

DETERMINISTIC AND STOCHASTIC MODELS FOR THE SPREAD OF CHOLERA

J. GANI¹ and R. J. SWIFT²✉

(Received 17 September, 2008; revised 5 February, 2010)

Abstract

In this note, we study deterministic and stochastic models for the spread of cholera. The deterministic model for the total number of cholera cases fits the observed total number of cholera cases in some recent outbreaks. The stochastic model for the total number of cholera cases leads to a binomial type distribution with a mean that agrees with the deterministic model.

2010 *Mathematics subject classification*: primary 92D30; secondary 60J27.

Keywords and phrases: cholera, water-borne bacterial infections, deterministic model, stochastic model.

1. Introduction

Cholera is a disease caused by the bacterium *Vibrio cholerae* often found in water contaminated by faecal matter. It has a short incubation period, from less than one day to five days, and produces an enterotoxin that causes copious watery diarrhoea in humans, quickly leading to severe dehydration and death if treatment is not promptly given.

According to the World Health Organization (WHO) website, cholera persists as a global threat and continues to challenge countries where access to safe drinking water and adequate sanitation is not assured. Almost every developing country faces cholera outbreaks or the threat of a cholera epidemic.

It was during the 1854 cholera epidemic in London that British physician John Snow (1813–1858) first determined the cause of cholera. He showed [4] that a nearby sewer had contaminated water from the Broad Street pump; once this had been put out of action, the epidemic subsided.

In this paper we develop and discuss deterministic and stochastic models for the spread of cholera. The models developed are applicable to other water-borne

¹Mathematical Sciences Institute, Australian National University, Canberra ACT 0200, Australia; e-mail: gani@maths.anu.edu.au.

²Department of Mathematics and Statistics, California State Polytechnic University, Pomona, CA 91768, USA; e-mail: rjswift@csupomona.edu.

© Australian Mathematical Society 2010, Serial-fee code 1446-1811/2010 \$16.00

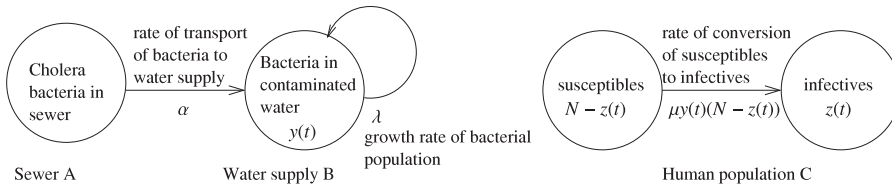


FIGURE 1. A diagrammatic representation of the spread of cholera.

gastrointestinal bacterial infections such as dysentery due to *Shigella* bacteria and typhoid due to *Salmonella typhi* bacteria. Deterministic models for the spread of cholera have also been considered by Das *et al.* [1], Pourabbas *et al.* [2] and Singh and Singh [3].

2. A deterministic model

Let us assume that cholera bacteria in a sewer *A* escape at a Poisson rate α into the water supply *B*, where $y(t)$, $t \geq 0$, bacteria multiply at the birth rate λ , with some of them infecting a susceptible human population *C* of size $N - z(t)$ at the rate $\mu < \lambda$, $z(t)$ being the number of infectives with cholera.

Figure 1 is a diagrammatical representation of this process and gives the deterministic equations

$$\frac{dy(t)}{dt} = \alpha + \lambda y(t) - \mu y(t) \quad \text{and} \quad \frac{dz(t)}{dt} = \mu y(t)(N - z(t)), \quad (2.1)$$

with initial conditions $y(0) = 0$ and $z(0) = 0$. We readily see that

$$y(t) = \frac{\alpha}{\lambda - \mu} \left(e^{(\lambda - \mu)t} - 1 \right), \quad (2.2)$$

$$z(t) = N \left[1 - \exp \left(-\frac{-\alpha \mu}{(\lambda - \mu)^2} \left(\left(e^{(\lambda - \mu)t} - 1 \right) - (\lambda - \mu)t \right) \right) \right].$$

Since $e^{(\lambda - \mu)t} > 1 + (\lambda - \mu)t$, the number of cholera infectives $z(t)$ is always less than or equal to N .

