

This Provisional PDF corresponds to the article as it appeared upon acceptance. Copyedited and fully formatted PDF and full text (HTML) versions will be made available soon.

Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling

Theoretical Biology and Medical Modelling 2006, **3**:38 doi:10.1186/1742-4682-3-38

Gabriel Sertsou (gabriel.sertsou@otago.ac.nz)
Nick Wilson (nick.wilson@otago.ac.nz)
Michael Baker (michael.baker@otago.ac.nz)
Peter Nelson (nelpe060@student.otago.ac.nz)
Mick G Roberts (m.g.roberts@massey.ac.nz)

ISSN 1742-4682

Article type Research

Submission date 26 September 2006

Acceptance date 30 November 2006

Publication date 30 November 2006

Article URL <http://www.tbiomed.com/content/3/1/38>

This peer-reviewed article was published immediately upon acceptance. It can be downloaded, printed and distributed freely for any purposes (see copyright notice below).

Articles in *TBiomed* are listed in PubMed and archived at PubMed Central.

For information about publishing your research in *TBiomed* or any BioMed Central journal, go to

<http://www.tbiomed.com/info/instructions/>

For information about other BioMed Central publications go to

<http://www.biomedcentral.com/>

Key transmission parameters of an institutional outbreak during the 1918 influenza pandemic estimated by mathematical modelling

Gabriel Sertsov¹, Nick Wilson^{1*}, Michael Baker¹, Peter Nelson¹, Mick G. Roberts²

¹Department of Public Health, Wellington School of Medicine & Health Sciences, University of Otago, Wellington, New Zealand

²Centre for Mathematical Biology, Institute of Information and Mathematical Sciences, Massey University, Auckland, New Zealand

*Corresponding author. Email: nick.wilson@otago.ac.nz

Abstract

Aim: To estimate the key transmission parameters associated with an outbreak of pandemic influenza in an institutional setting (New Zealand 1918).

Methods: Historical morbidity and mortality data were obtained from the report of the medical officer for a large military camp. A susceptible-exposed-infectious-recovered epidemiological model was solved numerically to find a range of best-fit estimates for key epidemic parameters and an incidence curve. Mortality data were subsequently modelled by performing a convolution of incidence distribution with a best-fit incidence-mortality lag distribution.

Results: Basic reproduction number (R_0) values for three possible scenarios ranged between 1.3, and 3.1, and corresponding average latent period and infectious period estimates ranged between 0.7 and 1.3 days, and 0.2 and 0.3 days respectively. The mean and median best-estimate incidence-mortality lag periods were 6.9 and 6.6 days respectively. This delay is consistent with secondary bacterial pneumonia being a relatively important cause of death in this predominantly young male population.

Conclusions: These R_0 estimates are broadly consistent with others made for the 1918 influenza pandemic and are not particularly large relative to some other infectious diseases. This finding suggests that if a novel influenza strain of similar virulence emerged then it could potentially be controlled through the prompt use of major public health measures.

Background

The 1918 influenza pandemic reached New Zealand with an initial wave between July and October [1]. This was relatively mild with only four deaths out of 3048 reported cases for the population of military camps [1]. The second wave in late October was much more severe and spread throughout the country causing over 8000 deaths [2]. One large military camp near Featherston (a town in the south of the North Island) also suffered from exposure to the second wave of the 1918 pandemic at approximately the same time as the rest of the country. Influenza cases were reported in the camp from 28 October to 22 November 1918, and reported mortality occurred between 7 November and 11 December 1918, with both incidence and mortality peaking in November 1918 [2]. A unique feature of this military camp outbreak was the systematic collection by medical staff of morbidity data as well as mortality data.

We undertook modelling of these data to understand better the transmission dynamics of the 1918 influenza pandemic in New Zealand.

