

Real Time Modelling of Pandemic Influenza

Raymond Gani, Ian Hall, Helen Hughes, Steve Leach

Health Protection Agency Centre for Emergency Preparedness and Response, Porton Down, Salisbury, Wilts, UK.

Global Health Security Action Group (GHSAG)
Influenza Modelling Workshop & Exercise 1-3 June 2005

Introduction

The threat posed by pandemic influenza is well documented. Many countries have devised contingency plans based on experience from seasonal influenza and previous pandemics. However, the characteristics of these have varied considerably leading to much uncertainty over what will be the precise characteristics and impacts of any new pandemic. In the early stages of a pandemic, previous contingency planning will need to be re-assessed which will mean re-parameterising the models that had been used in the formulation of these contingency plans and to optimise strategies to deal with the pandemic as it unfolds. This therefore highlights the need for real-time predictions of the future number of casualties and public health burdens and the need would be to provide rapid assessment of the potential impact of the pandemic.

The data required for such modelling can be broadly divided into disease specific and population specific. The disease specific data relate to the progression and consequences of infection within individuals. Examples of these would be the duration of the latent and infectious periods, and the case fatality rate. Whilst a number of factors may influence these, such as age and predisposing risk factors, it would be anticipated that the disease specific parameters would be determined from epidemiological investigations conducted early on in the pandemic and remain fairly consistent throughout it. The population specific parameters relate to the rate of disease transmission within the population through which the disease is moving. This may vary between different populations and at different spatial and social scales, and change over time. Two key reasons for this are that contact patterns are important to disease transmission and these are likely to vary intrinsically between different populations in relation to social structure; and second may change as a consequence of behavioural responses driven by concerns over the impact of the disease on the population as determined by reported incidence. It is the population specific parameters that we shall attempt to derive as these will enable us to estimate quantities such as the timing and amplitude of peak incidence, and public health burdens, across the duration of the pandemic.

A number of models have been proposed for real-time modelling of diseases that might not be satisfactory for modelling the impacts of pandemic influenza at a national scale. In particular, current models used for seasonal influenza rely on previous epidemics as a predictor of future epidemics (See Hall *et al.* for review), which would not be satisfactory for a pandemic with a completely new strain. Other methods, such as those developed for SARS (Wallinga & Teunis 2004) are unlikely to be applicable unless independent estimates of the ratio of reported incidence to serological incidence could be attained. Also, such methods often rely on estimating

the transmission rate from fine local scale epidemiological data, which may not either be available or extrapolate satisfactorily to provide predictions at a national scale.

Past Pandemic Data

The data available from previous pandemics is limited, with none found that were multiple measures of incidence across similar spatial scales. Therefore, the data that was used to investigate the approaches outlined in this paper comprised of single datasets from the UK from each of the three last pandemics in 1957, 1968 and 1918 (Fig. 1). It would be anticipated that in future pandemics multiple datasets would be available across similar spatial scales.

The data from the 1957 pandemic shown in Fig. 1a are the recorded deaths attributed to influenza in England and Wales from July 1957 until February 1958 (Anon. 1960). This pandemic strain was first identified in China in February 1957, with the first recorded imported cases in the UK being found in London during June 1957. This was followed by localised outbreaks during July and August. Fig. 1a shows that the epidemic appeared to take off in September, peaking in November then subsiding before a second wave, which appeared in January. It should be noted that these data represent only one measure of the deaths that might be attributable to pandemic influenza, and that excess deaths due to pandemic influenza may have been categorised elsewhere.

The characteristics of the 1968 pandemic in the UK varied considerably from those of the 1957. The first recorded cases in the UK were during August 1968 but the pandemic failed to take off. The records by the Royal College of General Practitioners of influenza-like-illness (ILI) from GP sentinel practices across England and Wales did not record unusually high incidence rates for the first small wave. These are shown in Fig. 1b along with the previous year's recorded seasonal influenza for comparison. It wasn't until the winter of 1969/70 that the pandemic really struck.

The 1918 pandemic was, again, markedly different from those that began in 1957 and 1968. It was characterised by three waves, of varying severity occurring at different times of the year. Fig. 1c shows the recorded deaths from influenza in England and Wales over this period, and shows the first wave during the summer of 1918, then the second in the autumn and the third in the spring of 1919. This pandemic was characterised by higher case fatality rates and different age-specific attack rates than the other two pandemics, which also varied between the three waves (Anon 1920).

