Estimation of Reproduction Number of Influenza in Greece using SIR Model

Cleo M. Skarpeti[1] and Dr. Michael G. Skarpetis[2]
[1]Hellenic-American Educational Foundation, Athens, Greece
[2]National and Kapodistrian University of Athens, General Department, Athens, Greece

SUMMARY

Infectious disease models are regularly used by epidemiologists to predict the course of diseases. However, the accuracy of such models remains hotly debated. In this study, we developed an algorithm to estimate the contact rate and the average infectious period of influenza using a Susceptible, Infected, and Recovered (SIR) epidemic model. The parameters in this model were estimated using data on infected Greek individuals collected from the National Public Health Organization. The data interval investigated spans 13 weeks of 2019 and the first 8 weeks of 2020. After estimating the nonlinear SIR epidemic model's parameter values, we computed the theoretical influenza reproduction number and simulated the evolution of the three SIR variables over time. We hypothesized that the reproduction number would be greater than one, which means that influenza is not like other simple diseases but is an epidemic. Our model labeled influenza as an epidemic with a basic reproduction value greater than one.

INTRODUCTION

In the 1960s, timely eradication of infectious diseases was a seemingly achievable goal. Nevertheless, contagious diseases are still a severe problem since they can lead to long-term suffering and mortality, worldwide but especially in developing countries which have a lack of resources, access to sanitation facilities, and healthcare education. As a result, pandemics or epidemics remain neglected or inadequately tackled in developing countries, which has detrimental repercussions to the population such as increased death rate. Underdeveloped countries are less likely than developed countries to combat diseases by deploying effective vaccinations to decrease transmission (1). Although Greece is not a developing country the healthcare system should be improved, because it lacks resources due to the financial crisis of 2010 (2).

Through viral evolution and microbe adaption to new circumstances and environments, diseases consistently emerge and reemerge. In particular, the influenza virus is a pathogen that evolves perpetually, usually appearing as a new variant every year. The influenza virus can adapt, change and evolve genetically, which is why it is found in various species. Viruses evolve by some basic mechanisms

such as RNA reassortment or genetic shift and genetic drift. In reassortment, two genetically different influenza viruses infect the same cell in a human organism and their gene segments swap and transfer genetic information (3). Mixing this information can create one hybrid virus. In genetic drift, mutational differences can develop in the genes of the virus through changes in its surface proteins. Genetic shift happens rarely. However, the difference is that the virus that originally infected animals now can infect human beings as well. A virus that undergoes drift and shift is probably the only case that can cause a global influenza pandemic associated with harmful implications (3). According to the World Health Organization, the average number of annual influenza cases is approximately 1 billion, and 3–5 million of which are severe with a death toll of 500,000 (4). Apart from the high mortality rate, influenza also provokes financial damages worldwide. Indicatively, the total financial loss of annual influenza epidemics in the United States is approximately $87.1 billion (5).

Though influenza outbreaks consistently occur every year, it is not known when the disease will emerge, the way it will spread among populations, how many deaths will occur, and which part of the population should be vaccinated. The answers to these questions can be approached by making use of experimental tools, such as mathematical models that enable us to test theories by simplifying reality.

To control the spread of a disease, we need to make general predictions and forecasts to prevent it from becoming an epidemic that will infect a substantial number of individuals within a community, population, or region in a short period of time. The Susceptible (S), Infected (I), and Recovered (R) (SIR) epidemic model first presented by Kermack and McKendrik in 1927 divides the population into three basic compartments (S, I, R) and studies the evolution of these quantities in time via nonlinear differential equations (6). Susceptible represents the part of the population which is at risk of becoming infected. Infected denotes the members of the population who are infectious and might transfer the disease to susceptible individuals. Finally, recovered is representative of the part of the population that was previously infected but recovered and has immunity to the disease. The SIR mathematical model indicates changes in the population of each of these three basic compartments based on two parameters, the contact rate and the average infectious period. The contact rate denotes the average number of contacts with susceptible individuals that might infect them.
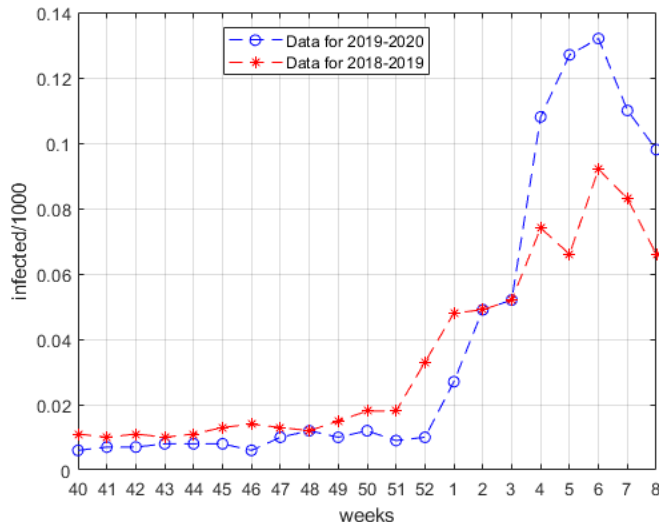
Figure 1. Number of cases of influenza infections per 1000 visits per week. Data taken from the National Public Health Organization regarding influenza infections per 1000 visits per week in Greece during the 2018-2019 (red line) and 2019-2020 (blue line) influenza season. Weeks represent week number in 2019 (40-52) and 2020 (1-8).