Methods

Data: The population of the Featherston Military Camp was that of a large regional town, comprising approximately 8000 military personnel of whom 3220 were hospitalised [3]. The camp policy was to hospitalise all those with diagnosed influenza and so we have used these hospitalisation data as the basis for the incidence of pandemic influenza in this population. An official report indicated a total of 177 deaths attributable to the outbreak [4]. However, this figure was actually the total number of men who died in the camp in 1918 from all causes as reported by the Principal Medical Officer at the camp [3]. Further examination of data on the cause of death and date-of-death suggests the total mortality attributable to this outbreak was 163 [5]. This revision gives a fairly conservative figure for the mortality impact and it is the one that we have used in this analysis.

Mathematical modelling approach: A susceptible-exposed-infectious-recovered (SEIR) model for infectious diseases can be applied to a hypothetical isolated population, to investigate local infection dynamics [6, 7]. The SEIR model allows a systematic method by which to quantify the dynamics, and derive epidemiological parameters for disease outbreaks. In this model, individuals in a hypothetical population are categorized at any moment in time according to infection status, as one of susceptible, exposed, infectious, or removed from the epidemic process (either recovered and immune or deceased). If an infected individual is introduced into the population, rates of change of the proportion of the population in each group (s , e , i , and r , respectively) can be described by four simultaneous differential equations:

$$\frac{ds}{dt} = -\beta si \quad (1)$$

$$\frac{de}{dt} = \beta si - ve \quad (2)$$

$$\frac{di}{dt} = ve - \gamma i \quad (3)$$

$$\frac{dr}{dt} = \gamma i \quad (4)$$

where β , v and γ are rate constants for transformation of individuals from susceptible to exposed, from exposed to infectious, and from infectious to recovered and immune states, respectively. Once the above equations have been solved, the parameters β and γ can be utilized to calculate the basic reproduction number (R_0) for the particular virus strain causing the outbreak. (The basic reproduction number represents the number of secondary cases generated by a primary case in a completely susceptible population). R_0 and the average latent period (T_E), and average infectious period (T_I), can be calculated using the following relationships:

$$R_0 = \frac{\beta}{\gamma} \quad (5)$$

$$T_E = \frac{1}{\nu} \quad (6)$$

$$T_I = \frac{1}{\gamma} \quad (7)$$

Other factors that are likely to affect the observed incidence of disease in a pandemic include the following: (i) the initial proportion of population that is susceptible (P_{is}); (ii) the proportion of infected cases who develop symptoms (P_{ids}); (iii) the infectivity of asymptomatic people relative to the infectivity of symptomatic people (Inf_{as}); and (iv) the proportion of symptomatic cases who present (P_{sp}).

In this study, the factors listed above were incorporated into an SEIR model to generate incidence and subsequent mortality models for the influenza pandemic that swept through this military camp. These specific models and the resulting estimates of R_0 and T_E and T_I are described below.

SEIR model of incidence: When the SEIR model was applied in this study, assumptions about additional factors that might influence the observed incidence were made. The parameters associated with these assumptions are summarised for 3 possible scenarios (Table 1). Parameters in Scenarios 1, 2, and 3 were chosen so that models would yield estimates of R_0 at the lower, mid-range and higher ends of a likely spectrum, respectively.

Equations 1 and 2 were modified to take the above parameters into account, as follows:

$$\frac{ds}{dt} = -\beta(P_{ids} + (1 - P_{ids})Inf_{as})si \quad (8)$$

$$\frac{de}{dt} = \beta(P_{ids} + (1 - P_{ids})Inf_{as})si - \nu e \quad (9)$$

Equations 3, 4, 8 and 9 are a system of non-linear differential equations, amenable to solution by the Runge-Kutta fourth order fixed step numerical method [8]. The population size was taken to be $N = 8000$. The initial value for s was $P_{is} - 1/N$, and initial values of e , i , and r were set at 0, $1/N$ and $1 - P_{is}$ respectively. The differential equation system solutions were used to calculate daily incidence, taking into account parameters in Table 1, using the following equation:

$$Incidence = P_{sp}P_{ids}N(s(t-1) - s(t)) \quad (10)$$

in which $s(t)$ and $s(t-1)$ are the proportion of susceptible individuals at t and $t-1$ days respectively after the introduction of a single symptomatic individual into the population.

