



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



An infectious way to teach students about outbreaks

Íde Cremin¹, Oliver Watson^{1,*}, Alastair Heffernan, Natsuko Imai, Norin Ahmed, Sandra Bivegete, Teresia Kimani, Demetris Kyriacou, Preveina Mahadevan, Rima Mustafa, Panagiota Pagoni, Marisa Sophiea, Charlie Whittaker, Leo Beacroft, Steven Riley, Matthew C. Fisher

Department of Infectious Disease Epidemiology, Imperial College London, UK

ARTICLE INFO

Keywords:

Teaching
Outbreak analysis
Pedagogical tool
Simulation analysis
Network reconstruction

ABSTRACT

The study of infectious disease outbreaks is required to train today's epidemiologists. A typical way to introduce and explain key epidemiological concepts is through the analysis of a historical outbreak. There are, however, few training options that explicitly utilise real-time simulated stochastic outbreaks where the participants themselves comprise the dataset they subsequently analyse. In this paper, we present a teaching exercise in which an infectious disease outbreak is simulated over a five-day period and subsequently analysed. We iteratively developed the teaching exercise to offer additional insight into analysing an outbreak. An R package for visualisation, analysis and simulation of the outbreak data was developed to accompany the practical to reinforce learning outcomes. Computer simulations of the outbreak revealed deviations from observed dynamics, highlighting how simplifying assumptions conventionally made in mathematical models often differ from reality. Here we provide a pedagogical tool for others to use and adapt in their own settings.

1. Introduction

An understanding of non-linear mechanisms that produce chains of transmission underlying outbreaks and epidemics is central to infectious disease epidemiology. Consequently, outbreak simulation is increasingly used as a teaching tool. Approaches to simulating real-life epidemiological data range from simple (e.g. rolling a dice) to complex (e.g. electronic barcode scanning), resulting in the generation of datasets that can then be analysed for epidemiological parameters of interest (Moore, 2017; Hayward, 2017).

Several recent infectious disease outbreaks have received much attention owing to their potentially high impact on public health. For instance, the Middle East Respiratory Syndrome Coronavirus (MERS-CoV), first reported in Saudi Arabia in 2012, currently circulates at a sub-critical level causing sporadic outbreaks (Cauchemez et al., 2014; Cauchemez et al., 2016). The West African Ebola epidemic caused over 28,600 cases between 2014 and 2016 (WHO Ebola Response Team et al., 2016) and its subsequent decline coincided with the emergence of Zika virus transmission in Brazil (Cugola et al., 2016). Following an outbreak of pathogens such as these, initial investigations include estimating the potential for onward transmission (i.e. the basic reproductive number, R_0 (Anderson and May 1992)) and understanding key epidemiological features to inform possible control policies. Such

epidemiological studies have been crucial to understanding and managing the transmission dynamics of important outbreaks such as Severe Acute Respiratory Syndrome (SARS) in 2003, swine flu in 2009, and more recently for MERS, Ebola and Zika (Cauchemez et al., 2014; Riley et al., 2003; Cauchemez et al., 2009; WHO Ebola Response Team, 2014).

Outbreak exercises are an established tool for field epidemiology training. Often employed by national and international public health agencies, different case studies such as foodborne outbreaks are used to train epidemiologists to track down the aetiological agent and then apply appropriate control measures (CDC, 2017a; Public Health England, 2017). Case studies of infectious disease outbreaks are used in classroom teaching to introduce students to key epidemiological concepts such as the incubation period, serial interval, attack rate, and the basic reproductive number (R_0) (Ponder and Sumner, 2009; Betancourt-Bethencourt et al., 2016). Other outbreak exercises are designed to train individuals as case investigation interviewers to increase the public health surge capacity (Gebbie et al., 2007). In addition, virtual exercises and complex computer simulations which incorporate network models have been used as teaching tools (Huang et al., 2010; Neulight et al., 2007; Hsieh et al., 2006). Furthermore, there are now numerous interactive online teaching resources that allow the user to 'solve' outbreaks by playing the epidemiologist (Barber and Stark, 2015; CDC,

* Corresponding author at: MRC Centre for Outbreak Analysis and Modelling, Department of Infectious Disease Epidemiology, Imperial College London, London W2 1PG, UK.
E-mail address: o.watson15@imperial.ac.uk (O. Watson).

¹ Contributed equally.

2017b; Center for Technology in Teaching and Learning, Rice University, 2017). The University of Cambridge provides a selection of online games where the aim of the game is to infect as many people as possible using the characteristics of a specific infectious disease such as influenza (Cambridge Infectious Diseases, 2017). More widely, the potential for virtual role-playing games which exhibit social dynamics, economies and player-objectives resulting in complex networks to be used as model case studies have also been considered (Castronova, 2003; Balicer, 2007).