Methods

This paper examines the potential to forward predict the onward trajectory of a pandemic of fitting procedures based on simple epidemic models applied to the types of data that have been routinely collected during past pandemics. Epidemic curves from previous influenza pandemics have typically followed dynamics that can be generalised using SEIR-type formulae. The basic assumption here is that the rate of change of the infected population is proportional to the product of the proportion of the population susceptible and the proportion infectious. The model is fitted to several consecutive weeks of incidence data (reported clinical cases or deaths in the first few weeks of the national epidemic) by minimising the RMS error by iteratively adjusting the transmission rate, R_s , the proportion of serological infections recorded, τ , and a

nominal start date, d . The procedure is repeated progressively as data from each new week is available. Further details can be found in Hall *et al.*

Results

The SEIR-type epidemic model was fitted independently to the different waves of the three pandemics as described above, and a selection of the results are shown here to demonstrate its forward predictive capabilities. The graphs shown in Figs 2-5 show the full epidemic curve, the real-time modelling fit and the early data points that were used to estimate parameters and generate that fit.

Figs. 2 and 3 are for the second and third waves of the 1918 pandemic. Data for the first wave was not fitted as there were insufficient data. Fig. 2 demonstrates a number of aspects about this model. The early parameter estimates based on four data points (Fig. 2a) over-predict the amplitude of the peak of deaths attributed to influenza, which is due to the estimate of R_ϵ based on the initial points being higher than the average across the epidemic. However, even at this early stage, the duration of the epidemic has been estimated well. As more data points are gathered, the estimates of the predicted amplitude improve markedly until by Fig. 2d, they coincide. Whilst this is the actual peak of the epidemic, it should be noted that this wouldn't have been known to have been the case until later in the epidemic. Thereafter, the predictions clearly confirm that the pandemic wave is in decline. The earliest prediction for the third wave based on 3 data points in Fig. 3a seriously underestimates both its amplitude and duration. This results from the relatively poor signal in the data at the start of the wave compared to the inherent noise. The addition of one extra data point however, markedly improved the prediction of both the amplitude and duration of this wave. Whilst amplitude continued to be predicted well, as was seen with the analysis of the previous wave, including data in the fitting procedure that followed the peak underestimated the extent of the tail of the epidemic with incidence remaining somewhat higher for longer than expected.

The results of the fitting procedure for the main waves of the 1957 and 1968 pandemics are shown in Figs. 4 and 5. For these, only R_ϵ and d were fitted, as estimates of the proportion of serological infections that report, τ , were obtained independently, which demonstrates this proposed approach when applied in conjunction with other relevant data sets. The results for the first wave of the 1957 pandemic are shown in Fig. 4. The earliest forward predictions of influenza deaths shown (Figs 4a and 4b) over-predict the amplitude of the real epidemic by about 50% due to the estimates of R_ϵ , being higher at the start than the average across the wave. Prediction of the duration and amplitude of the epidemic improves markedly as further data points are included in the analysis beyond 5 weeks into the epidemic and towards its peak. When applied to the second wave of the 1968 pandemic (Fig. 5) the fitting procedure provides quite reasonable estimates of the duration and amplitude of ILI reported to the RCGP even from the first 4 weeks of data onwards.

Conclusions

The approach demonstrated here is based on parameter estimation from simple classical epidemic modelling applied to national data of the type that are likely to be readily available in the event of a new pandemic. The approach has been extended to attempt to forward predict the likely timing and amplitude of peak influenza incidence as determined by deaths attributable to influenza or ILI reported to GPs, and the

duration of individual epidemic waves. Reasonable estimates of the amplitude and duration of epidemic waves were generally achieved, though this often required at least 4-5 weeks of data to achieve more convincing fits. This usually occurred close to the peaks of the real epidemic waves, which generally only lasted for about 10-12 weeks.

The data that is likely to be available is likely to be noisy, particularly at the tails of the epidemic where misdiagnosed cases will have their greatest impact on predictions. This is particularly important at the start of the epidemic when there is little new pandemic data, producing the greatest uncertainty surrounding the reliability of predictions. However, the data that was used in this study was collated at weekly intervals – finer temporal resolution in the collected data (i.e. daily or twice weekly data) would be expected to lead to better earlier predictions. The robustness of the forward predictions might also be improved by using modelling approaches that simultaneously fitted to multiple relevant data sets. For example, in the UK such data might include that from the Royal College of General Practitioners, NHS Direct, the Office of National Statistics and Hospital Episode Statistics.

