

# Applications Of Operations Research/Statistics In Infection Outbreak Management


Andrey Fendyur, University of Calgary, Canada

## ABSTRACT

*Operations Research (OR) can be identified as the discipline that uses statistics, mathematics, computer-modelling and similar science methodology for decision making (Luss, Rosenwein, 1997). OR, powered with statistics and models, is a high potential engine for use in many areas that require evidence-based or model-based decision making. One of the most promising areas is specifically the infection outbreak management. Surprisingly, very little OR/statistics research has been aimed at infection outbreak management; usually, other general epidemiology issues were tackled in models. However, OR/statistics models for use in the infection outbreak management exist and can be effectively used in public policy and outbreak management practice. Probably, key reasons for that little involvement of OR/statistics in the infection outbreaks management is low awareness among the specialist community of OR/statistics use and benefits for their decision making. Up to the moment, there is lack of contemporary review of OR/statistics-applied models used for the infection outbreak management decision making. The present paper aimed at filling that gap and providing two benefits to involved health care managers and academics: first, developing awareness on the use and benefits of OR/statistics models for the infection outbreak management decision making, and second, for plotting the current state of affairs to highlight research opportunities for developing the field by academics and epidemic control professionals.*

**Keywords:** OR; Statistics; Infection Outbreak Management; Applied Models

## INTRODUCTION

 Operations Research (OR) can be identified as the discipline that uses statistics, mathematics, computer modelling and similar science methodology for decision making (Luss, Rosenwein, 1997). According to consensual views, OR was born during World War II to support military operations. Applications of those days were radar defence systems, anti-submarine systems, bombing strategies, etc. Upon the end of the war, OR was employed by leading industrial organizations to re-engineer and improve efficiency of business and manufacturing processes (Luss, Rosenwein, 1997). However, in the 1960s, a mathematical bias started to prevail in the discipline and that resulted in loss of authority in the society (Luss, Rosenwein, 1997). As a result of that focus migration to mathematics per se, many applied functions of OR were acquired by organizations. Nowadays, OR is in the process of recovering those identity crisis period losses (Luss, Rosenwein, 1997). In general, OR practitioners can be classified into three categories (Corbett, van Wassenhove 1993): 1) theoreticians (they develop new theoretical methodologies), 2) management consultants (apply existing tools to resolve practical problems), and 3) operations engineers (modify or extend use of existing techniques for industrial applications).

All three groups of OR research and practice are presented in health care management research. A newly emerged (but old demand-wise) discipline within health care management is disease outbreak management, or epidemics control, which deals with resource allocation at different stages in the epidemics process, from preparedness toward eradication (Brandeau, 2008). There is enough room for all three types of OR professionals in preventing, preparing for, and eradicating epidemics.

Reviews of OR/MS application in the “real world” provide an interesting collection of facts. For example, a short review of OR/MS research into disaster operations management is provided by Altay, Green (2006). However, there is a need for review of OR/MS research into epidemics preparations/management. From my best recollection, there is little done to provide a review of OR/MS applied research for use in epidemics control. That review will be welcomed by both OR/MS community and the epidemics control and management authority. Such a review will benefit both academia, public administration and the general public.

OR can address a number of problems facing epidemics management: 1) how to allocate resources among options to better control outbreak or spread of a disease within population(s), 2) what resources are needed to control a disease within population(s), 3) which resources should be employed for disease control interventions, etc. (Brandeau, 2008).

Resource allocation is a popular research topic in OR. However, epidemics management is complicated with non-linear nature of disease proliferation. Preventing just one individual from being infected provides an exponential growth of individuals saved in the future (Brandeau, 2008). The course of epidemics is extremely variable and individual, even among the same pathogen.

Conceptually, tackling the resource allocation for epidemics control and management involves several directions of research. Relatively simple and general analytical and computer epidemic models are used to plot and forecast epidemics. Quantitative analysis of advanced models is used for quantifying exposure and forecasting resources needed. Decision-making techniques are used to help decision and policy-makers in setting up policies and making decisions (Brandeau, 2008).