However, there are few training modules that explicitly use real-time simulated stochastic outbreaks where the participants themselves comprise the dataset they subsequently analyse (Bellan et al., 2012; Darwiche and Bokor, 2016). Participation in a real-time simulated outbreak and analysing the resulting data reinforces and affords a better understanding of important concepts while gaining insight into the process of infectious disease transmission. To meet this need, Bellan et al. developed a pedagogical approach to bridge the divide between classical and dynamical epidemiology (Bellan et al., 2012). This approach involved real-time simulation of a stochastic outbreak among participants, which ranged from university undergraduate students to academic professors, attending an epidemiology training course. We adapted the approach developed by Bellan et al. and used it as a teaching tool for postgraduate epidemiology students by simulating a five-day outbreak among five consecutive student cohorts (2012–2016 inclusive). Here we describe how we conducted the outbreak simulation and how the data are analysed. We provide the teaching materials for others to adapt and use in their own settings.

2. Materials and methods

A stochastic outbreak was simulated over the course of five days. The outbreak was simulated within postgraduate students of five consecutive student cohorts (2012–2016 inclusive) of the MSc in Epidemiology at Imperial College London. The infectious agent was represented by a paper form (see Supplementary Material). Students acquired “dide-disease” (an imaginary infection) by receiving a form from a fellow student and transmitted infection by handing out new forms.

The outbreak was seeded on day one by discretely handing an infection form to three students at random. These paper ‘inocula’ contained a list of instructions. Upon receiving the infection form each student sent an e-mail to the dataset curator to notify that they had been infected. The number of new infections to be transmitted was determined by a random draw from a Poisson distribution with a mean of 1.8, corresponding to the basic reproductive number for the outbreak. Students then printed out the required number of infection forms, located on a shared drive. Students were instructed to do all onward transmitting within 48 h. Whether or not symptoms occurred was determined by a random draw from a Binomial distribution with a mean of 0.8. Symptoms were also reported to the dataset curator via e-mail. Each infected student listed on their infection form the names of those whom they had transmitted to and the time of transmission. Students recovered by placing their form in a ‘recovery box’ in their classroom. Students were instructed to place their form in the recovery box immediately after giving out all their infection forms. Infection was assumed to confer lifelong immunity.

The data were collated in a spreadsheet software (MS Excel®, Microsoft Corporation), and analysed by the students in a practical session to analyse the outbreak. During the practical, students plot the time series of the outbreak, calculate the basic and reproductive numbers as well as other key epidemiological parameters such as latent, incubation and infectious periods. Calculating these and other properties such as the proportion of infections transmitted asymptotically permits discussion regarding ease of control of “dide-disease”. The infection form, practical handout and dataset templates are provided for others to adapt for their own use (see Supplementary Information).

When calculating the effective reproductive number (R_t), we use day of acquisition, regardless of date of onward transmission. Two different cumulative attack rates are calculated: an attack rate based on whether students were infected at all, and a clinical attack rate based only on whether students displayed symptoms. We exclude the seeds when calculating the cumulative attack rate, as the aim is to describe the process in the population of interest rather than the external process that triggered the outbreak via the seeds. Given that the time of acquisition, onset of symptoms (if any), onward transmission and recovery are documented, the latent, infectious and incubation periods can easily be calculated. When estimating the generation time, we use only successful transmissions, rather than all attempted transmissions, as this is the closest analogue to what is observed in real outbreaks. However, in real outbreaks it is the serial interval, rather than the generation time that is typically observed.

For visualisation, analysis and simulation of the outbreak data, an R package and accompanying tutorial was developed and is freely available at <https://github.com/mrc-ide/outbreakteachR>. Visualisation of the temporal dynamics of the outbreak’s transmission network were animated at discrete intervals. Where contact tracing data was missing, the most probable source of an infection was imputed. Data imputation in this way first identified potential source individuals who were infectious at the time of the individual’s infection. The individuals most likely to have caused an additional infection were then extracted. This was based on how many onward infections they had already caused and the basic reproductive number for the outbreak given by a Poisson distribution with mean 1.8. From these individuals a random source individual was chosen.

For comparison, several outbreaks with identical initial conditions were simulated. Each simulation was seeded with three infections, the population size was set equal to the students’ class size and the outbreak was simulated for five days. In common with the paper-based outbreak, the number of secondary infections was drawn from a Poisson distribution with a mean of 1.8. The generation time and recovery period were sampled in two ways:

- (i) by sampling from a Poisson distribution with a mean equal to the observed mean generation time and recovery period respectively,
- (ii) by sampling from the observed distribution of generation times and recovery periods.

For each of the five years, 2000 outbreaks were simulated using a discrete-time approach with hourly intervals in line with the students’ timetable. The simulated time-series was then compared to the observed time-series.

