi_n'(\mu_n)\frac{\zeta\sigma_\varepsilon\mu_n\log 2}{\sqrt{N}} + o_\mathbf{P}(N^{-1/2})$$
$$= m + \psi_n'(\mu_n)\frac{\zeta\sigma_\varepsilon\mu_n\log 2}{\sqrt{N}} + o_\mathbf{P}(N^{-1/2}),$$

where $\zeta \sim N(0,1)$. Noting that $\psi_n'(\mu_n(m)) = 1/\mu_n'(m)$ we obtain the following.

Proposition 2. *Assume that $p_1 = 0$. As $x_0 \to \infty$ and $N \to \infty$, $\widehat{m}$ is a weakly consistent estimator of m, and*

$$\frac{\mu_n'(m)}{\sigma_\varepsilon\mu_n(m)\log 2}\sqrt{N}(\widehat{m} - m) \xrightarrow{\mathcal{D}} N(0,1).$$

4 The dependence of m on the antibiotic concentration

Assuming $p_1 \equiv 0$ we can estimate the mean for $c > 0$ fixed as described in Proposition 2. Next we combine our estimator for different concentrations.

We assume that the offspring mean as a function of c satisfies (1) for some unknown parameters $\alpha > 0$, $\beta > 0$. This is a quite flexible model, and we show that empirical data fits very well to this model. Rewriting (1)

$$\log\alpha + \beta\log c = \log\left(\frac{2}{m(c)} - 1\right).$$

Assume that we have measurements for $K \geq 2$ different concentrations $c_1 < c_2 < \ldots < c_K$, and we obtain the estimator for the offspring mean $\widehat{m}(c_i)$, $i = 1, 2, \ldots, K$. Using simple least squares estimator we obtain the estimates

$$\widehat{\beta} = \frac{K\sum_{i=1}^K f_i\ell_i - \sum_{i=1}^K f_iL_1}{KL_2 - L_1^2}$$
$$\widehat{\alpha} = \exp\left\{\frac{\sum_{i=1}^K f_i - \widehat{\beta}L_1}{K}\right\}, \tag{8}$$

where to ease notation we write

$$f_i = \log\left(\frac{2}{\widehat{m}(c_i)} - 1\right), \quad \ell_i = \log c_i,$$

9

and

$$L_1 = \sum_{i=1}^{K} \ell_i, \quad L_2 = \sum_{i=1}^{K} \ell_i^2. \tag{9}$$

Note that by the Cauchy–Schwarz inequality the denominator of $\widehat{\beta}$ is strictly positive for $K \geq 2$.

The *minimal inhibitory concentration* (MIC) is the smallest antibiotic concentration that stops bacteria growth. In mathematical terms

$$\vartheta := \mathrm{MIC} = \min\{c : m(c) \leq 1\},$$

which, under the assumption (1), $\vartheta = \mathrm{MIC} = \alpha^{-1/\beta}$. Define the estimator

$$\widehat{\vartheta} = \widehat{\alpha}^{-1/\widehat{\beta}}.$$

In the following statement we summarize the main properties of these estimators. Introduce the notation

$$k_i = -\frac{2}{m(c_i)(2 - m(c_i))} \frac{\sigma_\varepsilon \mu_n(m(c_i)) \log 2}{\mu_n'(m(c_i))}, \quad i = 1, 2, \ldots, K.$$

Proposition 3. *Assume that $x_0 \to \infty$ and $N \to \infty$. Then $\widehat{\alpha}, \widehat{\beta}$, and $\widehat{\vartheta}$ are weakly consistent estimators of the corresponding quantities. Furthermore,*

$$\sqrt{N}(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta) \xrightarrow{\mathcal{D}} (U, V),$$

where (U, V) is a two-dimensional normal random vector with mean 0 and covariance matrix

$$\begin{pmatrix} \sigma_\alpha^2 & \sigma_{\alpha\beta} \\ \sigma_{\alpha\beta} & \sigma_\beta^2 \end{pmatrix},$$