There are a number of issues that have not been addressed in this paper that if taken into account could also improve the applicability of such modelling approaches. Each of the 20th century pandemics has arrived in more than one wave, however the model has been applied to each of these waves separately. Factors that probably affect the transmission rate of pandemic influenza in populations include seasonality since inter-pandemic flu epidemics are seasonal and each of the previous pandemics has peaked in the autumn in the UK. However, the drivers for a new pandemic influenza strain may not be quite the same as for seasonal influenza which tends to peak later than this in the UK, during December and January. Therefore, although influenza transmission is considered to be affected seasonally, why these waves have occurred precisely when they have in the past is not well understood. If this could be explained mathematically with sufficient precision then such modelling approaches could be modified to account for this.

The data used in this study were based on aggregate national incidence rates. The predictions derived from them are therefore, similarly, for projected national incidence. If finer spatial resolutions to such predictions were required then corresponding observational data on an equivalent spatial or social scale would be likely to be required, as it would be unwise to extrapolate across spatial scales, or to apply results from one population to another.

In summary, the basic requirements for the methodology are the disease specific parameters that describe the kinetics of the disease and its outcomes within individuals. Required next is an invariant temporal marker of the incidence of the disease, such as number of deaths or clinical cases. An independent estimate of the ratio of reported incidence to serological infections is not essential as it can be estimated from the fitting procedure, but a prior estimate would enable more robust predictions to be made earlier in the outbreak. Reasonable forward predictions of the timing and amplitude of the measure of incidence being analysed as well as the duration of the national epidemic can then be provided, but for the last three pandemics this has usually required 4 to 5 data points based on aggregate weekly data. This might be improved by having more frequently reported data available and fitting to more than one measure of incidence at the same time. The best forward predictions are likely to come from the most reliable measures of incidence that are confounded least by misdiagnosed cases. Therefore, prior to any pandemic, it is vital that there are in place good surveillance systems that are able to return data promptly,

and robust diagnostic capability to ensure that the data being collected is as accurate as possible.

References

Anon. (1920) Reports on public health and medical subjects no. 4: report on the pandemic of influenza 1918–19. London: Her Majesty's Stationary Office.

Anon. (1960) The influenza epidemic in England and Wales 1957–58. Reports on public health and medical subjects no. 100. London: Her Majesty's Stationary Office.

Hall I, Gani R, Hughes H, Leach S. Epidemic forecasting and parameter estimation from real-time data feeds. In Press.

Wallinga J, Teunis P (2004) Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol.* 2004 Sep 15;160(6):509-16.

Figure 1: Previous influenza pandemics in the UK during the 20th Century. (a) Asian flu 1957, (b) Spanish flu 1918, (c) Hong Kong flu 1968