3. Results

Data for five outbreaks are presented (2012–2016) and together illustrate clearly how, despite key parameters being constant, stochastic effects and differences in behaviour from year-to-year affect the outcome of the outbreak. For each year, identical starting conditions were used (i.e. the outbreak was seeded with three infections and an R_0 of 1.8 was assumed). The class size increased over time from 56 students in 2012–84 in 2016. Each outbreak showed different temporal transmission dynamics (Fig. 1), however, each outbreak was broadly summarised by having one major chain of transmission resulting from one of the three seeds. Both 2012 and 2016 had a secondary smaller cluster of infections and a single infection event resulting in a dyad. In 2012 and 2016 the majority of transmission occurred within the first two days in 2016, whereas the majority of transmission occurred between the second and fourth days in 2012. The outbreaks from 2013, 2014 and 2015 were notably slower, with comparatively little transmission until the last two days. We also visualised the infection networks evolving over time (see Supplementary Movies).

A strong time effect was observed within each outbreak, with the

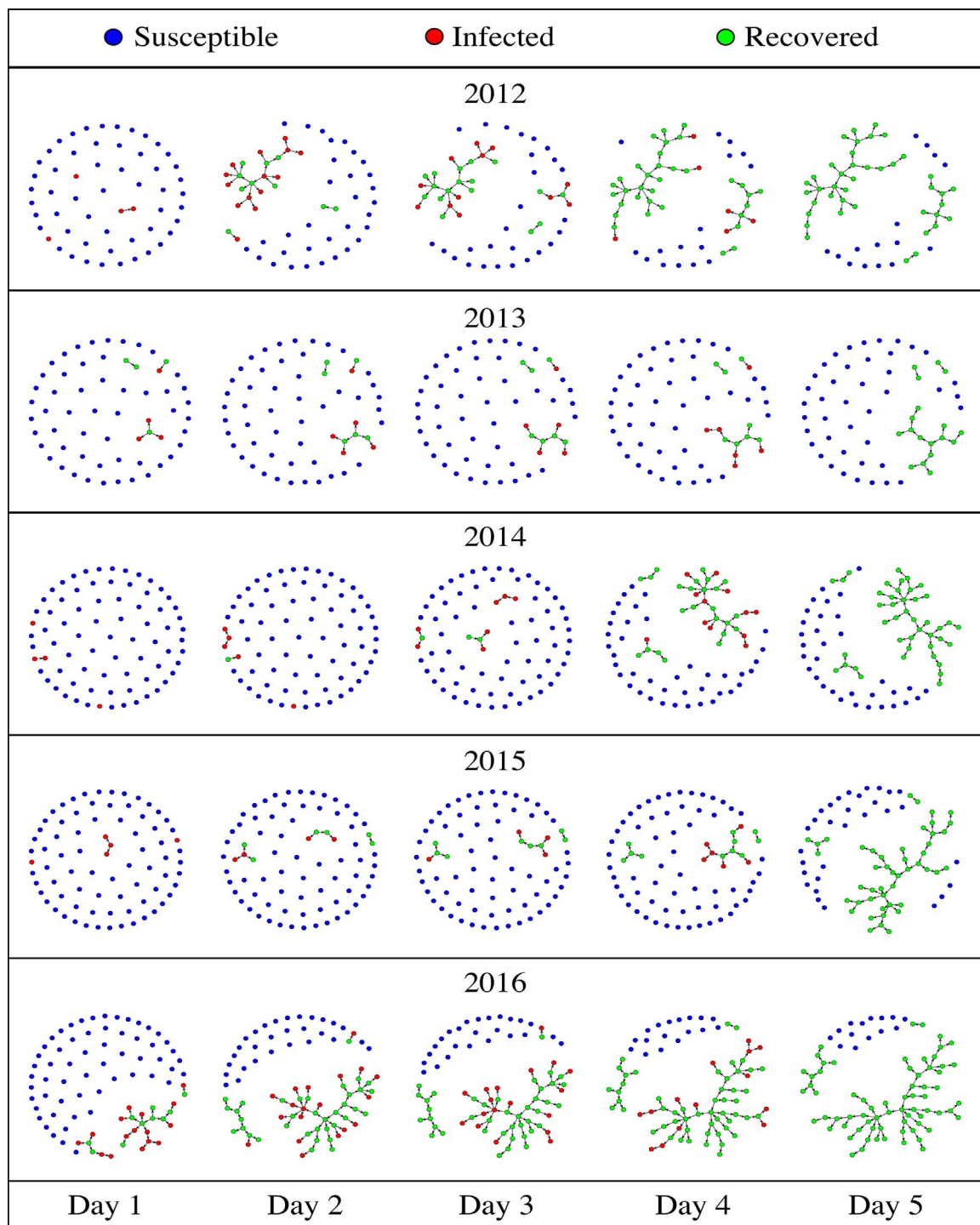


Fig. 1. Daily infection networks. The infection network is represented at daily intervals, with individuals either susceptible (blue), infected (red) or recovered (green). The networks are directional, indicating the source of onward transmission events. Individuals who remained uninfected throughout the five days are represented by unconnected blue nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

majority of transmission observed within short windows of time each day when frequent contacts occurred. This is highlighted within the visualisation of the outbreak, which show long periods of inactivity (Supplementary Movies). This effect can be seen in the strongly multimodal distributions describing the students' behaviour (Fig. 2).