where

$$\sigma_\alpha^2 = \frac{\alpha^2}{\left(KL_2 - L_1^2\right)^2} \sum_{i=1}^{K} k_i^2 (L_2 - L_1\ell_i)^2$$

$$\sigma_{\alpha\beta} = \frac{\alpha}{\left(KL_2 - L_1^2\right)^2} \sum_{i=1}^{K} k_i^2 (K\ell_i - L_1)(L_2 - L_1\ell_i)$$

$$\sigma_\beta^2 = \frac{1}{\left(KL_2 - L_1^2\right)^2} \sum_{i=1}^{K} k_i^2 \left(K\ell_i - L_1\right)^2,$$

and

$$\sqrt{N}(\widehat{\vartheta} - \vartheta) \xrightarrow{\mathcal{D}} N(0, \sigma_\vartheta^2),$$

with

$$\sigma_\vartheta^2 = \frac{\vartheta^2 (\log \alpha)^2}{\beta^2 (KL_2 - L_1^2)^2} \sum_{i=1}^{K} k_i^2 \left(\frac{L_2 - L_1\ell_i}{\log \alpha} - \frac{K\ell_i - L_1}{\beta}\right)^2.$$

5 Simulation study

Regardless of the fixed values $\mathbf{c} = (c_1, \ldots, c_K)$ the estimator $(\widehat{\alpha}, \widehat{\beta})$ is weakly consistent and asymptotically normal as $N \to \infty$. However, the asymptotic variances in Proposition 3 do depend on the specific choice of $K \geq 2$ and the values $c_1 < \ldots < c_K$. Intuitively, it is clear that we should choose values for the concentrations c_i such that $m(c_i)$ is not close to 0, nor to 2.

Consider the following example. Assume that

$$\alpha = 10, \quad \beta = 1, \quad n = 10, \quad x_0 = 10^4, \quad \sigma_\varepsilon = 0.2. \tag{10}$$

It turns out that this is a reasonable choice, see the azithromycin data in the next section. The mean offspring function $m(c)$ is given on Figure 2.
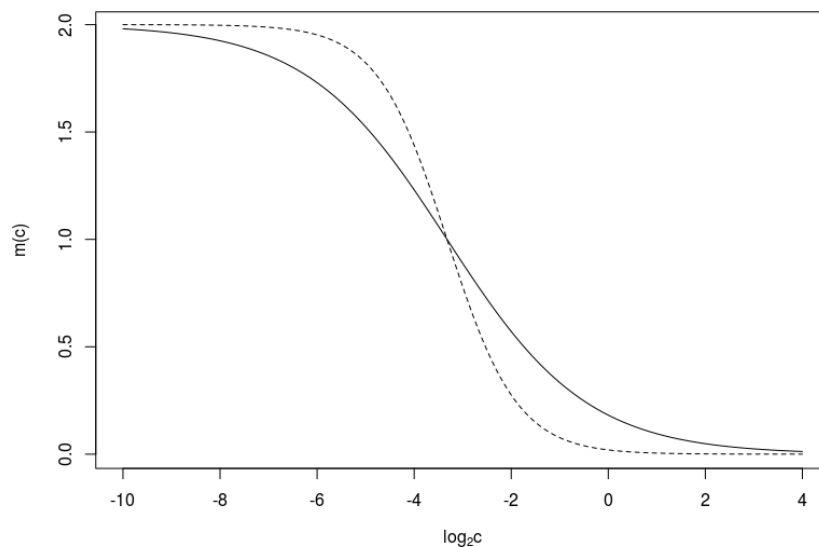

Figure 2: $m(c)$ in a logarithmic scale (solid $(\alpha, \beta) = (10, 1)$, dashed $(\alpha, \beta) = (100, 2)$).

