COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 110 (2013) 455–470

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A new surveillance and spatio-temporal visualization tool SIMID: SIMulation of Infectious Diseases using random networks and GIS

Lilia L. Ramírez-Ramírez[a,*], *Yulia R. Gel*[b,c], *Mary Thompson*[b], *Eileen de Villa*[d], *Matt McPherson*[e]

[a] *Instituto Tecnológico Autónomo de México (ITAM), Mexico City, Mexico*
[b] *Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada*
[c] *Department of Applied Mathematics and Statistics, Johns Hopkins University, USA*
[d] *Peel Public Health, 7120 Hurontario Street, Mississauga, ON, Canada*
[e] *Infonaut Inc., 255 Consumers Road, Suite 500, Toronto, ON, Canada*

ARTICLE INFO

Article history:
Received 7 February 2012
Received in revised form
24 October 2012
Accepted 11 January 2013

Keywords:
Software for epidemic surveillance
Disease outbreaks
Spatio-temporal models in epidemiology
Networks

ABSTRACT

In this paper we discuss the SIMID tool for simulation of the spread of infectious disease, enabling spatio-temporal visualization of the dynamics of influenza outbreaks. SIMID is based on modern random network methodology and implemented within the R and GIS frameworks. The key advantage of SIMID is that it allows not only for the construction of a possible scenario for the spread of an infectious disease but also for the assessment of mitigation strategies, variation and uncertainty in disease parameters and randomness in the progression of an outbreak. We illustrate SIMID by application to an influenza epidemic simulation in a population constructed to resemble the Region of Peel, Ontario, Canada.

1. Introduction

The prompt detection of respiratory and gastrointestinal infectious diseases is of critical interest in public health practice due to the rapid transmission of such illnesses and their potential burden on the community. There is growing evidence that dynamic space–time data constitute a key element in monitoring and forecasting the spread of certain infectious diseases (see, for example, reviews by [25,38,6,42,19]). In particular, the threat of a severe influenza pandemic raises significant concern which necessitates the early spatial detection of an outbreak for appropriate public health action [30,9,20,27,44]. However, traditionally most public health departments do not incorporate space–time dynamic information into disease surveillance, reducing the efficiency and speed with which outbreaks can be detected and interventions started. Hence, there exists a substantial demand for easy-to-use, map-based software and decision-making support services that could assist healthcare professionals in assessing and integrating actionable space–time information on infectious diseases.

* *Corresponding author*. Tel.: +52 55 5628 4000.
E-mail address: lilialeticia.ramirez@itam.mx (L.L. Ramírez-Ramírez).
0169-2607/$ – see front matter
http://dx.doi.org/10.1016/j.cmpb.2013.01.007

The classical tool for visualizing spatial information on disease spread is a geographic information system (GIS). However, tracking the dynamics of infectious diseases and detecting changes in disease processes are impossible without the development and implementation of statistical methodology for spatio-temporal disease surveillance, which is not routinely included in GIS and most statistical software packages. There exists a vast literature on statistical approaches for space–time epidemiology [45,25,48,7,43], which can be generally classified into the three main groups, namely, statistical tests, model-based methods and hybrid routines. The first group, which includes, for example, such methods as scan statistics and cumulative sums, is dominantly employed for detecting outbreaks by comparing a spatio-temporal subset of data vs. the expected rate of disease incidence [13,23,24]. The second group focuses on assessing the relative risk of disease occurrence within a domain of interest, with respect to various spatio-temporal environmental and demographic covariates. Some of the most widely used approaches in this group are spatio-temporal Bayesian methods [47,26] and references therein and generalized linear models [22]. The hybrid group of surveillance methods includes modern simulation-based routines that aim simultaneously to model and to detect the dynamics of infectious diseases, and includes such approaches as hidden Markov models [49], Bayesian nets [18], agent-based models [11,10] and references therein and network methodology [35,36,4,2,21,17], and references therein. As mentioned by [42], while many of these new hybrid methods appear to be promising, they are not yet broadly implemented in software and, hence, are not widely available for operational disease monitoring.

In this paper we describe a new simulation-based visualization program SIMID for infectious disease surveillance, based on modern network methodology[1] and implemented within the R and GIS frameworks [8]. Our new network epidemic model adapts and extends the models for networks of social contacts proposed by [35,4], by allowing us to consider not only a single random network that is described by a specific degree distribution, but also a population network of contacts made up of subnetworks. Moreover, each such subnetwork can follow a different degree distribution. The network describes the contacts that can result in disease transmission, and is defined in terms of a probability distribution for the number of contacts that each susceptible individual has with other individuals in the community. Moreover, the network methodology we use for disease dynamics generalizes some important epidemic models, relaxing the hypotheses that the population is homogeneously mixed, that the infection rate is constant between any two individuals, and that the latent and infectious period have lengths that are exponentially distributed [39–41]. Hence, the network algorithm incorporates not only the use of compartments as in [32,31,12] but also some of the most important everyday contacts that can be modeled through random networks, such as contacts within families, educational facilities, health care institutions and transportation hubs. In addition, our new surveillance simulation-based tool allows calculation of probability distributions for the total number of infected individuals and the replacement number[2] under four different control measures: mass vaccination and acquaintance vaccination that are implemented prior to the outbreak, and ring vaccination and isolation that are applied during the outbreak evolution. Thus, the SIMID tool enables not only the modelling of a single potential scenario of an epidemic outbreak but rather the generation of an ensemble of potential scenarios for infectious outbreaks and, hence, the assessment of various associated uncertainties and mitigation strategies. The developed methodology and software tool are applicable to any population with a possibly heterogeneous contact structure described as a random graph. The spatial view of social interactions that represent channels of diseases transmission and that are modeled with the random network methodology is placed within the GIS framework. This, along with the uncertainty quantification techniques, enables surveillance teams and other health care professionals to identify disease outbreaks in a prompt time frame (if the SIMID tool is initialized with currently observed data on influenza-like illness (ILI) occurrences or confirmed influenza infected individuals), and allows public health personnel to use the program for planning preventive interventions and training (if the SIMID tool is run based on simulated data). Although our current case study relates to visualization over time and space of the dynamics of influenza, including pH1N1 and seasonal influenza, the simulation tool is also applicable to a wide range of other infectious diseases, such as measles, meningococcal meningitis, enteric illnesses and sexually transmitted diseases, to name only a few.

The paper is organized as follows. In Section 2, we describe the process of generating the network of contacts, in its application to an influenza epidemic simulation in the Region of Peel, Ontario, Canada, taking into account demographic information from the 2006 census. In Section 3, we discuss the SIMID tool along with its most important methodological characteristics. Section 4 is devoted to technical software specifications and architecture for SIMID. Section 5 provides information on the workflow for building possible scenarios for infectious outbreaks and presents a sample run of the SIMID tool. The paper is concluded by discussion and future work.

[1] Throughout the text, "network model" refers to a specific population's network of contacts, while "network algorithm" refers to selection of subpopulations (using available information, e.g. census data) and the computational algorithm that is used for generation of subnetworks. In turn, "network epidemic model" refers both to the network model and the epidemic model spreading along this network.

[2] Here by replacement number we mean the expected number of new cases generated by one single infected secondary case [5,35]. Note that the basic reproduction number and the replacement number are generally not the same since the basic reproduction number describes the expected number of new cases generated by patient zero, i.e. the first infected individual. It can easily be shown that when the network of contacts is modeled as a random network with Poisson degree distribution, then the reproductive and replacement number coincide.