There was substantial variation in the cumulative attack rate, with between 26% – 83% students uninfected at the end of each outbreak. Saturation effects were also observed in each year, with an overall decline in the proportion of infectious contacts that led to successful infections observed from day 1 to day 5, except in 2013 reflecting the

observed low cumulative attack rate for that year.

Computer-simulations of a process similar to the classroom outbreak reveal potential explanations for some of the differences between years. For the 2016 outbreak, the number of infected individuals peaked at 28 on day 2 (Fig. 3a). Simulated numbers of susceptible, infected and recovered individuals broadly capture the magnitude of the outbreak. However, the simulated curves are delayed by approximately two days when the generation and recovery times are sampled from a Poisson distribution (Fig. 3b).

Sampling from the observed distributions provides an improved

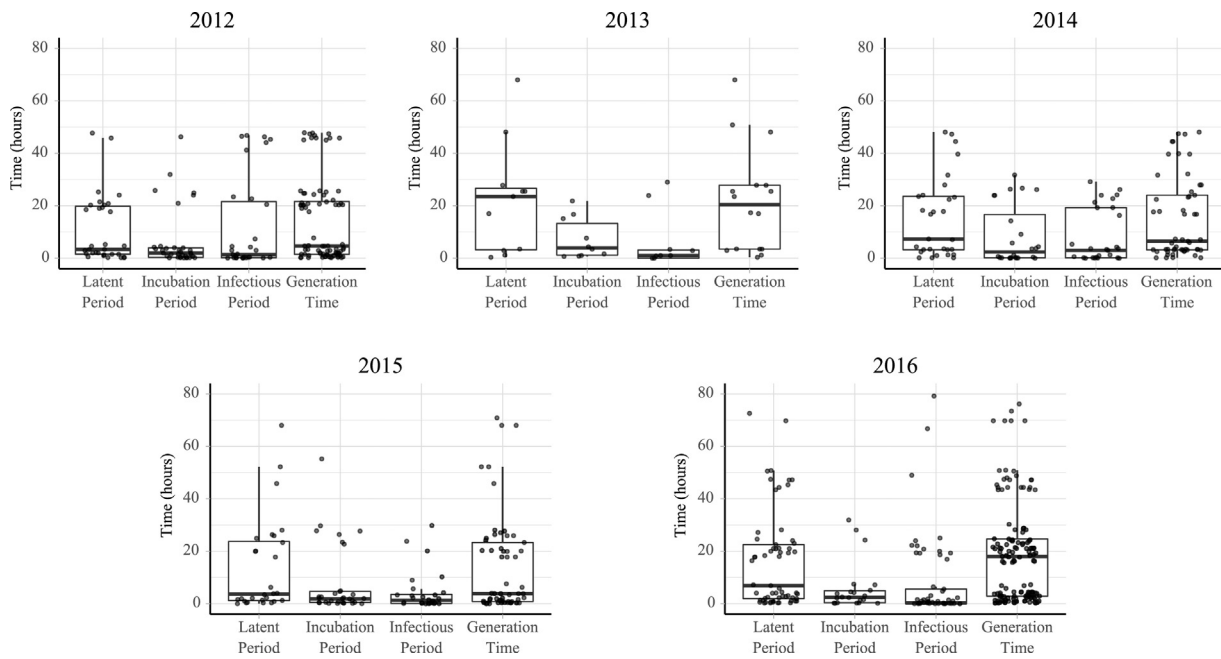


Fig. 2. Epidemiological parameters. Distributions for each year are presented using box and whisker plots, with each box representing the interquartile range (IQR) and the median. The whiskers extend to 1.5 * IQR above and below the upper and lower quartile respectively. The individual data are overlaid as a solid black circles. The plot spreads out data points to make them visible.

representation, with the diurnal patterns of the outbreak being better captured by the simulated mean, though a discrepancy remains (Fig. 3c). Subsequent comparisons between the simulated outbreaks and the observed outbreaks were consequently made with simulations using the empirical re-sampling method (Fig. 4).

Simulations captured the observed outbreak in 2012 well, however, simulations proceeded too quickly for 2013, 2014 and 2015, and failed to accommodate the slow transmission dynamics over the first 3 days followed by bursts of transmission in the final 2 days. In contrast, the simulation method lagged the 2016 outbreak and failed to capture the fast transmission dynamics on the first day.

4. Discussion

We successfully simulated an outbreak of an infectious agent among a class of epidemiology and public health students. Working with a perfectly observed outbreak allowed students to estimate directly key quantities such as the reproductive number (see Supplementary Fig. 1), generation time, and latent, incubation and infectious periods. In addition, key concepts including analysis of transmission networks and how this relates to the data time-series obtained from an outbreak of infection were explored. Comparing each of the five outbreaks provided

students with an awareness of stochastic effects that distinguish outbreaks amongst years.