The WHO’s Global Task Force on cholera control was launched in 1992 with the aim of reducing mortality and morbidity associated with the disease and addressing the social and economic consequences of cholera. The Task Force provides technical advice and support for cholera control and prevention; preparedness and response to diarrhoeal disease outbreaks; and the dissemination of information on cholera and other epidemic prone enteric diseases. The task force’s report series *Cholera Country Profile*, available on the WHO website, often contains data on cholera cases reported on a weekly basis.

The points in Figures 2(a) and (b) show the total number of cholera cases during outbreaks in Sierra Leone (20 August to 11 November 2006) and in N’Zerekore

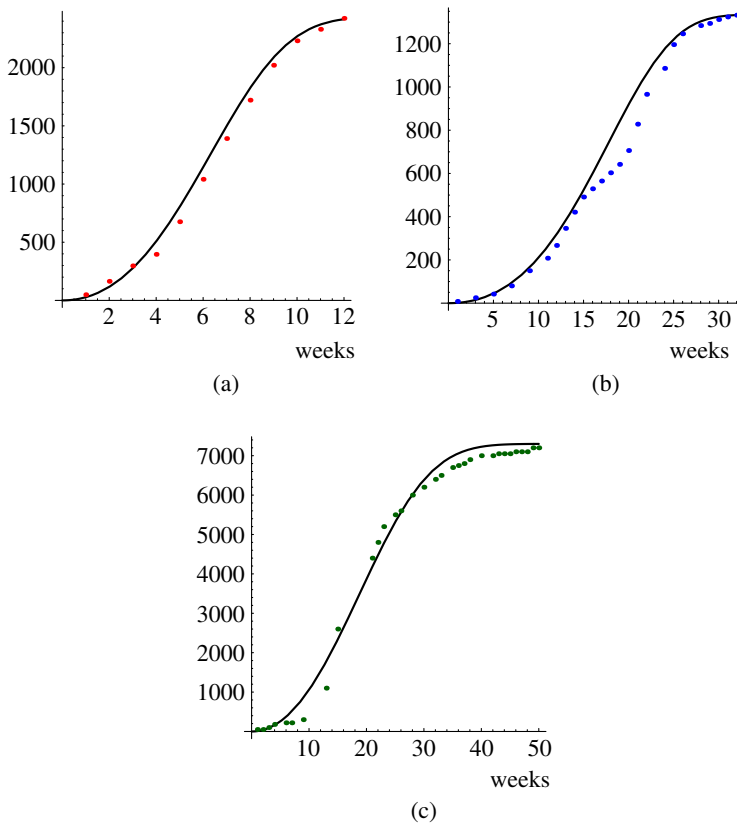


FIGURE 2. Total number of cholera cases shown as points and model predictions of $z(t)$ shown with lines for (a) Sierra Leone data, $N = 2435$, $\mu = 0.65$, $\lambda = 3$ and $\alpha = 3.2$, (b) Guinea data, $N = 1335$, $\mu = 0.4$, $\lambda = 1.6$ and $\alpha = 0.55$, (c) El Salvador data, $N = 7300$, $\mu = 0.35$, $\lambda = 0.85$ and $\alpha = 0.75$.

province in Guinea (1 January to 17 December 2006). Similar data sets are available for other countries. An additional set of data can be obtained from Pourabbas *et al.* [2] who present a method of estimating the seasonal variation of infection rate in cholera cases using empirical methods. The total number of cholera cases in El Salvador from 9 September 1993 to 2 September 1994 is shown with points in Figure 2(c). The solid curves in Figure 2 also show the predictions from deterministic model (2.2) for the total number of cholera cases $z(t)$ in Sierra Leone, Guinea and El Salvador. The deterministic model (2.2) for $z(t)$ provides a good representation of the behaviour shown in these data sets.