The basic model (Brandeau, 2008) used in epidemics modelling is

$$\begin{aligned} dx(t)/dt &= -\beta(t)x(t)y(t) - u(t), \\ dy(t)/dt &= \beta(t)x(t)y(t) - \gamma(t)y(t), \\ dz(t)/dt &= \gamma(t)y(t) - u(t). \end{aligned}$$

Here  $x(t)$  is the number of persons not infected at  $t$  time,  $y(t)$  is the number of infected persons at  $t$  time,  $z(t)$  is the number of persons taken away from infection outbreak,  $\beta(t)$  is the rate of getting infected at  $t$  time,  $u(t)$  is the immunization rate among susceptible persons at  $t$  time, and  $\gamma(t)$  is the removal from population rate at  $t$  time. The most important aspect in the model is the non-linear growth rate as a product of healthy and infected sub-populations.

On the basis of that model, many other ones were developed, breaking down elements on smaller variables to address, in detail, various aspects of disease outbreaks.

If the goal is to identify the optimal immunization rate or treatment stockpile, controls can be introduced into the model, such as vaccination, treatment or removal of patients, limitation of contacts, etc. The cost of disease is assumed to be linear (a product of cost per patient and the number of patients). The same is applicable to the cost of the control (Brandeau, 2008).

A brief illustration of that model’s use is provided in this section for the purposes of pre-informing the direction of the more detailed review in the next section. In this section, examples specify the method’s used, brief names, and outcomes. In the following section, examples contain a more detailed review of the modelling and methodological approach.

Lee & Pierskala (1988) developed a model to identify the best screening strategy for minimizing the number of patients. That strategy is a mass screening through certain time intervals.

May & Anderson (1984) considered allocation of a vaccine among various population subgroups to eradicate the disease. The model resulted in the optimal portion of each subgroup to be vaccinated.

Richter et al. (2000) developed a model for optimal disease control within various populations to minimize the number of disease cases.

A different type of framework is to use quantitative analysis in more complex models (Brandeau, 2008). For example, Longini et al. (1978) developed a model for identifying the optimal distribution of vaccine among various age groups during an outbreak of flu type A.

Another method was proposed by Tan and Yakowitz (1996); i.e., the use of a machine-learning algorithm for a Markov decision process in identifying the optimal epidemics control policies. The main idea was to continuously reallocate limited resources between programs aimed at decreasing the number of contacts or at lowering infecting capacity of patients. By the course of time, the model reaches equilibrium, which is the optimal point.

Brandeau (2008) noted that simulation models so far provide no exciting results due to two constraints: 1) sophisticated nature of epidemics not captured by models and 2) simplistic policies incorporated into models. However, she indicates that numerical techniques harvest practically applicable results if their limitation - hardness to collect data for model development - is overcome. For practitioners, Kaplan (1997) suggested the use of a production function construct – a sort of return on investment indicator, where the return is the incremental number of treatment units obtained for each project if funds are invested in it.

These methods and models help develop an understanding of the dynamics and outcomes of an epidemic. That makes epidemics control authorities capable of planning and exercising anti-epidemics actions and manage resources.

### **Review of currently employed models, methods and results**

The epidemics control service deals with a number of issues addressing different stages and elements of epidemics control. Preparation and planning constitute first service activities. Monitoring and identifying the epidemic complement those early stages. When an epidemic is detected, treatment and quarantine are in focus. That includes managing those affected, vaccinating still healthy population, and monitoring the outbreak development. Managing resources and addressing specific issues (such as the vector eradication, overlapping diseases, special population groups affected, etc.) are more examples of health care authority's challenges. Minimizing consequences and off-setting losses of epidemics are other tasks of the authority. In that sequence, the review of OR/MS research is organized. Highlights of research questions, methodology and results are given in a more detailed presentation than in the review section.

For the public, an epidemic starts when the actual germ spreads out and infects the population. For epidemic control authorities, the work starts much earlier – in the preparation stage. In its research, OR can address important planning issues confronting health care authorities.

Khans & Ashford (2001) argued for a wider spread of scientific knowledge on disease prevention, planning and management. They provided examples of the 1997 Nile virus outbreak in New York City and built up an argument that a better competence in disease control preparation and management would be an asset in many instances, from specialized public authorities to municipal entities. Thus, OR/MS techniques, targeting disease control preparation and management, have large potential for implementation in the health care industry.