For each scenario in Table 1, modelled incidence was compared to observed incidence over 26 days, and goodness of fit of the models was evaluated using sum of squared error (SSE) between modelled and empirical data. Optimum possible β , ν and γ values to one decimal place, in the range 0.1 to 20, were determined by finding values corresponding to a minimum SSE, utilizing an algorithm written in Mathcad [9].

The asymptotic variance-covariance matrix of the least squares estimates of β , ν and γ , was computed using the method described by Chowell et al. [10]. Equations 5, 6, and 7, together with elements of the variance-covariance matrix, and the Taylor series approximation for variance of quotients [11], were subsequently used to estimate best-fit values of R_0 , T_E and T_I , with associated standard deviations and confidence intervals.

Associated mortality model: As morbidity and mortality data are not linked at the individual level, case-fatality lag was modelled by using convolution. A least-squares gamma distribution was fitted to the observed incidence curve. A gamma distribution with the same scale parameter was then fitted to mortality data. Utilising these distributions and the convolution formula, a gamma distributed incidence-mortality lag distribution, with the same scale parameter, was obtained.

Gamma distributions with the same scale parameter were then fitted to the best-fit deterministic models of daily incidence. These distributions, convolved with the incidence-mortality lag distribution, yielded daily mortality distributions for each of Scenarios 1 to 3. A common scale parameter was used in the above convolutions in order to obtain closed-form (gamma) probability density functions.

Results

Best-fit incidence curves from the SEIR model for the three scenarios are shown in Figure 1. The corresponding best-fit β , ν and γ , and corresponding R_0 , T_E and T_I values, are shown in Table 2. The R_0 values ranged between 1.3, and 3.1, and corresponding average latent period and infectious period estimates ranged between 0.7 and 1.3 days, and 0.2 and 0.3 days, respectively.

The gamma distribution of incidence-mortality lag time obtained by convolution is shown in Figure 2. The mean, median, mode and variance of this distribution are 6.9, 6.6, 6.0 and 6.3 days respectively.

Observed mortality data, shown in Figure 3, indicate more variability around a best-fit gamma distribution than observed incidence data (see Figure 1). Mortality curves for each of Scenarios 1 to 3, obtained by convolution, all agree well with the best-fit gamma distribution of observed data.

Discussion

This analysis has demonstrated the potential for using historical disease epidemic data to derive plausible, and potentially useful, pandemic influenza parameter estimates.

This is the first time that these parameters have been reported for the 1918 pandemic outside of Europe, the USA and Brazil.

Limitations of this analysis: This work is limited by the very nature of using data from an event that occurred over eight decades ago. For example, the estimate of the camp's population was only approximate (at 8000). The mortality burden of this particular outbreak (at 20.4 per 1000) was also somewhat higher than that for the general male population of New Zealand (ie, at 10.0 per 1000 for 20-24 year olds) [2]. It was, however, similar to the pandemic influenza mortality burden of the armed forces as a whole (at 23.5 per 1000) and for other military camps at 22.0 and 23.5 (for Awapuni and Trentham camps respectively) [2]. It is plausible that higher death rates in military camps may have been related to both higher risk of infection (e.g. via crowding) and the poor living conditions involved (i.e. the extensive use of tents). Crowded troop trains may also have contributed to disease spread and in the weekend prior to the main outbreak in the camp many of the recruits had been away on leave, and were transported to and from the camp by troop trains. Furthermore, a severe storm struck the Featherston camp on 7 November (the day that influenza incidence peaked) and flattened many tents. This event placed additional stresses on accommodating men in huts that were already full and with some huts (and all institute buildings such as the YMCA, for example) being used as overflow wards to the main camp hospital to which the most severe cases were admitted. Less severe cases were admitted to makeshift wards in the so-called institute buildings, and the huts were used for convalescence. In his report, the Principal Medical Officer commented that this storm was likely to have exacerbated the impact of the outbreak and this is certainly plausible [3].