3. A stochastic equivalent

Let $Y(t)$ and $Z(t)$ be the random numbers of cholera bacteria in contaminated water, and cholera infectives at time $t \geq 0$ respectively, where $Y(0) = Z(0) = 0$. We define

the probabilities

$$P\{Y(t) = i, Z(t) = j \mid Y(0) = 0, Z(0) = 0\} = p_{ij}(t), \quad 0 \leq i < \infty, \quad 0 \leq j \leq N,$$

of the infection process, which satisfy the forward Kolmogorov equations

$$\begin{aligned} \frac{dp_{ij}(t)}{dt} = & -(\alpha + (\lambda + N\mu)i - ij\mu)p_{ij}(t) + (\alpha + (i-1)\lambda)p_{i-1j}(t) \\ & + (i+1)(N-j+1)\mu p_{i+1j-1}(t). \end{aligned} \quad (3.1)$$

Denoting the probability generating function (p.g.f.) of the distribution by

$$\phi(u, v, t) = \sum_{\substack{0 \leq j \leq N \\ 0 \leq i < \infty}} p_{ij}(t) u^i v^j, \quad 0 \leq u, v \leq 1,$$

we find from (3.1) that

$$\begin{aligned} \frac{\partial \phi(u, v, t)}{\partial t} = & \mu v(u-v) \frac{\partial^2 \phi(u, v, t)}{\partial u \partial v} + (\lambda u^2 - (\lambda + N\mu)u + \mu Nv) \frac{\partial \phi(u, v, t)}{\partial u} \\ & + \alpha(u-1)\phi(u, v, t). \end{aligned} \quad (3.2)$$

To simplify the partial differential equation (PDE) (3.1), we may replace the value $y(t) = i$ in (3.1) by the deterministic mean

$$y(t) = \frac{\alpha}{\lambda - \mu} \left(e^{(\lambda - \mu)t} - 1 \right).$$

This leads to

$$\begin{aligned} \frac{dp_j(t)}{dt} = & -(\alpha + (\lambda + N\mu)y(t) - y(t)j\mu)p_j(t) + (\alpha + (y(t)-1)\lambda)p_j(t) \\ & + (y(t)+1)(N-j+1)\mu p_{j-1}(t), \end{aligned} \quad (3.3)$$

where $p_j(t) = P\{Z(t) = j \mid Z(0) = 0\}$.

We may approximate (3.3) for large values of $y(t)$ by replacing $y(t) \pm 1$ by $y(t)$ to give

$$\begin{aligned} \frac{dp_j(t)}{dt} = & -(\alpha + \lambda y(t) + (N-j)\mu y(t))p_j(t) + (\alpha + y(t)\lambda)p_j(t) \\ & + y(t)(N-j+1)\mu p_{j-1}(t) \\ = & -y(t)(N-j)\mu p_j(t) + \mu(N-(j-1))y(t)p_{j-1}(t). \end{aligned}$$

The associated p.g.f. $\psi(v, t) = \sum_{j=0}^N p_j(t)v^j$ will then satisfy the PDE

$$\frac{\partial \psi(v, t)}{\partial t} + \mu y(t)v(v-1) \frac{\partial \psi(v, t)}{\partial v} = \mu N y(t)(v-1)\psi(v, t). \quad (3.4)$$

The auxiliary equations are

$$\frac{dt}{1} = \frac{dv}{\mu y v (v-1)} = \frac{d\psi}{N \mu y (v-1) \psi}. \quad (3.5)$$

Solving the first equality of (3.5) for v , we find that

$$\begin{aligned} \mu M(t) &= \mu \int_0^t y(\tau) d\tau = \mu \int_0^t \left[\frac{\alpha}{\lambda - \mu} \left(e^{(\lambda - \mu)\tau} - 1 \right) \right] d\tau \\ &= \frac{\mu \alpha}{(\lambda - \mu)^2} \left(e^{(\lambda - \mu)t} - 1 - (\lambda - \mu)t \right) \\ &= \int \left(-\frac{1}{v} + \frac{1}{v-1} \right) dv = \ln \left(\frac{v-1}{v} \right) + \text{constant} \end{aligned} \quad (3.6)$$

so that

$$\left(\frac{v-1}{v} \right) e^{-\mu M(t)} = \text{constant}.$$

Solving (3.5) for ψ ,

$$\frac{d\psi}{\psi} = N \frac{dv}{v} \quad \text{or} \quad \ln \psi = \ln v^N + \text{constant}.$$

Thus

$$\psi(v, t) v^{-N} = f \left(\left(\frac{v-1}{v} \right) e^{-\mu M(t)} \right), \quad (3.7)$$

where the function f is unknown.