The average infectious period denotes the time period during which an infected individual is likely to transmit the disease to a susceptible one.

The aforementioned SIR nonlinear dynamics that model the spread of influenza in Greece are presented with the following differential equations (6,7):

$$\frac{dS(t)}{dt} = -\beta S(t)I(t) \qquad (1)$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) \qquad (2)$$

$$\frac{dR(t)}{dt} = \gamma I(t) \qquad (3)$$

where S(t), I(t), and R(t) denote the susceptible, infected, and recovered individuals in a population infected from influenza $S(t) + I(t) + R(t) = N$ where N the total number of the individuals that belong to this population). The parameter β>0 denotes the contact rate and 1/γ denotes the average infectious period.

The basic reproduction number, R_0, is the predicted number of susceptible individuals that an infected one can infect. A value of R_0 greater than 1 is indicative of the persistence of the infection in the examined population, above its baseline, whereas a value of R_0 less than 1 is indicative of the eventual eradication of the disease. In an outbreak, it is of utmost importance to apply strategies and methods to control this number and limit it. Mitigation strategies include household isolation, school and store closures, vaccination, etc. (8). This is the most famous model in epidemiology and is used by many researchers (7-16).

In this study, we developed an algorithm in order to estimate the main parameters of an SIR epidemic model, which are the contact rate and the average infectious period. First, we collected data on infected Greek individuals from
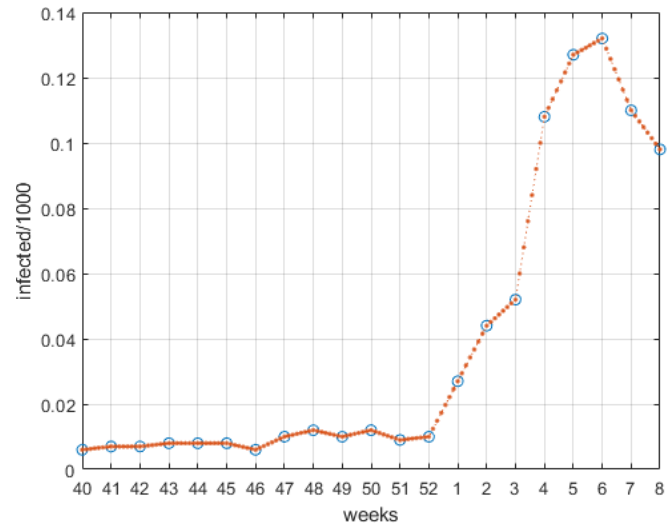

Figure 2. Number of cases of influenza infections per 1000 visits per day. Weekly case numbers underwent interpolation in MATLAB to generate visits per day.

the National Public Health Organization during the last 13 weeks of 2019 and the first 8 weeks of 2020 (17). The data is enriched via the interpolation function in MATLAB in order to be suitable for identification. Using the combined collected data, the parameter values are estimated and identified using grey box modelling methods (18). Finally, we presented the time evolution of the SIR model variables using simulation results.

The aim of this procedure is to answer the following question: Is influenza considered to be an epidemic in Greece in the specific time period 2019–2020 based on its reproduction number, and how does this time period compare to the same time period in 2018-2019? Using limited data from infected individuals, we determined that influenza is an epidemic in Greece, since the computed reproduction number, based on the parameter predictions, was indeed greater than 1 and the predicted infected individuals during the last week under investigation was greater than zero.