Choose $K = 3$ different concentrations such that $\mathbf{c}_1 = (2^{-6}, 2^{-4}, 2^{-2})$. Then for the asymptotic covariances we obtain

$$\sigma_\alpha^2 = 8.63, \quad \sigma_{\alpha,\beta} = 0.25, \quad \sigma_\beta^2 = 0.00767, \quad \sigma_\vartheta^2 = 0.00012. \tag{11}$$

However, as we see in Table 1 wrong choice of the concentrations might results much larger variances. For $\mathbf{c}_2$ we only observe the process at large

concentrations, killing all the bacteria, while in case $\mathbf{c}_3$ the concentration is small, the antibiotic does not have any effect. The combination of large and small values as in $\mathbf{c}_4$ does not help either. Less obvious is the fact that choosing too many points is contraproductive too. This is the case for $\mathbf{c}_5$.

concentrations	σ_α^2	$\sigma_{\alpha,\beta}$	σ_β^2	σ_ϑ^2
$\mathbf{c}_1 = (2^{-6}, 2^{-4}, 2^{-2})$	8.63	0.25	0.00767	0.00012
$\mathbf{c}_2 = (2^{-2}, 2^{-1}, 1)$	112	9.41	0.833	0.012
$\mathbf{c}_3 = (2^{-9}, 2^{-8}, 2^{-7})$	967	18.7	0.364	0.0298
$\mathbf{c}_4 = (2^{-8}, 2^{-7}, 2^{-1}, 1)$	58	1.17	0.0257	0.00179
$\mathbf{c}_5 = (2^{-9}, 2^{-8}, \ldots, 1)$	23	0.568	0.0157	0.00051

Table 1: Asymptotic variances for different choices of $\mathbf{c}$ for $(\alpha, \beta) = (10, 1)$.

Choosing the values as in (10), $K = 3$ and $\mathbf{c}_1 = (2^{-6}, 2^{-4}, 2^{-2})$ we simulated the process as follows. For a given concentration c_k, $k = 1, \ldots, K$, we calculate $m(c_k)$ from (1), and choose the offspring distribution

$$p_{0;k} = 1 - \frac{m(c_k)}{2}, \quad p_{1;k} = 0, \quad p_{2;k} = \frac{m(c_k)}{2}.$$

With this offspring distribution we simulate $n = 10$ generations of the two-type Galton–Watson process (X_n, Y_n) described in Section 2. Therefore we obtain $Z_{10;c_k,x_0}$. Independently, we repeat the simulation N times for each concentration c_k. Independent of the Z's take an iid sequence of Gaussian random variables $\{\varepsilon_{i;c_k} : i = 1, \ldots, N; k = 1, \ldots, K\}$ with mean zero and variance σ_ε^2. Take $a = 0$ in (2). The resulting sequence $\{C_i(c_k, x_0) : i = 1, \ldots, N; k = 1, \ldots, K\}$ is one simulated measurement. From each measurement we calculate the estimator $(\widehat{\alpha}, \widehat{\beta})$ as described in (8). We simulated the measurement this way 1000 times. The resulting means and empirical variances of $\sqrt{N}(\widehat{\alpha} - \alpha, \widehat{\beta} - \beta)$ and $\sqrt{N}(\widehat{\vartheta} - \vartheta)$ are given in Table 2. We see that the empirical values are very close to the theoretical ones in (11) even for $N = 3, 10$. It is somewhat surprising that the estimates work even for $N = 3$, which is the suggested number of measurements at each concentration in microbiology (see e.g. [15, 2]).

Next we investigate our estimator with a steeper killing curve. Let $\alpha = 100$ and $\beta = 2$, and the other values as in (10). This is also a possible choice, see the ciprofloxacin data. In Figure 2 wee see the mean offspring function $m(c)$ for $(\alpha, \beta) = (10, 1)$ and for $(\alpha, \beta) = (100, 2)$. In the latter case there are less relevant concentrations, so we expect larger variances. In Table 3 we see that this is partly true, however the estimate of ϑ is good.