Now for $t = 0$, $\psi(v, 0) = 1$, so that $v^{-N} = f((v-1)/v)$; or setting $w = (v-1)/v$ so that $v = 1/(1-w)$, we have $f(w) = (1-w)^N$. Hence

$$\begin{aligned} f \left(\left(\frac{v-1}{v} \right) e^{-\mu M(t)} \right) &= \left(1 - \left(\frac{v-1}{v} \right) e^{-\mu M(t)} \right)^N \\ &= v^{-N} \left(v \left(1 - e^{-\mu M(t)} \right) + e^{-\mu M(t)} \right)^N \end{aligned}$$

and

$$\psi(v, t) = \left(v \left(1 - e^{-\mu M(t)} \right) + e^{-\mu M(t)} \right)^N \quad (3.8)$$

so that $Z(t)$ has a binomial distribution with the p.g.f (3.8) satisfying the conditions

$$\psi(v, 0) = 1 \quad \text{and} \quad \psi(v, \infty) = v^N.$$

Note that the expectation of this process is $E(Z(t)) = N(1 - e^{-\mu M(t)})$ which is precisely the deterministic $z(t)$ of (2.2).

Using the values $\mu = 0.65$, $\lambda = 3$, $\alpha = 3.2$ and $N = 2435$ from the data on the 2006 Sierra Leone cholera outbreak, a graph of some of the transient probabilities

$$p_j(t) = \binom{N}{j} \left(1 - e^{-\mu M(t)} \right)^j e^{-\mu(N-j)M(t)}$$

is shown in Figure 3.

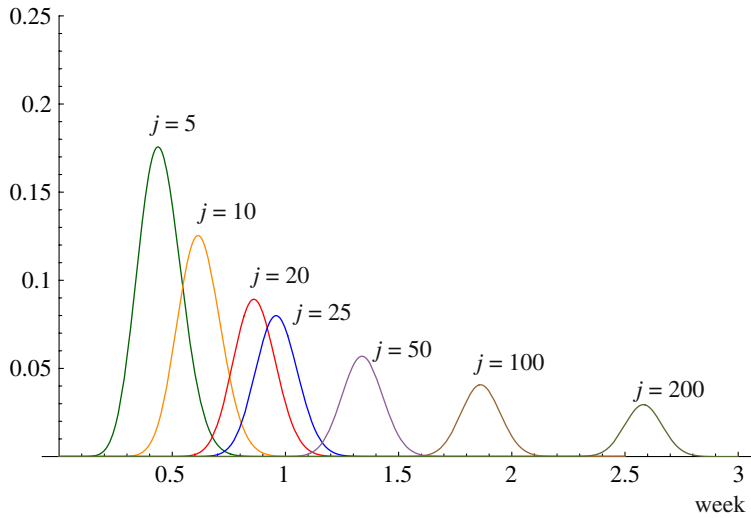


FIGURE 3. A graph of the probabilities $p_j(t) = P\{Z(t) = j \mid Z(0) = 0\}$ for the total number of cholera cases $Z(t)$ in the 2006 Sierra Leone outbreak.

