Procedia Computer Science

Procedia Computer Science 101, 2016, Pages 282 – 291

YSC 2016. 5th International Young Scientist Conference on Computational Science

Influenza outbreaks forecasting in Russian cities: is Baroyan-Rvachev approach still applicable?

Vasiliy N. Leonenko[1], Yulia K. Novoselova[2], and Karen M. Ong[3]

[1] ITMO University, Saint Petersburg, Russian Federation
vnleonenko@yandex.ru
[2] ITMO University, Saint Petersburg, Russian Federation
juliaeternity@gmail.com
[3] New York University School of Medicine, New York, the United States of America
karen.ong@med.nyu.edu

Abstract

In the current work we assess the accuracy of the outbreak peak prediction expected in a fixed Russian city with the help of the populational SEIR models fitted to the past local outbreaks happened earlier in the country during the same season. This approach was successfully used in USSR by Baroyan and Rvachev to predict flu outbreaks throughout 1970's. However, in early 1980's it became increasingly inaccurate and due to that the approach was not applied since. Our aim is to find out, using the long-term data on acute respiratory infections incidence for Moscow, Saint Petersburg and Novosibirsk as an input, to what extent this method of outbreak peak prediction is applicable nowadays, and discuss the methods of increasing its accuracy. We found that SEIR population modeling is still valid for prediction of peak heights, but not peak days.

Keywords: data analysis, mathematical epidemiology, acute respiratory infection, seasonal influenza, Python

1 Introduction and motivation

Acute respiratory infections (or, shortly, ARIs) are known as one of the oldest and the most popular human infectious diseases. The most notorious of them, influenza, has an ability to cause repetitive epidemic outbreaks with ARI incidence dramatically exceeding the average seasonal level. The outbreaks of influenza result in 3 to 5 million cases of severe illness annually worldwide, and the mortality is from 250 to 500 thousand individuals [14]. Influenza also causes increase of heart attacks and strokes [4], as well as other disease complications. It's important to mention that even during an epidemic outbreak only 15 to 20% of the total ARI cases are attributed to influenza [11], and the certain diagnosis whether a person is infected by influenza or by another acute respiratory infection with similar symptoms is possible only

Peer-review under responsibility of organizing committee of the scientific committee of the 5th International Young Scientist Conference on Computational Science

doi:10.1016/j.procs.2016.11.033

through laboratory testing [3]. Due to that issues the common clinical diagnosis 'influenza-like illness' (ILI) is often used, which includes all severe ARI cases fitting a certain description (the term often used in Russian healthcare is 'clinical influenza')[1]. The criteria of ILI vary slightly in different national healthcare systems. According to WHO, ILI is an acute respiratory infection with measured fever of ≥ 38 C and cough with onset within the last 10 days [15].

In the late 1960's, multiple mathematical models of influenza-like illness outbreaks were created. One of the most prominent studies of that time, accomplished by Baroyan and Rvachev, was connected with the flu propagation within the cities of Soviet Union [2] (later applied to worldwide propagation of the pandemic flu [12]). The Baroyan-Rvachev model was a combination of the Kermack-McKendrick SEIR model and a linear model of inter-city migration flows. Although the structure itself was not novel, it matched the true incidence data and achieved accurate forecasts of the outbreaks starts and peaks for one hundred Soviet cities included into the modeling system. For instance, the day of the outbreak start was predicted without mistakes in 56.1% of cases and with a bias less than a week in 92.2% of cases, the same numbers for the day of the outbreak peak were 53.0% and 87.4% correspondingly [7]. The accuracy of forecasts was achieved due to the efficient Soviet system of ARI cases registration [5] which gathered and collected reports on ARI incidence from the local healthcare units. The data for ILI outbreak model calibration was extracted from the obtained overall ARI and ILI incidence using statistical methods [7].

However, since early 1980's the Soviet modeling complex for flu forecasting, created in Research Institute of Influenza [6] and based on Baroyan-Rvachev model, showed the signs of growing incoherence with the epidemic outbreak patterns observed in Soviet cities [8]. The reason for that, according to one of the versions, is in the growing levels of collective immunity to flu due to increasing speed of its circulation around the globe. The dynamics of growth of collective immunity could be dependent from different factors, including the structure of contact networks within an urban area, that's why the original assumption that fraction of non-immune individuals is the same for all Soviet cities and depends only on currently circulating virus strain, seems to be less applicable [8]. According to [7], to maintain the original prediction accuracy, the values of parameters α and β (the fraction of non-immune individuals in the population and the force of infection correspondingly) was to be assessed separately for every city according to available *a priori* information on past outbreaks in this area. Dramatic events of the late 1980's – early 1990's, including the deaths of Baroyan and Rvachev and the collapse of the Soviet Union, adversely affected ARI surveillance and led to the cessation of use of the ARI forecasting system. The plans for testing and validating the model predictions came to a halt. For the past thirty years, the incidence data has not been fitted to the Baryoan-Rvachev model, so its predictive accuracy for modern data is unknown.

In this paper the authors are adapting to the modern data some of the ideas used in Baroyan-Rvachev forecast modeling complex for the sake of predicting the day and height of the flu outbreak peaks in Russian cities. Our aim was to assess the prediction accuracy of the named outbreak parameters and to find out to what extent the assumptions of Baroyan and Rvachev could be still applied for the sake of modeling the influenza propagation within Russia.

2 Research statement

The original idea of the Baroyan-Rvachev modeling procedure for a fixed flu epidemic season was based on the idea that influenza is spread between the Soviet cities via transport network,

[1] Although is is not completely accurate from the epidemiological point, for the sake of simplicity further in this study we consider 'flu', 'influenza' and 'ILI' synonyms.