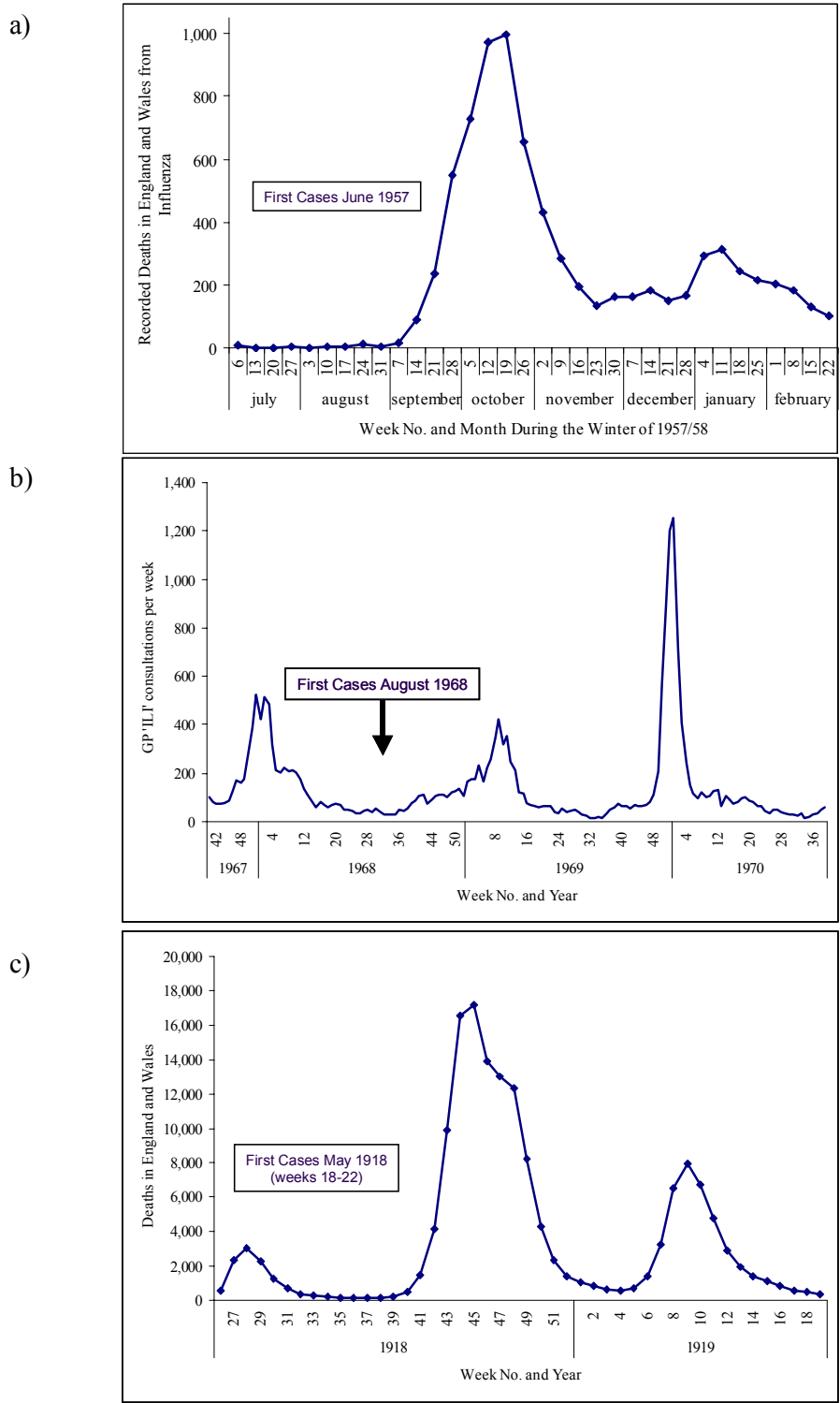


Figure 2: Model fitted to the second wave of the 1918 pandemic by adjusting R_0 , τ and d . The horizontal axis is time in weeks and the vertical axis is number of recorded influenza deaths per week in the UK. The solid blue line is the observations, the blue dots the data points to which the model was fitted and the red line the prediction.