RESULTS
Data Collection and Estimation For 2019–2020

We used the data of the last 13 weeks of 2019 and the first 8 weeks of 2020 collected from National Public Health Organization (13) to visualize the number of cases of influenza infections per 1000 visits per week (**Figure 1**). We also applied the 1-D interpolation function in MATLAB, in order to enrich the data from weeks to days to create necessary data for the identification algorithm (18) (**Figure 2**).

Applying data after interpolation and using the proposed procedure (see Methods), the parameters of the epidemic model were identified using grey box modelling methods (18). Using the nonlinear SIR model equations 1–3, the differential equation-based model structure is specified. The unknown model parameters (contact rate and average infectious period) were estimated via nonlinear least squares

minimization techniques. Finally, the Euler numerical method was applied to solve these differential equations. The fit to estimation data was 84.53% and the final prediction error was 4.79567 x 10^{-5} (**Figure 3**). These values are quite satisfactory considering the form of the initial data. These estimation results were independent from the initial values of the predicted parameters.

The predicted parameter values of the SIR model are:

$\beta=0.2044$, $\gamma=0.1113$

Using the aforementioned parameters and replacing them in equations 1–3, we present the evolution of the three variables in time (S, I, & R) (**Figure 4**). The evolution of the influenza infection in Greece considering the average infectious period as nearly $1/0.1113 \sim 9$ days, which means that an infected individual can transmit the virus for 9 days. The contact rate is $\beta=0.2044$ which means that an average infected individual has the potential to infect 0.2044 susceptible per day or equivalently 1 susceptible per 5 days since $1/0.2044 \sim 5$.

Since there was an increase in the number of infected people as well as the recovered ones, there was an eventual decrease in the number of susceptible individuals. We noticed that the blue line (S) did not hit zero and the green line (R) never reached 1, which meant that not every single individual in this particular population had been infected so far. Furthermore, due to the fact that the reproduction number $R_0 = \beta/\gamma=1.8360>1$, the disease can be characterized as an epidemic that eventually invades the population (19). It is worth mentioning that since influenza turned out to be an epidemic, the number of infected individuals is always greater than zero. The predicted infected individuals during the last week under investigation is 0.7447 .

Using 2018 data (last 13 weeks of 2018 and the first 8 weeks of 2019) of infected individuals collected from the national public health organization (**Figure 1**), we estimated the parameter values of the SIR model as follows: $\beta=0.1254$ and $\gamma=0.0791$, which was used to compute the basic reproduction number, $R_0 = \beta/\gamma=1.585>1$. Since this reproduction number is lower than the estimated reproduction number of 2019-2020, we conclude that the influenza outbreak was more severe in 2019-2020.

DISCUSSION

In this paper, the data from only one of the three basic involved variables (i.e. infected individuals) was used to estimate the value of the parameters β and γ of the SIR model and compute the basic reproduction number. The results concern the outbreak of influenza in 2019–2020. The initial hypothesis, that influenza is an epidemic in Greece, is validated since the basic reproduction number is greater than 1. Influenza in Greece is characterized by a rapid spread, with approximately 13.2% of the population examined during the 6th week of 2020 becoming infected (17). The model can be further validated when more data becomes available. This will support that the forecasted evolution is the same as the real one.