Simulating an infectious disease outbreak in the classroom is an engaging exercise as students actively participate and comprise the dataset they subsequently analyse. Analysing the outbreak data, in both excel and R, is a useful teaching exercise as it reinforces definitions previously covered in lectures. Active learning has been found to be more effective than traditional exposition-centered (“teaching by telling”) approaches. Active learning can take many forms and has been defined as “engaging students in the process of learning through activities and/or discussion in class. It emphasizes higher-order thinking and often involves group work” (Freeman et al., 2014). A recent meta-analysis of 225 studies found that use of active learning in undergraduate science, technology, engineering and maths courses, improves exam performance (Freeman et al., 2014). Active learning was also found to be particularly beneficial at improving performance on “concept inventories”. Thus, our constructivist approach, whereby students experience and then reflect, lends itself well to this particular teaching exercise, as the objective is to reinforce concepts and definitions.

Panel: How to run the outbreak exercise

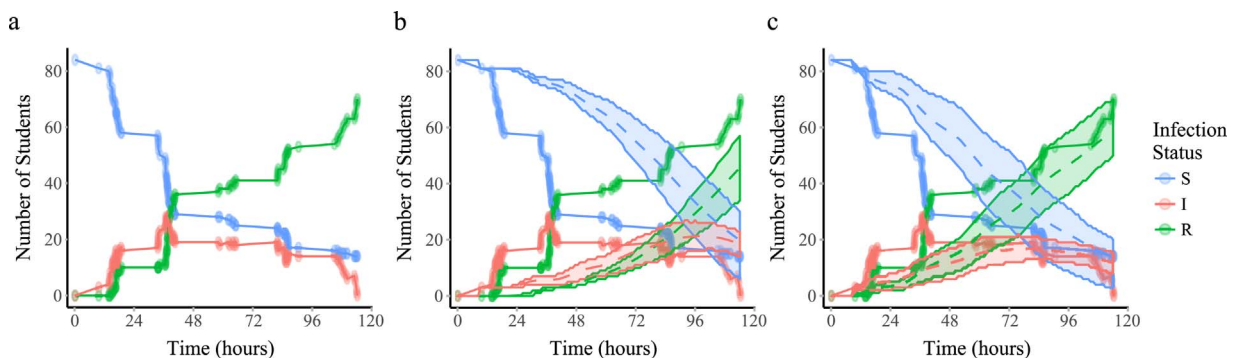


Fig. 3. 2016 outbreak time series. Observed data (a), simulated data with Poisson sampling (b) and simulated data with empirical re-sampling (c). Shaded areas show the interquartile range of the simulated outbreaks, with the simulated mean shown as a dashed line.