N	$\overline{\alpha}$	$\overline{\beta}$	$\overline{\vartheta}$	$\widehat{\sigma}^2_\alpha$	$\widehat{\sigma}_{\alpha,\beta}$	$\widehat{\sigma}^2_\beta$	$\widehat{\sigma}^2_\vartheta$
3	10.359	1.004	0.0998	12.95	0.325	0.00891	0.000121
10	10.106	1.002	0.1	9.27	0.262	0.00789	0.000116
50	10.03	1.0005	0.1	9.3	0.265	0.008	0.000124
100	9.999	0.9999	0.1	8.83	0.258	0.008	0.000117
∞	10	1	0.1	8.63	0.25	0.00767	0.00012

Table 2: Empirical mean and variances for $(\alpha, \beta) = (10, 1)$.

concentrations	σ^2_α	$\sigma_{\alpha,\beta}$	σ^2_β	σ^2_ϑ
$\mathbf{c}_1 = (2^{-6}, 2^{-4}, 2^{-2})$	11298	35.6	0.0124	0.000364
$\mathbf{c}_6 = (2^{-5}, 2^{-4}, 2^{-3})$	1431	5.49	0.0216	0.0000126
$\mathbf{c}_7 = (2^{-7}, 2^{-6}, \ldots, 2^{-1})$	42490	129.3	0.429	0.00142

Table 3: Asymptotics variances for different choices of $\mathbf{c}$ for $(\alpha, \beta) = (100, 2)$.

6   The experiment

In the experiment $50,000$ mother cells were infected by *Chlamydia trachomatis*. The multiplicity of infection (MOI) value, the ratio of the initial number of bacteria and number of mother cells is 0.2. That is $x_0 = 10,000$. The measurements correspond to 12 different antibiotic concentrations using two-fold dilution technique, meaning that $c_i = 2^i c_0$, $i = 0, 1, \ldots, 11$. For each concentration 3 measurements were done. For the technical details of the experiment we refer to [2].

We analyze two antibiotics: azithromycin and ciprofloxacin. These antibiotics have different antimicrobial effects: azithromycin is a bacteriostatic antibiotic, meaning that it does not necessarily kill the bacteria, only prevents growth, while ciprofloxacin is a bactericide antibiotic, which usually kills bacteria. In Figures 3 and 4 we see the qPCR measurements as a function of $\log_2 c$.

If c is large enough, i.e. at very high antibiotic concentration $m(c)$ is close to 0, that is $Z_{n;x_0,c} \approx x_0$, since all the bacteria dies without offspring. Therefore, for c large enough we can estimate the constant a in (2) as

$$\widehat{a}_N = \frac{1}{N} \sum_{i=1}^{N} C_i(c, x_0) + \log_2 x_0.$$

Then $\widehat{a}_N$ is normally distributed with mean a and variance σ^2_ε/N. Furthermore, σ_ε can also be estimated from these data. For azithromycin we used
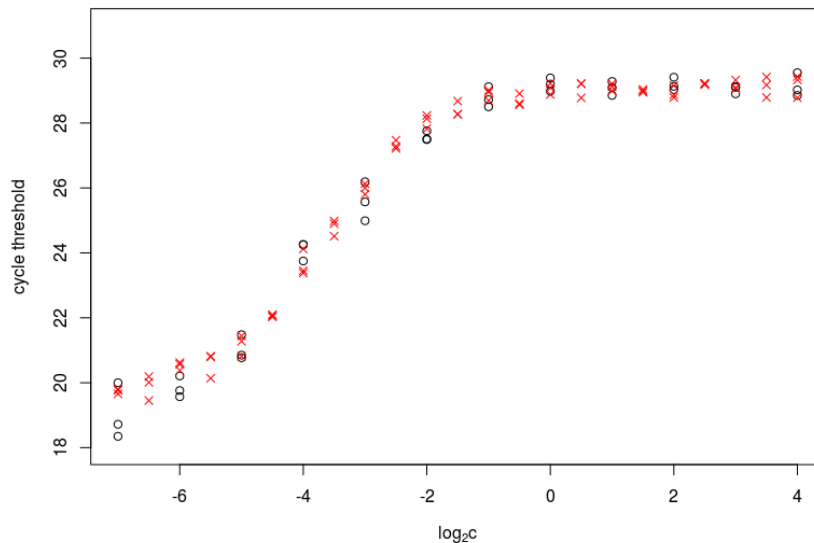
Figure 3: Measured ($\circ$) and simulated ($\times$) Ct values for azithromycin.

the measurements where $c \geq 2^{-1}$, while for ciprofloxacin $c \geq 1$.