Pierskalla & Brailer (1994) supported the argument of large potential use of OR/MS in the health care environment. One example they quoted was the use of AIDSPLAN, a spreadsheet model for planning resources needed in HIV/AIDS health care services in the United Kingdom. The model helped in demand forecasting by patient categories, care protocols, and resource and budget requirements. Both local and federal levels of resources could be summarized. Although simple, the application is helpful in routine duties.

The preparation/planning application of OR/MS was illustrated by Arinaminpathy & McLean (2009). They researched into preparing for viral epidemics. The scope of the paper was to test what anti-viral

treatment/prophylactic approaches had produced better results for both treatment and prevention. Limited stockpile of anti-viral agents was under assumption to make the research more applicable to real life. Also, specific risk groups' management strategy was researched. Specifically, the authors studied a scenario of response to a flu infection in the UK where 25% of the population could be treated with an anti-viral stockpile. What were the best treatment coverage strategies to reduce mortality and stop a disease outbreak? The compartmental model of homogeneous population was used first and then it was modified by inclusion of high risk groups in the population studied. The findings of the paper were leading to better effects of aggressive treatment/prophylactic, even if that exhausts the medicine stockpile early. In general, early and massive anti-viral treatment cut disease length in people and saved both finance and lives, and it also reduced the spread-out of epidemics. No matter if a stockpile is sufficient or not, results were the same. Peculiarities existed in addressing the needs of high risk groups, however. First, they benefited more if the drug was given to those who got better results, thus reducing transmission and the spread of disease. Second, enough stockpiles should have been held for those at particular risk. Overall, those findings supported aggressive anti-viral treatment should an epidemic emerge.

The health care authority's responsibility is to monitor infectious disease prevalence to recognize in time eruption of epidemics in order to commence appropriate actions. Delay of time for starting anti-epidemic actions can cost lives and substantial expenditures. An automated system can be used for recognizing the arrival of epidemics. The appropriate model should power that system. An approach to construct such a model was proposed by Rath et al. (2003). They developed a method for automatic, human-free, decision-making upon the start of an influenza outbreak. Their model is powered by the Hidden Markov Model technique, coupled with "an Exponential-Gaussian mixture, to characterize the non-epidemic and epidemic dynamics in a time series of influenza-like" cases. Validation of the method on actual data suggested reducing the incidence of wrong detections of epidemics compared to currently used approaches. Robustness of their method versus variability of epidemics scenarios was also higher.

Currently, the Serfling's method of cyclical regression is used to measure proportions of mortality rate due to pneumonia and influenza, adjusted for seasonal effects. Delays are due to lead time required to observe and confirm a death. Some other techniques were developed; for instance, a ratio of influenza-like illness (ILI) among calls on doctors. That method is also based on cyclical regression and has to deal with a need for non-epidemic data to the model baseline distribution. Also, data is considered independent and identically distributed, which is never true. That makes the development of an automated system problematic.

Hidden Markov Model is a sequence of pairs of random variables -  $(Y_t, S_t)$ ,  $t=1\dots n$  - that satisfy conditional independence assumptions. The use of the Hidden Markov Model enables utilization of historical time series for both epidemic and non-epidemic periods. Then this method assumes independence of observations contrary to marginal independence in the case of Serfling's method. The change between Gaussian and exponential distributions enables omission of complete modelling of trends and makes predictions more accurate. Also, if modelling non-epidemic rates with Gaussian distribution, a non-zero probability for negative rates is assumed. Thus, exponential distribution should be used for non-epidemic and Gaussian for epidemic periods in the model. According to Rath et al. (2003), the model proposed delivered better accuracy than previous ones and can function in automatic mode.

Another research and practical question OR can address is the way emerging epidemics will evolve. Will it die out or explode? The scenario is of crucial importance since resources involved in both of them are tremendously different in scale. The issue can be addressed through partially observed branching processes for stochastic epidemics (Panaretos, 2006). Branching processes are, for instance, death/birth or die-out/epidemic explosion. The author assumed that, quite often, observation of the epidemics had differed from reality (the measurement error). If so, the model developed could be wrong. Thus, deployment of controls and initialization of action plans could be missed or triggered for no actual reason. The model worked through analyzing the conditional probability of extinction of epidemics and the conditional time for that extinction from the point of knowing existing parameters of epidemics. Stochastic processes were used to model the construct. The author concluded that the use of randomness in partial observation for a given branching process could result in better stochastic models of epidemics. The technique used in the model was binomial thinning for introducing partial observations. The core of the model was making the infected and observed patients a Markov chain process. Transitional probabilities and conditional distributions were developed for and by the model too. The model described the probabilistic behaviour of the

epidemics and could be modified by the course of obtaining new data for initial stages. It could be used in policy making in cases of emerged epidemics – to launch anti-epidemic actions expecting outbreak of a disease - or wait until the epidemics die-out.