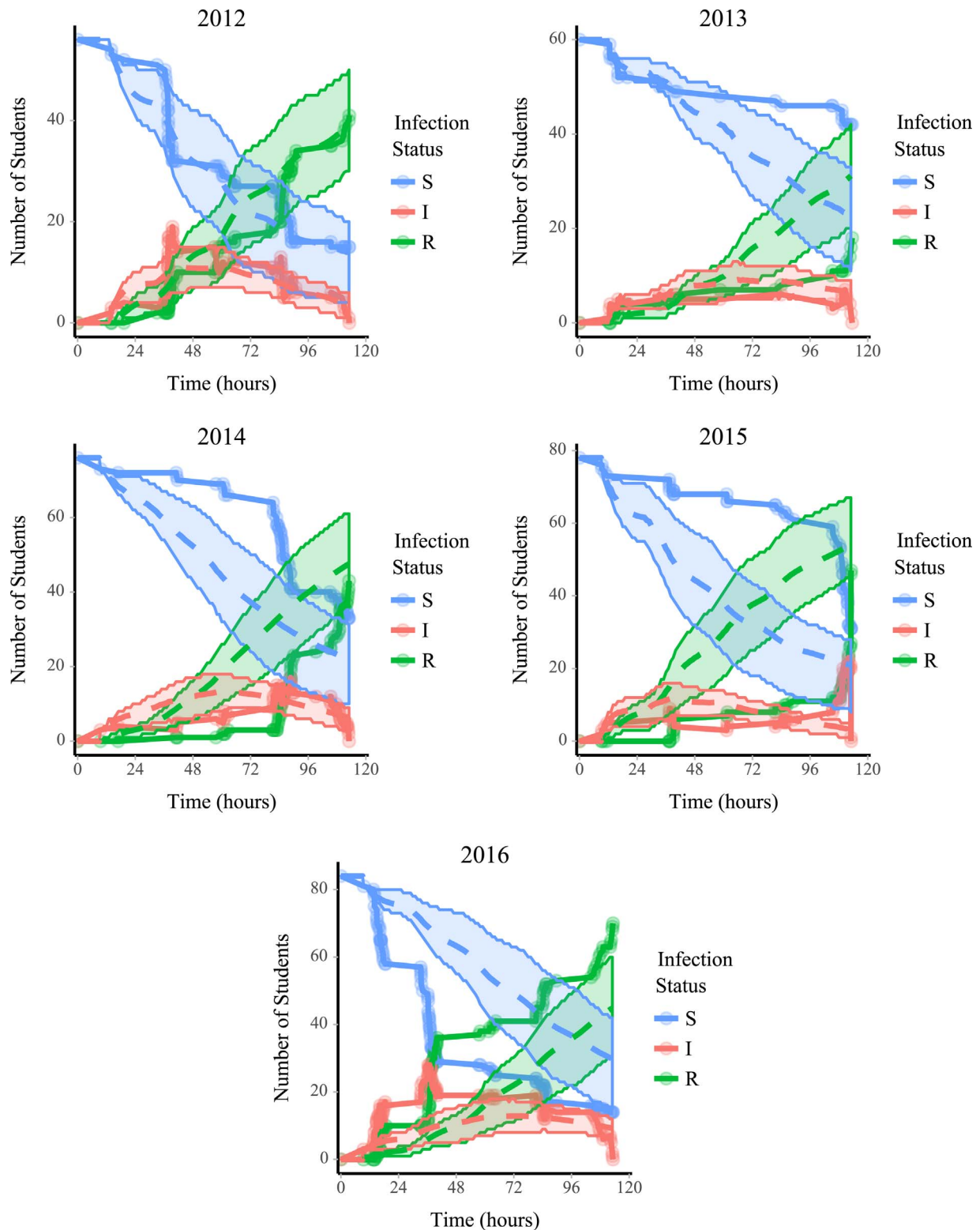


Fig. 4. 2012–2016 Outbreak time series and comparative simulations. Observed outbreak data is presented for each year with solid lines, and the output of 2000 replicate simulations data with empirical re-sampling shown as a dashed line with the interquartile range shown as shaded areas.

What is required?

- An infection form
- A list of all students in the class
- An easily accessible shared drive where the infection form and class list can be accessed
- A means by which students can draw a number from a distribution (e.g. a Poisson distribution)

- A location to store ‘recovered’ infection forms, or use an online form

Practical tips for running the outbreak:

- Finalise a class list of all participating individuals before seeding the outbreak
- Emphasise that not everyone in the class will get infected

- Students can only acquire infection if someone hands them the paper form i.e. they cannot infect themselves by downloading the infection form
- If infected, students need to follow all the instructions carefully
- When transmitting the infection, students must only hand the infection form to someone listed on the class list
- Students need to make sure to put their name on the form and the names of those they have infected

Over the five years, the class size increased, however, this did not appear to affect the ability for an outbreak to be sustained, with the smallest and largest class sizes (2012 and 2016 respectively) yielding the highest cumulative attack rate. Importantly, the outbreak never experienced stochastic fadeout across the five days of the outbreak, with new infections still occurring on day five in all years. Overall, the outbreaks presented with short incubation and infectious periods, but comparatively longer and more variable latent periods. The increased variance in the latent period, and subsequently the generation times, was typified by the strongly diurnal patterns in the behaviour of the students, with most events occurring within their timetable. It is interesting to observe the decline in transmission exhibited on the third day of the 2012 and 2016 outbreak, which could reflect the absence of afternoon classes on Wednesday. However, this effect is hard to be definitive about, with the outbreaks in 2013–2015 exhibiting low levels of transmission over the first three days.

The human simulation is different from a computer simulation of the same system. Sampling the generation time and recovery time directly from the observed distributions, rather than sampling from a Poisson distribution with a mean based on the observed data, markedly improves the closeness of the simulated outbreak to the observed outbreak. This effect can be explained by the multi-modal nature of the distributions seen within the outbreaks, reflecting the strong diurnal time-effects observed, which weakens the suitability of a unimodal distribution such as the Poisson distribution.

Although the simulated outbreaks largely failed to capture the 2013–2016 outbreaks, we are able to use these discrepancies as a teaching exercise, through both highlighting the weaknesses of the mathematical model we used, and to highlight phenomena within the observed outbreaks that are typical within real outbreaks. Firstly, the simulation lagged behind the observed outbreak in 2016, due to the large amount of transmission that occurred on day 1 (see Supplementary Fig. 2). The lengthening generation times in 2016 could represent a degree of waning enthusiasm from the students, which would invalidate the time-independent sampling used within the simulations. Secondly, the long latent times exhibited at the beginning of the 2013–2015 outbreaks, combined with a burst of shorter generation times at the end of these outbreaks, yields another time-dependant effect upon the distribution of generation times. These effects, combined with the lower population mean number of secondary infections observed in these years, offer a potential explanation for the poor predictions afforded by the simulations. Additionally, these effects also demonstrate a teaching point concerning the potential for time dependent effects to occur in outbreaks, representative of real outbreaks (Chowell et al., 2015; Nishiura, 2010), and the difficulties they present in mathematical modelling. The slower transmission in 2013–2015 is illustrative of commonly observed stuttering chains of transmission that occur prior to sustained chains of transmission (Dibble et al., 2016).

