Herald waves of cholera in nineteenth century London

Joseph H. Tien1,*, Hendrik N. Poinar2,3, David N. Fisman5 and David J. D. Earn3,4

1 Department of Mathematics, Ohio State University, Columbus, OH 43210, USA
2 Departments of Anthropology, Biology and the Ancient DNA Centre, 3 M.G. DeGroote Institute for Infectious Disease Research, and 4 Department of Mathematics and Statistics, McMaster University, Hamilton, Ontario L8S 4K1, Canada
5 Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario M5T 3M7, Canada

Deaths from cholera in London, UK, were recorded weekly from 1824 to 1901. Three features of the time series stand out: (i) cholera deaths were strongly seasonal, with peak mortality almost always in the summer, (ii) the only non-summer outbreaks occurred in the spring of 1832, the autumn of 1848 and the winter of 1853, and (iii) extraordinarily severe summer outbreaks occurred in 1832, 1849, 1854 and 1866 (the four ‘great’ cholera years). The non-summer outbreaks of 1832, 1848 and 1853 appear to have been herald waves of newly invading cholera strains. In addition, a simple mathematical model confirms that a non-summer introduction of a new cholera strain can result in an initial herald wave, followed by a severe outbreak the following summer. Through the analysis of the genomes of nineteenth-century specimens, it may be possible to identify the strains that caused these herald waves and the well-known cholera epidemics that followed.

Keywords: London cholera; herald wave; waterborne disease model; disease seasonality; John Snow
evolution of new cholera strains appears to be facilitated by horizontal gene transfer among strains in different serogroups ([17]; see the electronic supplementary material). This mechanism might also account for the invasion of antigenically novel strains in the nineteenth century. Cyclical replacement of the predominant cholera serotype has also been observed [18], and may be relevant for London cholera.

Figure 1. Weekly London cholera deaths from 1824 to 1901. (a) Weekly cholera deaths versus time and (b) cholera deaths versus time of year.

Figure 2. Weekly cholera deaths, normalized by all-cause mortality and plotted against time of year. Open circles correspond to Bills of Mortality data (years prior to 1842), filled circles to Registrar General data (1842 onwards). Filled light blue circles, 1832 herald; solid light blue line, 1832 summer; filled red circles, 1848 herald; solid red line, 1849 summer; filled dark blue circles, 1853 herald; solid dark blue line, 1854 summer; solid green line, 1866 summer.
We used a simple mathematical model to verify theoretically that a non-summer introduction of a new cholera strain can result in a herald wave followed by a severe outbreak in the summer. The model extends the classical ‘susceptible–infectious–recovered’ (SIR) framework to include a water compartment (W), with transmission occurring through both person–person and person–water–person pathways ([19]; see appendix A for model equations).

As an example, figure 3 compares the London cholera mortality time series from 1845 to 1852 (figure 3a) with an SIWR simulation (figure 3b). Assuming the transmission rate in the water varies seasonally (peaking in the summer, as in figure 3c), simulated cholera time series resemble the typical pattern in London, with mild annual summer outbreaks (as in the period 1845–1848 in figure 3b). The arrow in figure 3b indicates the time at which we introduced a new cholera strain into the SIWR model (the end of September 1848). The population is completely susceptible to the new strain, resulting in an initial outbreak when the strain is introduced. However, owing to the season [10], disease transmissibility is waning at the time of introduction. Eventually transmission decreases past the point where an outbreak can be sustained (and the initial herald wave terminates), but the new cholera strain persists at low levels in the water (figure 3d).

When transmissibility from the water rises again the following summer, it triggers an unusually severe epidemic owing to the large number of susceptible individuals (simulation details given in appendix A).

We used a simple mathematical model to verify theoretically that a non-summer introduction of a new cholera strain can result in a herald wave followed by a severe outbreak in the summer. The model extends the classical ‘susceptible–infectious–recovered’ (SIR) framework to include a water compartment (W), with transmission occurring through both person–person and person–water–person pathways ([19]; see appendix A for model equations).

As an example, figure 3 compares the London cholera mortality time series from 1845 to 1852 (figure 3a) with an SIWR simulation (figure 3b). Assuming the transmission rate in the water varies seasonally (peaking in the summer, as in figure 3c), simulated cholera time series resemble the typical pattern in London, with mild annual summer outbreaks (as in the period 1845–1848 in figure 3b). The arrow in figure 3b indicates the time at which we introduced a new cholera strain into the SIWR model (the end of September 1848). The population is completely susceptible to the new strain, resulting in an initial outbreak when the strain is introduced. However, owing to the season [10], disease transmissibility is waning at the time of introduction. Eventually transmission decreases past the point where an outbreak can be sustained (and the initial herald wave terminates), but the new cholera strain persists at low levels in the water (figure 3d).