In simulating epidemic models, hierarchical networks could be of benefit (Quax et al., 2009). The essence of the method offered was parameterization of separate dynamics operators. The operators then were applied iteratively to the contact network (in this case, the epidemic model). That approach decreased the network's generator computational complexity (Quax et al., 2009), increased cache efficiency, and parallelized the simulator (making it quicker and more powerful). The authors tested the approach on a data set of more than a million homosexual men to model HIV epidemics and found the model to be quicker and simple to use. The authors identified that the features (or parameters) of an epidemic be network properties (along with community structure, etc.). Currently, standard mean-field approximations, or master equations, are used for modelling. The problem is to extend them for precise reflection of models. Hierarchical networks solve that problem. In their model called SEECN, Quax et al. (2009) organized nodes and edges hierarchically. They attributed arbitrary properties to nodes and edges, which rule the temporal evolution of the network (and epidemics). That evolution was managed by a complex of dynamics operators. Those could be parameterized independently and with regard to node and edge properties. The hierarchical organization reduced complexity of that dynamic system. Quax et al. (2009) showed that the running time for the model dropped by a factor of 2. Hence, hierarchical networks could offer flexibility and speed in simulating epidemic models compared to traditional mathematical models of epidemics.

Moreno et al. (2002) researched into out-breaking infections across complex population networks with acquired immunity. The issue of immunity is a powerful factor influencing the spread of infections. Acquired immunity could be due to previous exposure to the agent or immunization. Thus, the study added to the understanding and management of immunization and benefits/timing for vaccination. A new epidemics management framework was formed due to increasing incidences of outbreaks because of large connectivity fluctuations across those populations. The authors discovered that scale-free networks with diverging connectivity fluctuations and a large number of nodes limit produced a lower epidemic threshold and exhibited a finite portion of affected persons. That issue resulted in a high variance of response from systems upon introduction of infected individuals with various connectivities. The model added to understanding the spread of epidemics across complex heterogeneous networks. Examples of those systems could be populations with sexually transmitted diseases, such as AIDS, or respiratory diseases, such as flu or SARS. In those systems, the number of contacts per infected individual varied and created hardness in modelling epidemics. Scale-free (SF) networks were those where connectivity fluctuations were diverging, which were cases of heterogeneity. The spread-out of infections, hence, was much bigger in those systems. In practice, that meant that those systems were virtually never below the epidemic threshold and ever produced risk of an outbreak. That research shifted understanding of sexually transmitted and respiratory diseases toward treating them as in an ever-in-dangerous state. The way of eradication was to limit contacts for infected individuals which supported immunization, quarantine, etc.

Eames et al. (2009) addressed the issue of the contact network's use in epidemiological research. Many models used contact networks as a general approach. However, in actual epidemics, it is important to know the direction of interaction - who interacts and the effect of strengths for those interactions. The authors constructed a weighted interaction network, which reflected those factors mentioned above. Weights indicated the amount of time the individual spent in contact or proximity of contacts. Those were factors influencing the infective potential of a person and were used in assessing travel flow infectious properties. That allowed switching from idealized plain networks to more realistic weighted ones. Then the SIR stochastic model was constructed. The new model was used to assess effectiveness of intervention strategies. The principle was to assess the influence of each individual (element of the network) on the dynamics of epidemic which would allow making an informed decision on what path to take for prevention; for instance, in assessing effectiveness of vaccination for preventing an outbreak, etc. A novel and interesting development was incorporation of a diary study's results on social mixing behaviour for modelling patterns of individuals' interaction during epidemics. The authors came out with a practical recommendation on the circle of first-priority contacts for vaccination to suppress the outbreak. The authors justified incorporation of social behaviour data into models for vaccination effectiveness assessment. Among possible improvements coming from future research, the authors highlighted two. First, a way to incorporate the individual's

response to the disease would improve weights in the network. Second, collection of behavioural and epidemiological data during actual epidemics would improve predictive capacity of the model.