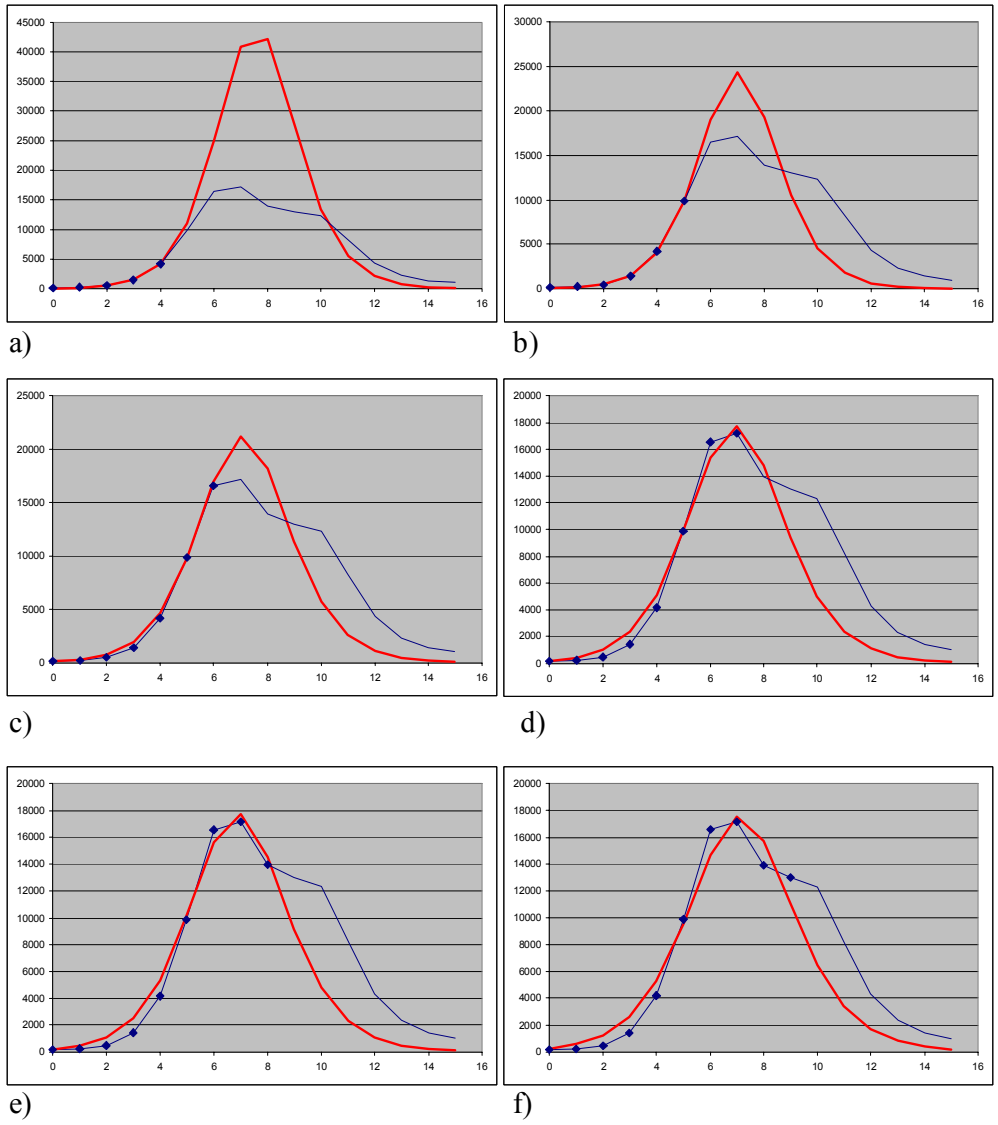
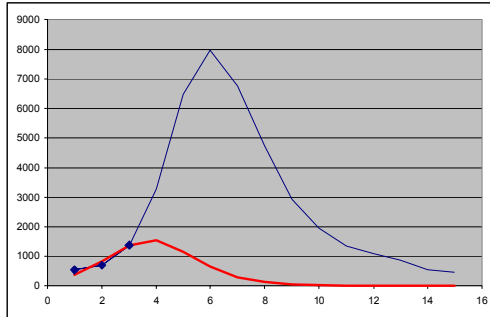
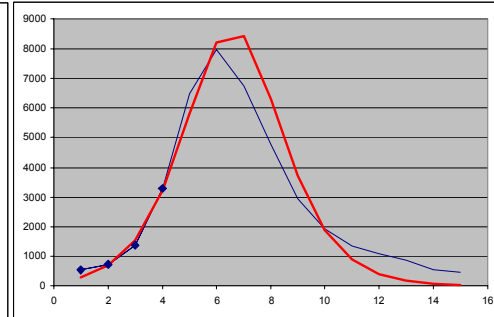


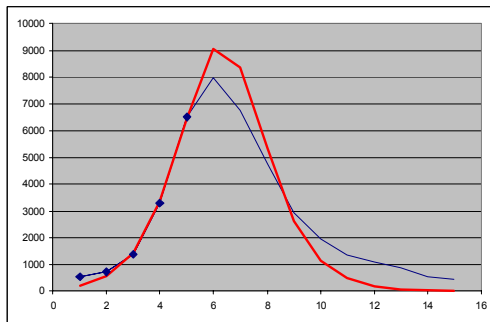
Figure 3: Model fitted to the third wave of the 1918 pandemic by adjusting R_e , τ and d . The horizontal axis is time in weeks and the vertical axis is number of recorded influenza deaths per week in the UK. The solid blue line is the observations, the blue dots the data points to which the model was fitted and the red line the prediction.