Table 1. Variables for system (A 1), together with initial conditions used for simulations of endemic and introduced London cholera (figure 3).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Initial Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>susceptible individuals</td>
</tr>
<tr>
<td>I</td>
<td>infected individuals</td>
</tr>
<tr>
<td>R</td>
<td>recovered individuals</td>
</tr>
<tr>
<td>W</td>
<td>pathogen concentration in water reservoir (cells ml⁻¹)</td>
</tr>
<tr>
<td>D</td>
<td>individuals who have died from the disease</td>
</tr>
<tr>
<td>N</td>
<td>total population size</td>
</tr>
</tbody>
</table>

When transmissibility from the water rises again the following summer, it triggers an unusually severe epidemic owing to the large number of susceptible individuals (simulation details given in appendix A).

While the strain invasion hypothesis is simple—and appealing from the point of view of parsimony—many other explanations are possible, including misdiagnosis involving other diarrhoeal diseases. A direct test of the strain invasion hypothesis would require genetic information from the circulating cholera strains. Relevant
tissue samples from patients who died of cholera do exist in museum collections, and recent advances in the recovery and sequencing of DNA as well as the reconstruction of complete genomes from fossil materials [20, 21] make sequencing substantial portions of the genomes of nineteenth-century cholera strains a realistic goal (see the electronic supplementary material).

Given the fame and historical importance of the four major London cholera epidemics in the nineteenth century, it is surprising that the herald waves we have identified here have not been highlighted previously. Unravelling the mechanisms behind these herald waves will deepen our understanding of the evolutionary and ecological history of this important disease, and in turn help us understand the factors underlying severe cholera outbreaks in modern times. Our study of London cholera also suggests that herald waves may occur for more diseases than has been previously realized, and emphasizes the need for further work examining the relationship between the timing and magnitude of seasonal outbreaks [10]. The systematic digitization of lengthy historical records of disease incidence and mortality will be invaluable for this endeavour. In particular, evidence that herald waves have preceded major epidemics of other diseases may be hidden in untapped historical sources.

Many people deserve thanks for contributing to the acquisition and digitization of London’s weekly mortality records, especially Kelly Hancock, Susan Marsh-Rollo, James McDonald and David Richardson. We thank Alison Devault for valuable discussions regarding cholera evolution and archival cholera samples. This project was funded primarily by a grant to D.J.D.E. from the J. S. McDonnell Foundation. D.J.D.E. and H.N.P. also thank NSERC and CIHR.

APPENDIX A. MATHEMATICAL MODEL

A.1. Model equations

Our SIWR model for waterborne disease modelling is expressed as a simple system of ordinary differential equations [19],

\[
\begin{align*}
    \dot{S} &= vN - b_W(t)SW - b_I SI - \mu S, \\
    \dot{I} &= b_W(t)SW + b_I SI - \gamma I - \mu I, \\
    \dot{W} &= \alpha I - \xi W, \\
    \dot{R} &= (1 - f)\gamma I - \mu R, \\
    \dot{D} &= f\gamma I,
\end{align*}
\]  
(A 1)

where the host population is divided into susceptible (S), infectious (I) and recovered (R) compartments. The variable W tracks pathogen concentration in a water compartment (e.g. the River Thames and natural wells), and D is the number of individuals killed by the disease. Recovered individuals are immune to further infection. The total host population size is \( N = S + I + R \). The parameter \( v \) is the birth rate, \( \mu \) is the natural death rate, \( 1/\gamma \) is the mean infectious period and \( 1/\xi \) is the mean pathogen lifetime in the water compartment [23]. The parameter \( \alpha \) is the pathogen shedding rate per person–person contact, with rate parameter \( b_I \) or through the water, with rate parameter \( b_W(t) \). Seasonality in waterborne transmission is modelled using sinusoidal forcing,

\[
b_W(t) = B \left( 1 + A \cos \left( \frac{2\pi(t - t_f)}{T} \right) \right) \tag{A 2}
\]

A.2. Simulation details

Model variables and parameters for system (A 1) are summarized in tables 1 and 2, together with initial conditions and parameter values for simulating endemic and introduced cholera (figure 3). The birth and natural death rates were chosen to match London’s population growth between 1801 and 1901 (http://www.demographia.com/dm-lon31.htm). An expected infectious period of 3 days was used in the model (the typical infectious period is 1–5 days for cholera patients [22]). The ability of Vibrio cholerae to persist outside of human hosts depends upon environmental factors such as salinity [23] and temperature
Under appropriate conditions, *V. cholerae* can persist for extended periods of time in environmental water sources [25]. Here, we model the expected pathogen lifetime in the water to be of the order of one to two weeks. Case fatality rates for cholera in modern times range from a few percent to as high as 50 per cent [26], and was fixed at 10 per cent for our model. We set $\alpha$, the rate at which infected individuals shed pathogen into the water compartment, to 10 cells ml$^{-1}$ d$^{-1}$, a value that has been used in previous modelling efforts [27,28]. The transmission parameters were tuned to give reasonable fits to the cholera mortality data. This tuning was accomplished by first locating a periodic orbit for the model when $\nu = \mu$, such that this orbit matched the general seasonal pattern of London cholera in ‘typical’ years. Initial conditions for endemic cholera were taken from this periodic orbit. An initial population size of 100 000 was used in the simulations, rather than the population size of the entire city of London, since only a portion of the city’s population was at risk for cholera (e.g. John Snow’s finding that Vauxhall and Southwark Waterworks customers were at greatly elevated risk of infection [29]).

REFERENCES