We considered several limitations of the SIR model used. The model assumes that everyone has the same probability of being in contact with infected individuals, which does not represent society since not all civilians come in contact with the same number of people. Moreover, the model considers the total population to be constant. As a result, births and deaths do not affect the outcome. Another limitation is that infected individuals are considered to have an equal transmission rate regardless of official diagnosis or quarantine status. Furthermore, more complex versions of the SIR model exist that consider tunable inputs, such as vaccination, educational campaigns, and immunity. Such variables affect the results and are closer to a real social
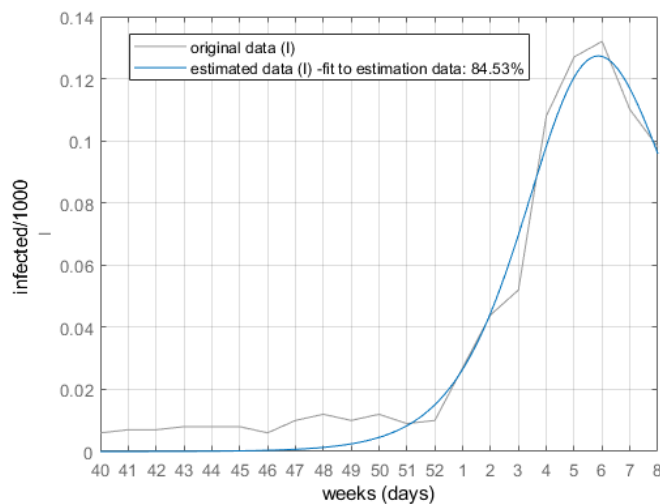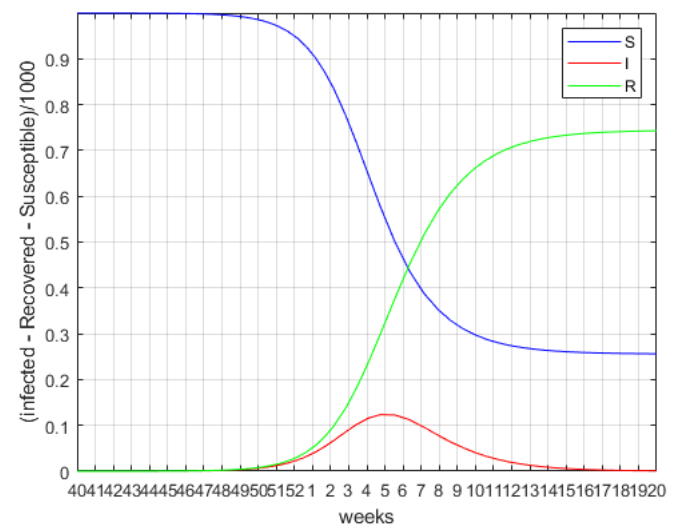

Figure 3. Predicted response of infected individuals. Comparison of original data (grey line) and estimated data (blue line). The fit to estimation data was 84.53% and the final prediction error was 4.79567 x 10^{-5}.

Figure 4. Simulation of SIR predicted epidemic dynamics. Model of the 2019-2020 influenza season in Greece. Weeks represent week number in 2019 (40-52) and 2020 (1-20). S, susceptible (blue); I, infected (red); R, recovered (green).
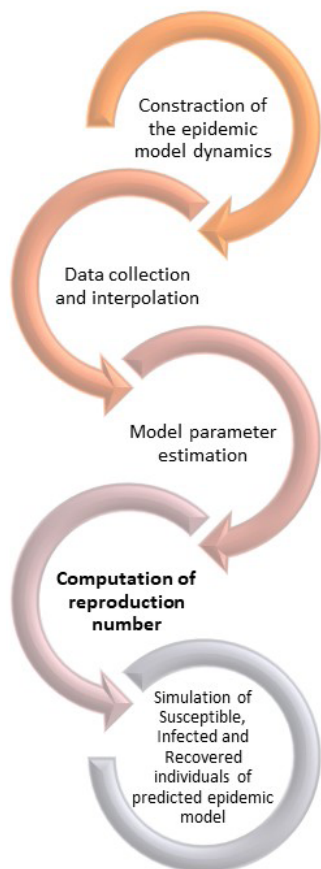
Figure 5. Flow chart for the computation of reproduction number for 2019-2020 influenza in Greece.

structure. These complex models can propose the most optimal solution to limit the repercussions of an epidemic or even pandemic. Nevertheless, the SIR model is quite useful and acceptable despite these limitations since it is accurate enough to enable the researchers to easily approximate infectious disease evolution. Using this particular model, with only a few measurements of infected individuals, general predictions about the outbreak can be made enabling experts to take action fast to limit the implications as soon as possible.

According to the results of the estimated basic reproduction numbers of influenza in 2018–2019 and 2019–2020, we understand that influenza outbreak in 2019–2020 was more severe than the outbreak in 2018–2019 (7.4% of the examined population during the 6th week of 2019 was infected). According to the estimated reproduction numbers, we observed that in 2019-2020 the predicted number of susceptible individuals that an infected one can infect is 1.836, as opposed to 2018-2019 which was 1.585. This comparison highlights that the influenza virus in 2019–2020 was transmitted more rapidly and ending up infecting a quite large number of individuals, as opposed to the previous year. Further research into the reasons why these differences exist year to year could lead to development of better strategies to mitigate and prevent further epidemics, such as influenza, in the future.