The number of generations n is typically a fixed small number, in our experiments around 10. If c is small then there is no antibiotical effect so the bacterial population grows freely, that is $Z_{n,x_0,0} \approx 2^n x_0$. We can estimate n as

$$\widehat{n}_N = \widehat{a}_N - \log_2 x_0 - \frac{1}{N} \sum_{i=1}^{N} C_i(c, x_0).$$

Then $\widehat{n}_N$ is normally distributed with mean n and variance $2\sigma_\varepsilon^2/N$. To estimate $\widehat{n}_N$ we used the smallest possible concentration, $c = 2^{-7}$.

Using Proposition 2 we estimate $m(c)$. In Figures 5 and 6 we see the estimated means and the corresponding fitted curve $m(c)$, where the parameters α, β are estimated as described in (8). In the previous section we showed that the best strategy is to choose few concentration where the mean offspring is not close to 0, nor to 2. For the azithromycin we chose $\mathbf{c} = (2^{-5}, 2^{-4}, 2^{-2}, 2^{-1})$ and obtained $\widehat{\alpha} = 9.1$, $\widehat{\beta} = 1.12$, and $\widehat{\vartheta} = 0.139$. (We obtain similar estimates for various reasonable choices.) For ciprofloxacin in Figure 4 we see a rapid drastic change; for $c \geq 2^{-2}$ the population dies out, while for $c \leq 2^{-4}$ the population freely grows. We chose $\mathbf{c} = (2^{-4}, 2^{-3}, 2^{-2})$
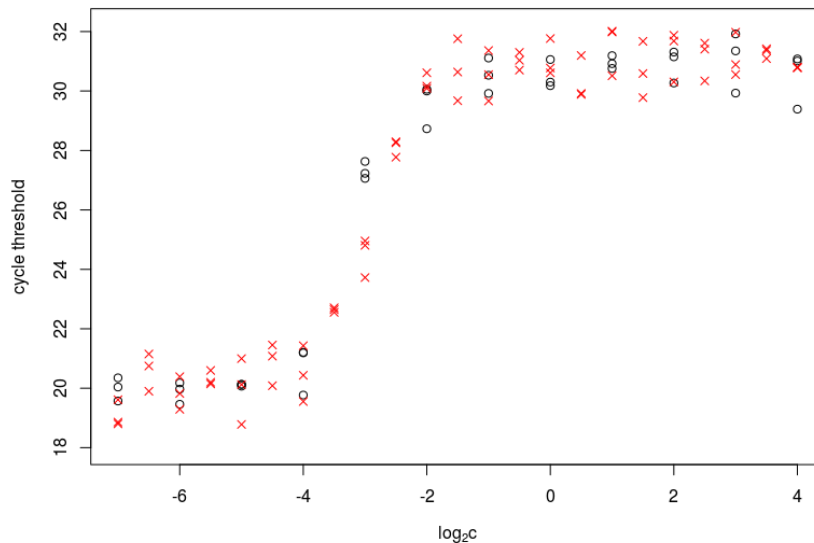
14

Figure 4: Measured ($\circ$) and simulated ($\times$) Ct values for ciprofloxacin.

and obtained $\widehat{\alpha} = 71.8$, $\widehat{\beta} = 2.46$, $\widehat{\vartheta} = 0.175$. (These values are less stable to the change in **c**.) Simulated measurements with the estimated values are given in Figures 3 and 4, where the circles are the real measurements and the crosses are the simulated ones. In both cases we obtain a remarkably good fit.