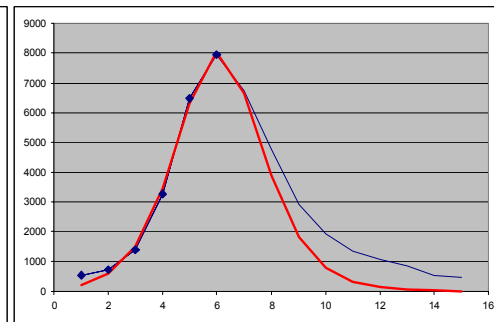
a)



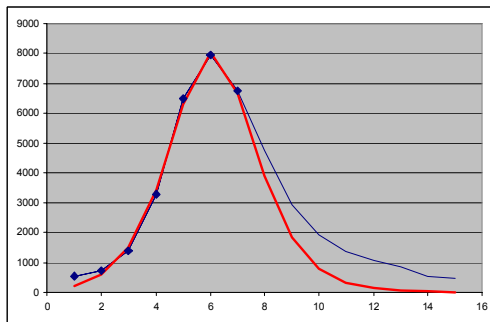
b)



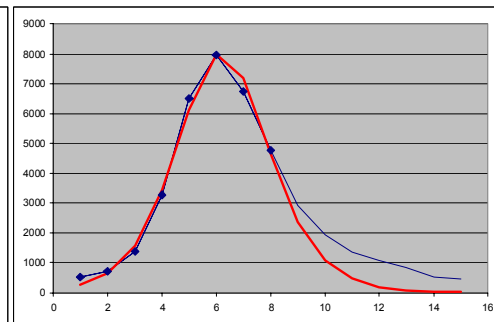
c)



d)

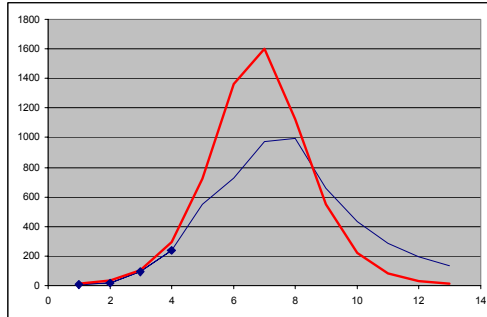


e)

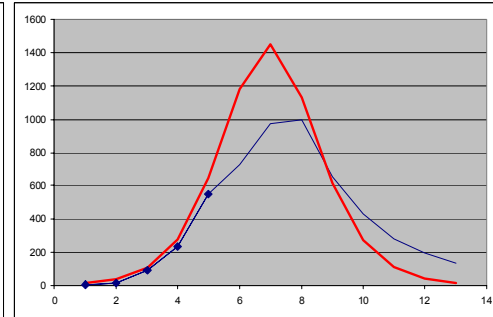


f)

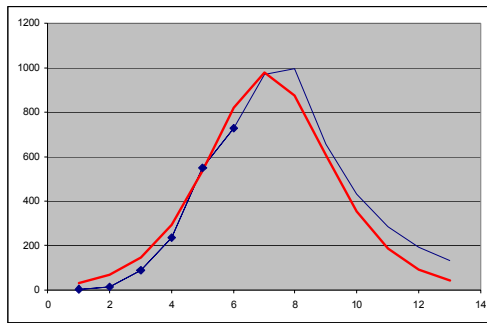
Figure 4: Model fitted to the first wave of the 1957 pandemic by adjusting R_e and d . The horizontal axis is time in weeks and the vertical axis is number of recorded deaths attributed to influenza per week in the UK. The solid blue line is the observations, the blue dots the data points to which the model was fitted and the red line the prediction



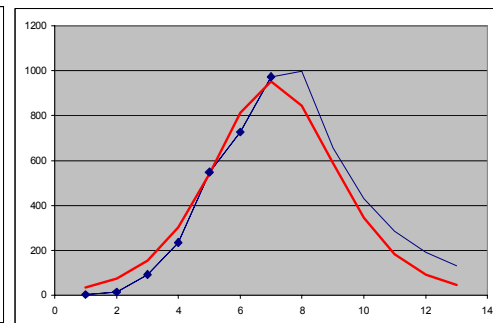
a)



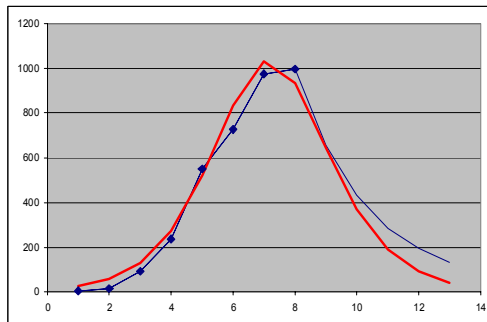
b)