This teaching exercise provides a flexible template which can easily be modified for other audiences, such as public-health professionals and policy makers. For example, a more advanced class with prior experience of epidemiological modelling could calculate clustering coefficients and path lengths to assess deviation from a random mixing assumption. For a larger class, students could calculate R_0 from the final class size and growth rate, and compare to the observed R_0 . The outbreak itself could be adapted to run over a shorter or longer time period (by altering R_0 or time in which onward transmission should occur) and could also be elaborated by looking at the impact of interventions to

reduce transmission, such as vaccination, whereby some students are immune to infection. As noted by Bellan et al., the outbreak could be elaborated to include multiple pathogen strains with different properties (such as infectiousness, symptomatic proportion and infectious period).

These data collected over five years highlighted a number of limitations within the methodology of the practical. Firstly, the paper ‘inocula’ form was largely well completed by students, however, there were a number of missing data. A large degree of the missing information could be imputed once all the data was collected, but there is no guarantee that participants will always fill the form in fully or accurately. This represents a key issue in data collection (Owada et al., 2016), one which is often addressed with the design of more intuitive and robust surveillance tools and software. To address this, an online form, utilising both automated data fields and required data entry fields, has been designed for use in the future as an alternative to the paper form (see Supplementary Form). It is hoped that this will also strengthen participant engagement and will help prevent students ‘forgetting’ about the outbreak, which appeared to occur on a few occasions when observed latent periods were greater than 48 hours (see Supplementary Fig. 3). Secondly, the initial practical spreadsheet was designed to emphasise how epidemiological quantities such as the generation time and latent periods are calculated, as well as for plotting the outbreak time series. While this was helpful for introducing these calculations, we found that this analysis was overly time consuming and reduced the ability to explore additional epidemiological concepts such as the cumulative attack rate and saturation effects. Consequently, we redesigned the spreadsheet to automate specific sections of the analysis, thereby improving the pedagogical experience (see Supplementary Datasets).

In conclusion, we present five years of data in response to the approach presented by Bellan et al. (Bellan et al., 2012). Furthermore, we have introduced several new features to extend the pedagogical insights afforded by this approach, in particular the incorporation of a networks component, and the focus on comparing modelled results to the data itself in order to demonstrate both the strength and potential limitations of mathematical models. The approach could readily be modified for other settings, for example over a longer or shorter period of time and is a useful method to engage students and reinforce key epidemiological concepts in a “real-world” setting. In addition, the R package is freely available for extension and adaptation to consider alternative epidemiological questions and analysis and provides a framework to introduce discrete-time models.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

None.

Acknowledgements

We thank Claire McCormack and Lucy Li for help in collating the student data sets. We also thank the MSc student cohorts 2012–2016 for taking part in the practical. For funding, we thank the Medical Research Council (UK, Project MR/K000373/1, MR/J008761/1, MR/K010174/1B), the Wellcome Trust (UK, Project 109312/Z/15/Z, 105268/Z/14/Z), the Natural Environment Research Council (UK, Project NE/K014455/1), the Bill and Melinda Gates Foundation (OPP1067133) and the Indonesia Endowment Fund for Education (0000530/KK/M/3/lpdp2015).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <https://doi.org/10.1016/j.epidem.2017.12.002>.

References

Anderson, R.M., May, R.M., 1992. *Infectious Diseases of Humans: Dynamics and Control*. Vol. 28. Oxford university press, Oxford.

Balicer, R.D., 2007. Modeling infectious diseases dissemination through online role-Playing games. *Epidemiology* 18, 260–261.

Barber, N.C., Stark, L.A., 2015. Online resources for understanding outbreaks and infectious diseases. *CBE Life Sci. Educ.* 14, fe1.

Bellan, S.E., Pulliam, J.R.C., Scott, J.C., Dushoff, J., 2012. MMED Organizing Committee. How to make epidemiological training infectious. *PLoS Biol.* 10, e1001295.

Betancourt-Bethencourt, J.A., Acao-Francois, L., González-Ronquillo, Y., 2016. Entrenamiento analítico en investigaciones epidemiológicas para estudiantes de medicina. *REE* 20, 1.

Cambridge Infectious Diseases, 2017. *Games*. ([Internet]. [cited 10 December 2017]. Available: <http://www.infectiousdisease.cam.ac.uk/schoolzone/games>).

Castronova, E., 2003. On virtual economies. *Int. J. Comput. Game Res.* 3 (2). Available: <http://www.gamestudies.org/0302/castronova>. (Accessed 10 December 2017).

Cauchemez, S., Donnelly, C.A., Reed, C., Ghani, A.C., Fraser, C., Kent, C.K., et al., 2009. Household transmission of 2009 pandemic influenza A (H1N1) virus in the United States. *N. Engl. J. Med.* 361, 2619–2627.

Cauchemez, S., Fraser, C., Van Kerkhove, M.D., Donnelly, C.A., Riley, S., Rambaut, A., et al., 2014. Middle east respiratory syndrome coronavirus: quantification of the extent of the epidemic, surveillance biases, and transmissibility. *Lancet Infect. Dis.* 14, 50–56.