In addition to data limitations, the parameters used for the SEIR model also involve uncertainties; for example, we have no good data on the proportion of the young male population who were likely to be susceptible to this strain in 1918 (e.g. based on the possible residual immunity from the first wave of the pandemic or from previous influenza epidemics and pandemics). Also, the SEIR model involves a number of simplifying assumptions, including a single index case, homogeneous mixing, exponentially distributed residence times in infectious status categories, and isolation of the military camp.

Estimating R_0 : The estimates for R_0 in the range from 1.3 to 3.1 are the first such estimates for the 1918 pandemic outside Europe, the United States and Brazil, so far as we are aware. However, given the unique aspects of the military camp (crowded conditions and a young population with low immunity) it is quite likely that the R_0 values estimated in our analysis might tend to over-estimate those for the general population. Nevertheless, this effect may have been partly offset by the camp policy of immediate hospitalisation upon symptoms, effectively reducing infective contacts.

Our estimated range for R_0 is broadly consistent with estimates for this pandemic in the United States (a median R_0 of 2.9 for 45 cities) [12]. Other comparable figures for the 1918 pandemic are: 1.7 to 2.0 for the first wave for British city-level mortality data [13]; 2.0, 1.6 and 1.7 for the first, second and third waves in the UK respectively [14]; 1.5 and 3.8 in the first and second waves in Geneva respectively [15]; and 2.7 for Sao Paulo in Brazil [16]. The upper end of our estimated range ($R_0 = 3.1$) may reflect the differences between disease transmission in the general population (as per

the above cited studies) and transmission in a crowded military camp with a predominance of young males.

Considered collectively, these R_0 estimates for pandemic influenza in various countries are not particularly high when compared to the R_0 estimates for various other infectious diseases [17]. This observation provides some reassurance that if a strain of influenza with similar virulence were to emerge, then there would be scope for successful control measures. Indeed, one model, using R_0 values in the 1.1 to 2.4 range, has suggested the possibility of successful influenza pandemic control [18]. This was also the case for a model using $R_0 = 1.8$ [19]. Nevertheless, at the upper end of the estimated range for R_0 , control measures may be more difficult, especially if public health authorities are slow to respond and they have insufficient access to antivirals and pandemic strain vaccines.

The latent and infectious periods: The average latent and infectious periods were estimated to be in the range between 0.7 to 1.3 days, and 0.2 to 0.3 days, respectively. The infectious period is short compared to the period of peak virus shedding known to occur in the first 1 to 3 days of illness [20]. Other modelling work has used longer estimates, e.g. a mean of 4.1 days used by Longini et al. [18].

The fast onset and subsequent decline of the outbreak in the Featherston Military Camp, as compared to a national or city-wide outbreak, might possibly be due to relatively close habitation and a high level of mixing. The average time for infection between a primary and secondary case (the serial interval) is greatly shortened in this case. This could explain a short apparent infectious period, and a relatively large proportion of the serial interval in the latent state. Another possible explanation of the relatively short apparent infectious period for this outbreak is that it may reflect the limited transmission that occurred once symptomatic individuals were hospitalised on diagnosis – which was the policy taken in this military camp for all cases.

The lag period from diagnosed illness to death: This analysis was able to estimate an approximate seven-day delay from reported symptomatic illness to the date of death at a population level. This result is suggestive that even in this relatively young population (largely of military recruits), an important cause of death was likely to have been from secondary bacterial pneumonia – as opposed to the primary influenza viral pneumonia or acute respiratory distress syndrome (for which death may have tended to occur more promptly). This finding is consistent with other evidence that a large proportion of deaths from the 1918 pandemic was attributable to bacterial respiratory infections [21]. This picture is also somewhat reassuring as it suggests that much of this mortality could be prevented (with antibiotics) if a novel strain with similar virulence emerged in the future.