If we assume that the survival rate from cholera is $0 < p < 1$, then

$$\begin{aligned}
 P\{Z(t) = k\} &= \sum_{j=k}^N \binom{N}{j} (1 - e^{-\mu M(t)})^j (e^{-\mu M(t)})^{N-j} \binom{j}{k} p^k q^{j-k} \\
 &= \sum_{j=k}^N \frac{N! e^{-\mu N M(t)}}{(N-j)!(j-k)!k!} \left[(e^{\mu M(t)} - 1)q \right]^j \left(\frac{p}{q} \right)^k \tag{3.9}
 \end{aligned}$$

with p.g.f.

$$\chi(v, t) = \sum_{k=0}^N v^k \sum_{j=k}^N \binom{N}{j} \binom{j}{k} \left((1 - e^{-\mu M(t)})q \right)^j e^{-\mu M(t)(N-j)} \left(\frac{p}{q} \right)^k. \tag{3.10}$$

After a change in the order of summation in (3.10),

$$\begin{aligned}
 \chi(v, t) &= \sum_{j=0}^N \binom{N}{j} \left((1 - e^{-\mu M(t)})q \right)^j e^{-\mu M(t)(N-j)} \sum_{k=0}^j \binom{j}{k} \left(\frac{pv}{q} \right)^k \\
 &= \sum_{j=0}^N \binom{N}{j} \left((1 - e^{-\mu M(t)})q \right)^j e^{-\mu M(t)(N-j)} \left(1 + \frac{pv}{q} \right)^j \\
 &= \sum_{j=0}^N \binom{N}{j} (1 - e^{-\mu M(t)})^j e^{-\mu M(t)(N-j)} (q + pv)^j \\
 &= \left[(1 - e^{-\mu M(t)})(q + pv) + e^{-\mu M(t)} \right]^N \tag{3.11}
 \end{aligned}$$

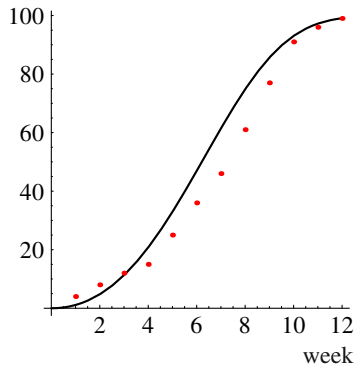


FIGURE 4. A graph of the expected total number of deaths predicted by Equation (3.12) (line) compared to the actual number of deaths (points) in the 2006 Sierra Leone cholera outbreak, with $\mu = 0.65$, $\lambda = 3$, $\alpha = 3.2$ and $N = 2435$.

which also follows directly from (3.8). The expected number of fatalities in a cholera outbreak is

$$qE(Z(t)) = qN\left(1 - e^{-\mu M(t)}\right). \quad (3.12)$$

In the 2006 Sierra Leone cholera outbreak, there were 99 reported fatalities, so that the approximate probability of death q is about 0.041. Figure 4 shows the actual number of fatalities of this outbreak, together with the number predicted by Equation (3.12). Although the curve fits the data reasonably, it seems inadequate between weeks four and nine, possibly because the death rate varies rather than remaining constant at 0.041.

4. Concluding remarks

In this note, we have considered simple deterministic and stochastic models for the spread of cholera. We describe the spread of cholera by modeling the bacterial population in contaminated water and the human interaction with the bacteria in the water supply. Although the models are relatively simple, they reflect reasonably the behaviour of the observed total number of cases. The value from the deterministic model for the total number of cholera cases fits the observed data in three recent outbreaks. The stochastic model leads to a binomial type distribution for the number of cholera cases, with a mean that agrees with the deterministic model. We hope that these models will prove useful to health agencies.

References

- [1] P. Das, D. Mukherjee and A. K. Sarkar, “Study of a carrier dependent infectious disease—cholera”, *J. Biol. Syst.* **13** (2005) 233–244.
- [2] E. Pourabbas, A. d’Onofrio and M. Rafanelli, “A method to estimate the incidence of communicable diseases under seasonal fluctuations with application to cholera”, *Appl. Math. Comput.* **118** (2001) 161–174.
- [3] L. S. Singh and K. Singh, “SIS epidemic model of cholera”, *Vikram Math. J.* **18** (1998) 56–64.
- [4] J. Snow, “The cholera near Golden Square, and at Deptford”, *Medical Times and Gazette* **9** (1854) 321–322.