Cauchemez, S., Nouvellet, P., Cori, A., Jombart, T., Garske, T., Clapham, H., et al., 2016. Unraveling the drivers of MERS-CoV transmission. *Proc. Natl. Acad. Sci. U. S. A.* 113, 9081–9908.

CDC, 2017a. *Solve The Outbreak*. ([Internet]. [cited 10 December 2017]. Available: <https://www.cdc.gov/mobile/applications/sto/web-app.html>).

CDC, 2017b. *Epidemiologic Case Studies*. ([Internet]. [cited 10 December 2017]. Available: <https://www.cdc.gov/epicasestudies/index.html>).

Center for Technology in Teaching and Learning, Rice University, 2017. *MEDMYST — Explore Science — Infectious Diseases*. ([Internet]. [cited 10 December 2017]. Available: <http://medmyst.rice.edu/>).

Chowell, G., Viboud, C., Hyman, J.M., Simonsen, L., 2015. The western africa ebola virus disease epidemic exhibits both global exponential and local polynomial growth rates. *PLoS Currents* 7.

Cugola, F.R., Fernandes, I.R., Russo, F.B., Freitas, B.C., Dias, J.L.M., Guimarães, K.P., et al., 2016. The Brazilian Zika virus strain causes birth defects in experimental models. *Nature* 534, 267–271.

Dibble, C.J., O’Dea, E.B., Park, A.W., Drake, J.M., 2016. Waiting time to infectious disease emergence. *J. R. Soc. Interface* 13, 20160540.

Darwiche, H., Bokor, J.R., 2016. Ebola epidemic using current events to teach authentic inquiry science. *Am. Biol. Teach.* 78, 190–197.

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P., 2014. Active learning increases student performance in science, engineering, and mathematics. *PNAS* 111 (23), 8410–8415.

Gebbie, E.N., Morse, S.S., Hanson, H., McCollum, M.C., Reddy, V., Gebbie, K.M., et al., 2007. Training for and maintaining public health surge capacity: a program for disease outbreak investigation by student volunteers. *Public Health Rep.* 122, 127–133.

Hayward, J., 2017. *Outbreak of Pal Flu Observed at Science Museum*. ([Internet]. [cited 10 December 2017]. Available: http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/medicine/publichealth/mrscentreforoutbreakanalysisandmodelling/newssummary/news_19-6-2013-17-29-45).

Hsieh, J.-L., Sun, C.-T., GY-M, Kao, Huang, C.-Y., 2006. Teaching through simulation: epidemic dynamics and public health policies. *Simulation* 82 (11), 731–759.

Huang, C.-Y., Tsai, Y.-S., Wen, T.-H., 2010. Simulations for epidemiology and public health education. *J. Simulation* Palgrave Macmillan UK 4, 68–80.

Moore, M.J., 2017. *Bacterial Games at the Big Bang Fair*. ([Internet]. [cited 10 December 2017]. Available: http://www3.imperial.ac.uk/newsandeventspggrp/imperialcollege/medicine/newssummary/news_19-3-2013-13-28-18).

Neulight, N., Kafai, Y.B., Kao, L., Foley, B., Galas, C., 2007. Children’s participation in a virtual epidemic in the science classroom: making connections to natural infectious diseases. *J. Sci. Educ. Technol. Kluwer Acad. Pub.* 16, 47.

Nishiura, H., 2010. Time variations in the generation time of an infectious disease: implications for sampling to appropriately quantify transmission potential. *Math. Biosci. Eng.* 7, 851–869.

Owada, K., Eckmanns, T., Kamara, K.-B.O., Olu, O.O., 2016. Epidemiological data management during an outbreak of ebola virus disease: key issues and observations from Sierra Leone. *Front Public Heal* 4, 2–5.

Ponder, M.A., Sumner, S., 2009. Use of case studies to introduce undergraduate students to principles of food microbiology, molecular biology, and epidemiology of food-borne disease. *Biochem. Mol. Biol. Educ.* 37, 156–163.

Public Health England, 2017. *Health Protection and Epidemiology Training – Course List*. ([Internet]. [cited 10 December 2017]. Available: <https://www.phe-protectionservices.org.uk/cms/article.php?article=1656&details=true&course=65>).

Riley, S., Fraser, C., Donnelly, C.A., Ghani, A.C., Abu-Raddad, L.J., Hedley, A.J., et al., 2003. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science* 300, 1961–1966.

WHO, 2014. Ebola virus disease in west africa — the first 9 months of the epidemic and forward projections. *N. Engl. J. Med.* 371, 1481–1495.

WHO, Agua-Agum, J., Allegranzi, B., Ariyarahaj, A., Aylward, R.B., Blake, I.M., et al., 2016. After ebola in west Africa—Unpredictable risks, preventable epidemics. *N. Engl. J. Med.* 375, 587–596.