7   Conclusion

To model the growth of a bacterial population and its dependence on the antibiotic concentration we proposed a simple Galton–Watson model, where the offspring distribution depends on the antibiotic concentration via (1). A stochastic model is more natural compared to the previous deterministic model in [7], because we are able to estimate the parameters of the model and investigate the properties of the estimator. Taking into account the measurement error using qPCR technique, from the measurements at different antibiotic concentrations we obtained a weakly consistent asymptotically normal estimator for the unknown parameters (α, β) in (1).

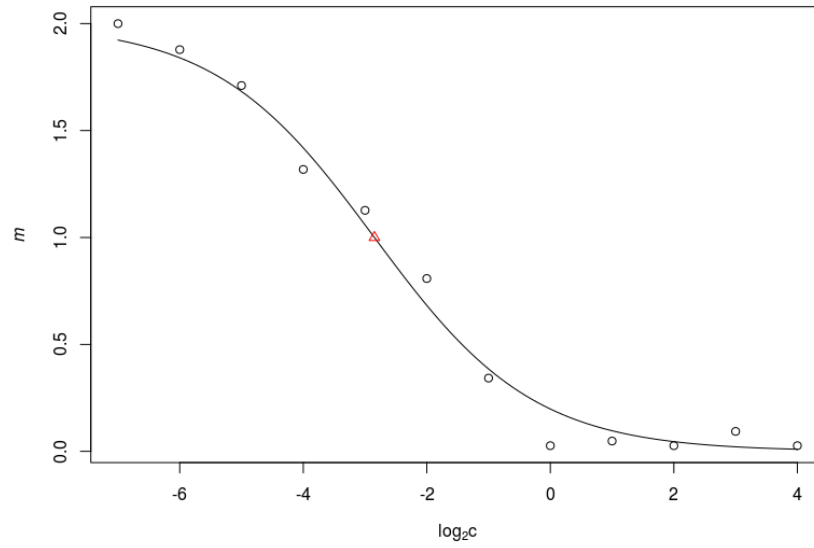The *minimal inhibitory concentration* (MIC), the smallest concentration

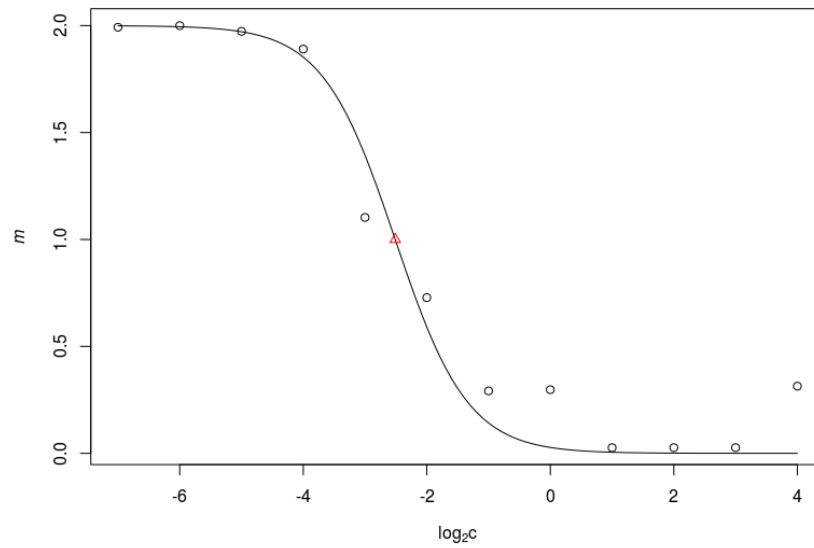Figure 5: Estimated means and the fitted curve for azithromycin.

Figure 6: Estimated means and the fitted curve for ciprofloxacin.

16

of antibiotic that prevents bacterial growth, is a very important parameter in pharmacology. Its estimation is rather troublesome, since due to the usual two-fold dilution technique one can observe only the bacterial growth under antibiotic concentration $c_0, 2c_0, \ldots, 2^k c_0$. Therefore one can claim only that the MIC belongs to some interval $[c, 2c]$, or give an upper bound for it. The vast majority of the literature does not provide a proper mathematical model for the growth of the bacterial population, only determines the MIC value as the smallest concentration without visible bacterial growth. In our framework an explicit mathematical definition of the MIC is given, and we constructed an estimator for it.