Conclusions

The R_0 estimates in the 1.3 to 3.1 range are broadly consistent with others made for the 1918 influenza pandemic and are not particularly large relative to some other infectious diseases. This finding suggests that if a novel influenza strain of similar virulence emerged then it could potentially be controlled through the prompt use of major public health measures. These results also suggest that effective treatment of

pneumonia could result in better outcomes (lower mortality) than was experienced in 1918.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Three of authors were involved in initial work in identifying the data and analysing it from a historical and epidemiological perspective (PN, NW and MB). The other two authors worked on developing and running the mathematical model (GS, MR). GS did most of the drafting of the first draft of the manuscript with assistance from NW. All authors then contributed to further re-drafting of the manuscript and have given final approval of the version to be published.

Acknowledgements

We thank the following medical students for their work in gathering information on the outbreak in the Featherston military camp: Abdul Al Haidari, Abdullah Al Hazmi, Hassan Al Marzouq, Melinda Parnell, Diana Rangihuna, Yasotha Selvarajah. We also thank the journal's reviewers for their helpful comments. Some of the work on this article by two of the authors (NW & MB) was part of preparation for a Centers for Disease Control and Prevention (USA) grant (1 U01 CI000445-01).

References

1. Maclean F: *Challenge for Health: A history of public health in New Zealand*. Wellington: Government Printer; 1964.
2. Rice G: *Black November: The 1918 influenza pandemic in New Zealand*. Christchurch: Canterbury University Press; 2005.
3. Carbery A: *The New Zealand Medical Service in the Great War 1914-1918*: Whitcombe & Tombs Ltd. pp. 506-509; 1924.
4. Henderson R, *New Zealand Expeditionary Force. Health of the Troops in New Zealand for the year 1918*, in *Journal of the House of Representatives, 1919*. Wellington Marcus F. Mark; 1919.
5. Al Haidari A, Al Hazmi A, Al Marzouq H, Armstrong M, Colman A, Fancourt N, McSweeney K, Naidoo M, Nelson P, Parnell M, Rangihuna-Winekerei D, Selvarajah Y, Stantiall S: *Death by numbers: New Zealand mortality rates in the 1918 influenza pandemic*. Wellington: Wellington School of Medicine and Health Sciences; 2006.
6. Anderson R, May R: *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press; 1991.
7. Diekmann O, Heesterbeek J: *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*. Chichester: Wiley; 2000.
8. Zill D: *Differential Equations with Boundary-Value Problems*. Boston: PWS-Kent Publishing Company; 1989.

9. Mathsoft: *Mathcad version 13*. Cambridge, MA: Mathsoft Engineering & Education, Inc; 2005.
10. Chowell G, Shim E, Brauer F, Diaz-Duenas P, Hyman JM, Castillo-Chavez C: **Modelling the transmission dynamics of acute haemorrhagic conjunctivitis: application to the 2003 outbreak in Mexico**. *Stat Med* 2006, **25**:1840-1857.
11. Mood A, Graybill F, Boes D: *Introduction to the Theory of Statistics (3rd Ed)*. Chichester: McGraw-Hill; 1982.
12. Mills CE, Robins JM, Lipsitch M: **Transmissibility of 1918 pandemic influenza**. *Nature* 2004, **432**:904-906.
13. Ferguson NM, Cummings DA, Fraser C, Cajka JC, Cooley PC, Burke DS: **Strategies for mitigating an influenza pandemic**. *Nature* 2006, **442**:448-452.
14. Gani R, Hughes H, Fleming D, Griffin T, Medlock J, Leach S: **Potential Impact of Antiviral Drug Use during Influenza Pandemic**. *Emerg Infect Dis* 2005, **11**:1355-1362.
15. Chowell G, Ammon CE, Hengartner NW, Hyman JM: **Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: Assessing the effects of hypothetical interventions**. *J Theor Biol* 2006, **241**:193-204.
16. Massad E, Burattini MN, Coutinho FA, Lopez LF: **The 1918 influenza A epidemic in the city of Sao Paulo, Brazil**. *Med Hypotheses* 2006, **Sep 28**; [Epub ahead of print].
17. Fraser C, Riley S, Anderson RM, Ferguson NM: **Factors that make an infectious disease outbreak controllable**. *Proc Natl Acad Sci U S A* 2004, **101**:6146-6151.
18. Longini IM, Jr., Nizam A, Xu S, Ungchusak K, Hanshaoworakul W, Cummings DA, Halloran ME: **Containing pandemic influenza at the source**. *Science* 2005, **309**:1083-1087.
19. Ferguson NM, Cummings DA, Cauchemez S, Fraser C, Riley S, Meeyai A, Iamsirithaworn S, Burke DS: **Strategies for containing an emerging influenza pandemic in Southeast Asia**. *Nature* 2005, **437**:209-214.
20. WHO Writing Group: **Non-pharmaceutical interventions for pandemic influenza, international measures**. *Emerg Infect Dis* 2006, **12**:81-87.
21. Brundage JF: **Interactions between influenza and bacterial respiratory pathogens: implications for pandemic preparedness**. *Lancet Infect Dis* 2006, **6**:303-312.