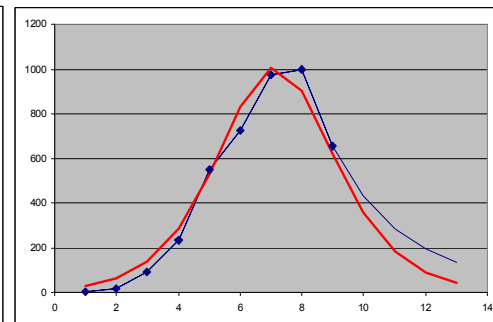
c)



d)

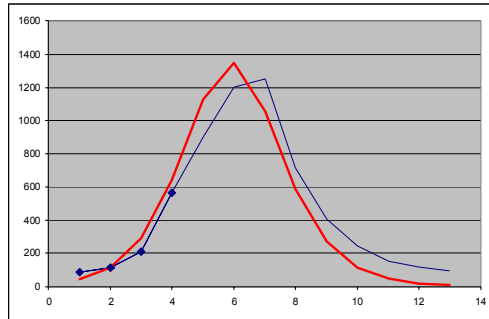


e)

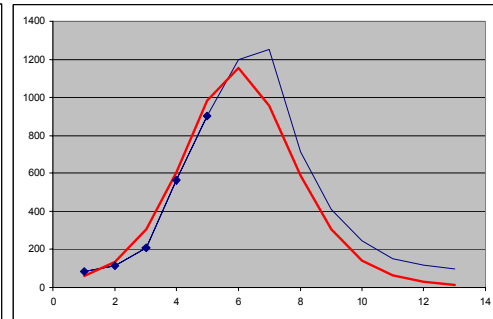


f)

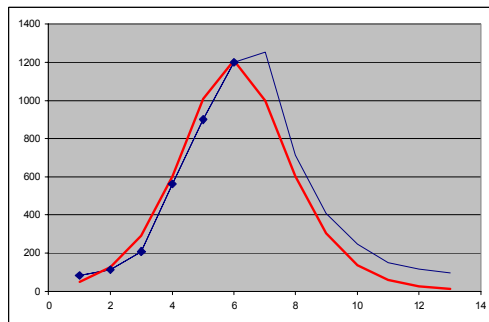
Figure 5: Model fitted to the second wave of the 1968 pandemic by adjusting R_e and d . The horizontal axis is time in weeks and the vertical axis is number of recorded GP consultations for influenza-like-illness per 100,000 per week in the UK. The solid blue line is the observations, the blue dots the data points to which the model was fitted and the red line the prediction



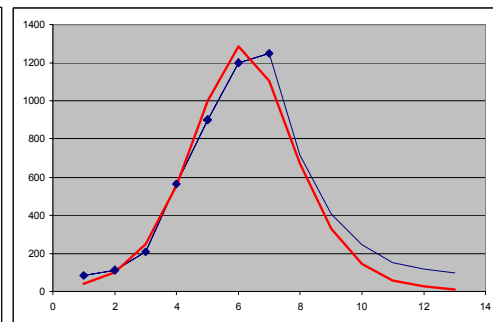
a)



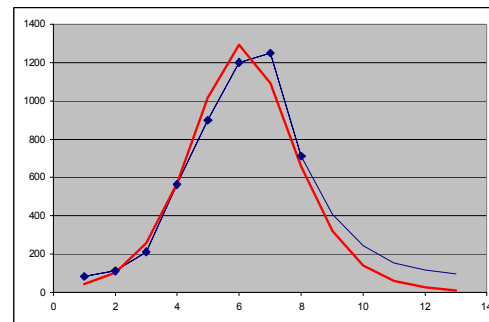
b)



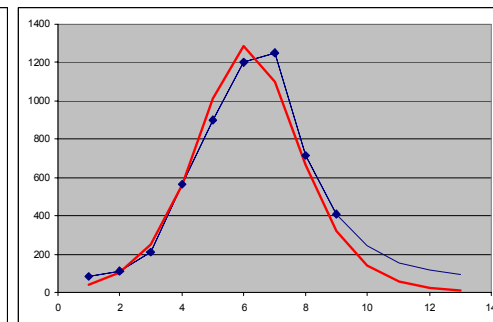
c)



d)



e)



f)