Simulation study showed that the estimators work surprisingly well even if the number of measurements at different concentration is 3, which is the suggested number in microbiology (see e.g. [15, 2]).

We applied the model to real measurements, where growth of *Chlamydia trachomatis* was analyzed treated by two different antibiotics. Although the mathematical model has only 2 parameters, we found extremely good fitting to the real data for both the bactericide and the bacteriostatic antibiotic.

Appendix

Proof of Lemma 1. Conditioning on $\mathbf{X}_n$

$$\mathbf{E}\left[\mathbf{X}_{n+1}|\mathbf{X}_n\right] = \begin{pmatrix} mX_n \\ p_0 X_n + Y_n \end{pmatrix} = \mathbf{X}_n \mathbf{M},$$

thus

$$\mathbf{E}\mathbf{X}_n = \mathbf{X}_0 \mathbf{M}^n.$$

We have, by induction on n that

$$\mathbf{M}^n = \begin{pmatrix} m^n & p_0(1 + \ldots + m^{n-1}) \\ 0 & 1 \end{pmatrix},$$

thus

$$\mathbf{E}Z_n = m^n + p_0(1 + m + \ldots + m^{n-1})$$
$$= \begin{cases} m^n \left(1 + \frac{p_0}{m-1}\right) - \frac{p_0}{m-1}, & \text{if } m \neq 1, \\ 1 + np_0, & \text{if } m = 1, \end{cases}$$

as claimed. $\qquad\qquad\square$

Proof of Proposition 3. Again by Proposition 2

$$f_i = \log\left(\frac{2}{m(c_i)} - 1\right) - \frac{2}{m(c_i)(2 - m(c_i))}\frac{\sigma_\varepsilon \mu_n(m(c_i))\log 2}{\sqrt{N}\mu'_n(m(c_i))}\zeta_i + o_{\mathbf{P}}(N^{-1/2})$$

$$= \log\left(\frac{2}{m(c_i)} - 1\right) + \frac{k_i}{\sqrt{N}}\zeta_i + o_{\mathbf{P}}(N^{-1/2}),$$

where ζ_i's are iid $N(0,1)$, $i = 1, 2, \ldots, K$. Recall the notation in (9). Then

$$\sum_{i=1}^{K} f_i\left(K\ell_i - L_1\right) = \sum_{i=1}^{K}\left[\log\left(\frac{2}{m(c_i)} - 1\right) + \frac{k_i\zeta_i}{\sqrt{N}}\right](K\ell_i - L_1) + o_{\mathbf{P}}(N^{-1/2}).$$

Substituting back into (8)

$$\widehat{\beta} - \beta = \frac{1}{\sqrt{N}}\sum_{i=1}^{K}\zeta_i\, k_i\frac{K\ell_i - L_1}{KL_2 - L_1^2} + o_{\mathbf{P}}(N^{-1/2}), \tag{12}$$

and similarly

$$\log\widehat{\alpha} - \log\alpha = \frac{1}{K\sqrt{N}}\sum_{i=1}^{K}\zeta_i\, k_i\left(1 - \frac{L_1(K\ell_i - L_1)}{KL_2 - L_1^2}\right) + o_{\mathbf{P}}(N^{-1/2})$$

$$= \frac{1}{\sqrt{N}}\sum_{i=1}^{K}\zeta_i\, k_i\frac{L_2 - L_1\ell_i}{KL_2 - L_1^2} + o_{\mathbf{P}}(N^{-1/2}), \tag{13}$$

which implies

$$\widehat{\alpha} - \alpha = \frac{\alpha}{\sqrt{N}}\sum_{i=1}^{K}\zeta_i\, k_i\frac{L_2 - L_1\ell_i}{KL_2 - L_1^2} + o_{\mathbf{P}}(N^{-1/2}).$$