Table 1. Parameters used in the SEIR incidence model*

Parameter	Scenario 1	Scenario 2	Scenario 3
Initial proportion of the population susceptible (P_{is})	1.0	0.9	0.8
Proportion of infected cases who develop symptoms (P_{ids})	0.95	0.81	0.67
Infectivity of asymptomatic/infectivity of symptomatic people (Inf_{as})	0.6	0.5	0.4
Proportion of symptomatic cases who present and are diagnosed as infected with influenza (P_{sp})	0.95	0.88	0.8

*Based on plausible ranges for pandemic influenza with Scenario 1 being closer to a worse case for impact on health and Scenario 3 being less severe. For example, Scenario 3 assumes 20% of the population may have had immunity from previous influenza pandemics that may have reached New Zealand in the late 19th century – as suggested by Rice [2] and supported by the unusually low mortality rates in the older age groups for this pandemic in New Zealand [2].

Table 2. Rate constants and epidemiological parameters corresponding to the best-fit models shown in Figure 1 (associated standard deviation or 95% confidence interval is given in brackets).

Scenario	β (days⁻¹)	ν (days⁻¹)	γ (days⁻¹)	R_0	Latent period T_E (days)	Infectious period T_I (days)
1	5.3 (0.50)	1.5 (0.08)	4.2 (0.33)	1.3 (0.02)	0.67 (0.60, 0.74)	0.24 (0.21, 0.28)
2	6.5 (0.27)	1.2 (0.04)	3.6 (0.11)	1.8 (0.04)	0.83 (0.78, 0.89)	0.28 (0.26, 0.30)
3	10.1 (1.55)	0.8 (0.11)	3.3 (0.36)	3.1 (0.18)	1.25 (0.99, 1.69)	0.30 (0.25, 0.38)

Figure 1. Observed and best-fit modelled incidence (ill cases per day) for Scenarios 1 to 3, and best-fit gamma distribution.

Figure 2. Incidence-mortality lag time distribution.

Figure 3. Observed and best-fit modelled mortality (deaths per day) for Scenarios 1 to 3.