MATERIALS & METHODS

In order to compute the reproduction number of influenza, data that represent the estimated number of cases of influenza infections per 1000 visits per week, was collected from the National Public Health Organization. According to the National Public Health Organization, the number of cases per 1000 visits was estimated by weighting the country's permanent population by geographical department and urbanity (17). This information comes from the network of private doctors, the network of doctors Health Centers and the network of doctors of Health Units. The data was enriched by the interpolation function in MATLAB so as to concern influenza infections per 1000 visits per day. From the interpolated dataset, the nonlinear model with unknown parameters was identified and the unknown model parameters (namely contact rate and average infectious period) were estimated using the nonlinear SIR model and suitable functions in MATLAB. The identification was accomplished using grey box modelling methods and nonlinear least squares minimization techniques. Finally, the differential equations were solved using the Euler numerical method.

Based on these estimations, the basic reproduction number was computed. After the parameter estimation, simulation was performed. The outbreak of influenza for this specific time period was visualized by showing the evolution of the three basic variables of the model in time. The aforementioned procedure is illustrated in **Figure 5**. The proposed methodology was accomplished using a Matlab code. Full code can be provided upon request.

REFERENCES

1. Peabody, John W., and Mario M. Taguiwalo. "Improving the Quality of Care in Developing Countries: Disease Control Priorities in Developing Countries." Oxford University Press, 2006.
2. Economou C, Kaitelidou D, Karanikolos M, Maresso A. "Greece: Health System Review." Health Syst Transit., vol. 19, no. 5, 2017.
3. Balgopal, Meena M. and Bondy, Cindi. "Antigenic Shift and Drift" Science teacher (Normal, Ill.), Jan. 2011, doi: 10.1038/283524b0
4. Clayville, Lisa R. "Influenza update: a review of currently available vaccines." P & T : a peer-reviewed journal for

formulary management, vol. 36, no. 10, 2011.

5. Mao, Liang, et al. "Annual Economic Impacts of Seasonal Influenza on US Counties: Spatial Heterogeneity and Patterns." International Journal of Health Geographics, vol. 11, no. 1, 2012, doi:10.1186/1476-072x-11-16.

6. Kermack, WO and McKendrick, AG. "A contribution to the mathematical theory of epidemic." Proceeding of the Royal Society of London, Mathematical and Physical Sciences, vol. 115, 1927.

7. Merler, S, Ajelli, M, Pugliese, A, and Ferguson, NM. "Determinants of the spatiotemporal dynamics of the 2009 H1N1 pandemic in Europe: implications for real-time modelling." PLoS Comput Biol, 2011.

8. Biggerstaff, Matthew, et al. "Estimates of the Reproduction Number for Seasonal, Pandemic, and Zoonotic Influenza: a Systematic Review of the Literature." BMC Infectious Diseases, vol. 14, no. 1, 2014, doi:10.1186/1471-2334-14-480.

9. Liu, Lei. "The spread model of H1N1 influenza." Journal of Gansu Sciences, vol. 22, 2010.

10. Anderson, RM. ed. "Population Dynamics of Infectious Diseases." Chapman and Hall, London, 1982.

11. Hethcote, Herbert W. "The Mathematics of Infectious Diseases." SIAM Review, vol. 42, no. 4, 2000.

12. Brauer, F and Castillo-Chávez, C. "Mathematical Models in Population Biology and Epidemiology." Springer, 2012.

13. Bailey, NTJ. "The Mathematical Theory of Infectious Diseases." Grin, London, 1975.

14. Li, Gui-Hua, and Yong-Xin, Zhang. "Dynamic Behaviors of a Modified SIR Model in Epidemic Diseases Using Nonlinear Incidence and Recovery Rates." Plos One, vol. 12, no. 4, 2017.

15. Zhang, Xiang, HE, Yilin, Yang, Haiyu. "A retrospective evaluation of protective effect of seasonal influenza vaccine and influenza A(H1N1) vaccine." Chinese Journal Disease Control Prevention, vol. 16, 2012.

16. Allen, LJS. "Some discrete-time SI, SIR, and SIS epidemic models." Mathematical Biosciences, Elsevier, vol. 124, Issue 1, 1994.

17. National Public Health Organization, Weekly Report Epidemiological Surveillance of Influenza Week 8/2020 (17–23 February 2020)

18. Bohlin, Torsten P. "Practical Grey-box Process Identification: Theory and Applications." Springer Science & Business Media, 2006.

19. Chaturvedi, Ojaswita. "MATLAB - A Successful Tool for Epidemic Modelling Simulation." International Journal for Research in Applied Science and Engineering Technology, vol. 5, no. 8, 2017.