From (12) and (13) we obtain

$$\widehat{\vartheta} - \vartheta = -\frac{\vartheta\log\alpha}{\sqrt{N}\beta(KL_2 - L_1^2)}\sum_{i=1}^{K}\zeta_i k_i\left(\frac{L_2 - L_1\ell_i}{\log\alpha} - \frac{K\ell_i - L_1}{\beta}\right).$$

$\square$

References

[1] G. Enciso, C. Sütterlin, M. Tan, and F. Wan. Stochastic chlamydia dynamics and optimal spread. *Bull Math Biol*, 83(24), 2021.

[2] I. Eszik, I. Lantos, K. Önder, F. Somogyvári, K. Burián, V. Endrész, and D.P. Virok. High dynamic range detection of Chlamydia trachomatis growth by direct quantitative PCR of the infected cells. *J. Microbiol. Methods*, 120:15–22, 2016.

[3] Patsy Haccou, Peter Jagers, and Vladimir A. Vatutin. *Branching processes: variation, growth, and extinction of populations*, volume 5 of *Cambridge Studies in Adaptive Dynamics*. Cambridge University Press, Cambridge; IIASA, Laxenburg, 2007.

[4] H. Kesten and B. P. Stigum. Limit theorems for decomposable multidimensional Galton-Watson processes. *J. Math. Anal. Appl.*, 17:309–338, 1967.

[5] Marek Kimmel and David E. Axelrod. *Branching processes in biology*, volume 19 of *Interdisciplinary Applied Mathematics*. Springer, New York, second edition, 2015.

[6] J.K. Lee, G.A. Enciso, D. Boassa, C.N. Chander, T.H. Lou, Pairawan S.S., M.C. Guo, Wan F.Y.M., M.H. Ellisman, C. Sütterlin, and M. Tan. Replication-dependent size reduction precedes differentiation in chlamydia trachomatis. *Nat Commun.*, 45(9):3884–3891, 2018.

[7] Y. Q. Liu, Y. Z. Zhang, and P. J. Gao. Novel concentration-killing curve method for estimation of bactericidal potency of antibiotics in an in vitro dynamic model. *Antimicrob. Agents Chemother.*, 48(10):3884–3891, 2004.

[8] Karl E. Miller. Diagnosis and treatment of Chlamydia trachomatis infection. *Am Fam Physician*, 8(73):1411–1416, 2006.

[9] M.E. Panzetta, R.H. Valdivia, and H.A. Saka. Chlamydia persistence: A survival strategy to evade antimicrobial effects in-vitro and in-vivo. *Front Microbiol.*, 9(3101), 2018.

[10] I.K. Paterson, A. Hoyle, G. Ochoa, C. Baker-Austin, and N.G.H. Taylor. Optimising antibiotic usage to treat bacterial infections. *Sci. Rep.*, 6(37853), 2016.

[11] Fabian Svara and Daniel J. Rankin. The evolution of plasmid-carried antibiotic resistance. *BMC Evolutionary Biology*, 11(130), 2011.

[12] Vu T.H., N.G. Ha-Doung, A. Aubry, E. Capton, P. Fechter, P. Plésiat, P. Verbeke, and N. Serradji. In vitro activities of a new fluoroquinolone derivative highly active against Chlamydia trachomatis. *Bioorg Chem*, (83):180–185, 2019.

[13] Frederic Y. M. Wan and Germán A. Enciso. Optimal proliferation and differentiation of *chlamydia trachomatis*. *Stud. Appl. Math.*, 139(1):129–178, 2017.

[14] D. P. Wilson. Mathematical modelling of chlamydia. In *Proc. of 11th Computational Techniques and Applications Conference CTAC-2003*, volume 45, pages C201–C214, 2004.

[15] J. S. Yuan, A. Reed, F. Chen, and C. N. Stewart. Statistical analysis of real-time PCR data. *BMC Bioinformatics*, 7(85), 2006.