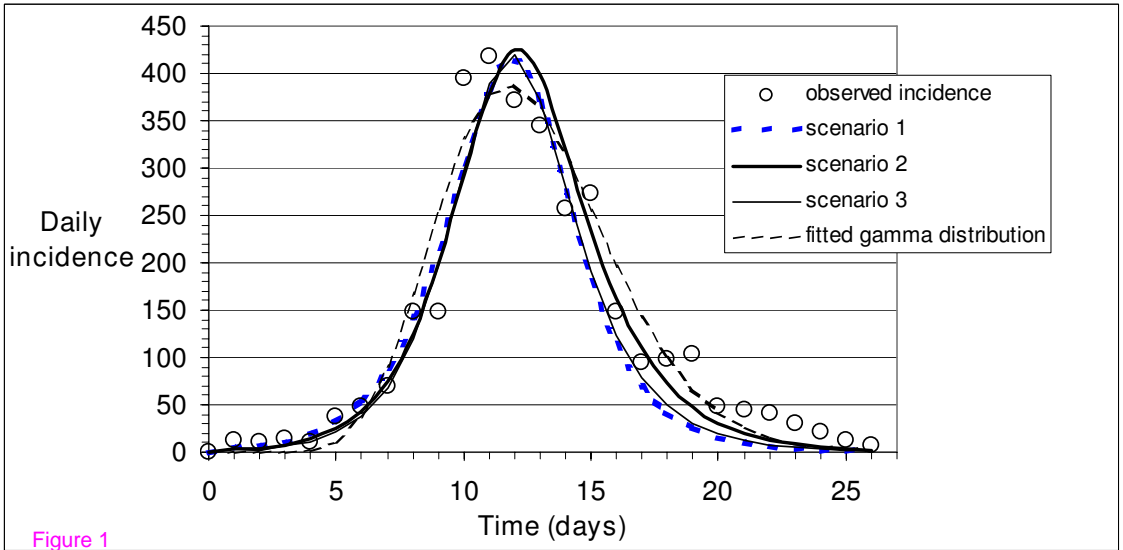


Figure 1

Probability density

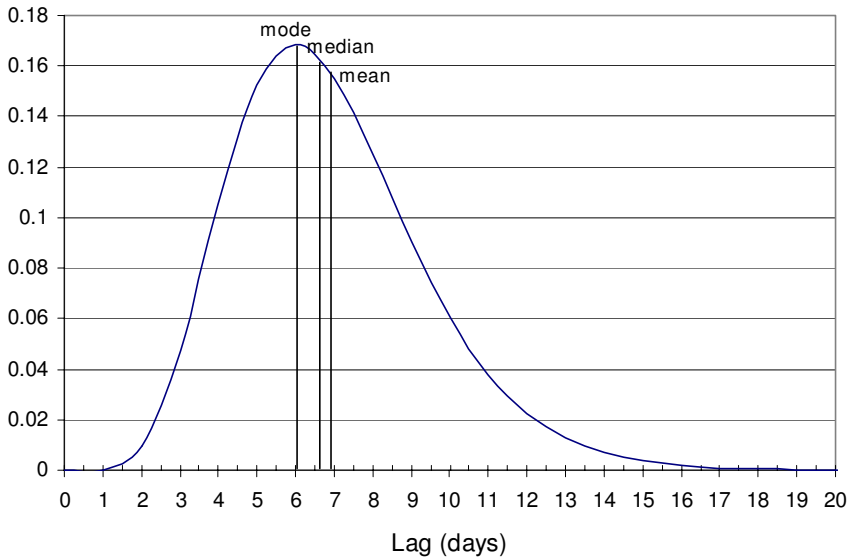


Figure 2

Daily mortality

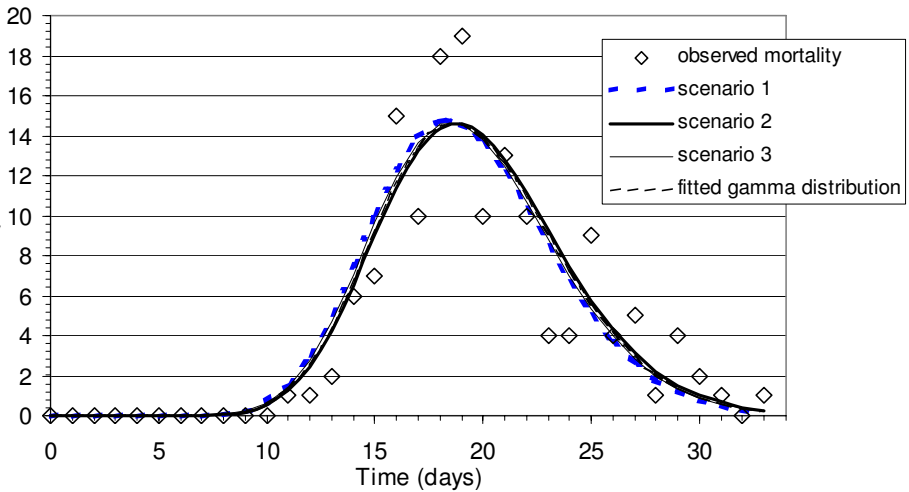


Figure 3