Yan et al. (2008) added to studies of scale-free weighted networks with asymmetric infection while focusing on both the epidemic threshold and critical behaviours. The study demonstrated that asymmetry could be redistributed across parameters influencing the spread-out of the infection, and the threshold could be recovered to make the infection more manageable. That finding was practically important. On the other hand, authors developed an argument that lack of threshold could mean a slow growth in prevalence. That argument sounded a bit questionable in epidemics management, since uncertainty could mean both low and high potential for an outbreak. The authors detailed on the role of airlines in the spread-out asymmetric rate, which was an interesting development in the research. That meant infecting probability from a node to its neighbours was not homogeneous and it was different to each direction from the node. The assumption was important since asymmetry of the infection spread complicated outbreak management in real life. Although both simulation and analysis were performed, the model was abstract and further research could yield more practical findings.

Rao and Kakehashi (2004) developed a non-linear dynamical method for modelling an epidemic. Various statistical distributions were tested. The model developed was robust as enough given inputs were accurate. Measurement errors affected capacity of the model significantly, which highlighted the importance of accurate and timely data collection on epidemics. The authors built up on the approach of convoluting statistical functions and likelihood estimation. The value of the method was high if information about early stages of epidemics development was not available. That is, for instance, the case of highly dangerous “imported” infections, such as Denge fever, West Nile fever, Ebola fever, etc. In such cases, the method could provide a viable estimate of risk (a product of uncertainty and consequences). A drawback was the method’s pure mathematical nature, not taking into account biology of the agent and spread-out (such as weather, demographics, etc.). This method combined a mathematical model with statistical estimation of variable, which was new to epidemiology and enabled to recover missing pieces of information while forecasting or planning anti-epidemics measures.

Coutinho et al. (2006) addressed the outbreak of diseases transmitted by mosquitoes. They developed a non-autonomous dynamic model which incorporated seasonal fluctuations in mosquito population. That model aimed at explaining Dengue fever’s seasonal pattern. The authors introduced a time-dependant threshold  $R(t)$ . The model assumed Dengue outbreak when the mean  $R(t)$  was more than 1 and no outbreak if below 1. The model explained the gap between the peak in mosquito population and the Dengue fever outbreak. The Dengue fever re-appeared worldwide as a major health care challenge in the 1990s, after decades of relatively local episodes of disease. The pattern of Dengue is at its peak in the wet season (3-4 months after the rains start) and drops to almost nothing in the dry season due to the mosquito population fluctuation. An interesting research question, which surprisingly the model provided insight on, was how the virus survived the dry season. According to the model, the virus survived in some mosquitoes alive and in their eggs – the pattern of mosquito population matched the mosquito number and hatching time. The egg reservoir seemed to be more plausible. The model in itself consisted of three components to incorporate all major locations of the virus – in people, mosquitoes, and their eggs. The model differed from the previously used time-independent threshold for vector-dependant epidemics that was developed by McDonald (1952). The model stipulated that the delay of epidemics, peak versus peak in mosquito population, was due to two factors – immunity of population and increase in number of infected mosquito bites (it takes time to get mosquito-infected from infected individuals). The extinction of disease outbreak was explained through the immunity parameter growth by the course of epidemics and a seasonal decrease in mosquito population. The model was an interesting application of OR techniques in tropical epidemiology.

Another application of OR in managing dangerous tropical epidemics of the Dengue and Chikungunya viruses came from Seyler et al. (2009). The researchers investigated the risk of import and outbreak of those diseases in Europe. The authors used a stochastic Monte-Carlo simulation model to assess the arrival of infected individuals and mosquitoes in Europe, particularly in those regions where the *Aedes albopictus* (the transmitter vector of those viruses) resided (Greece, Albania, Montenegro, Croatia, Italy, France, and Spain). For the simulation, researchers selected the following inputs: 1) incidence-based probability of Dengue and Chikungunya in those areas, 2) air traffic arrivals from endemic areas to Europe, 3) duration of viremia, 4) probability of being infected upon arrival, 5) distribution, and 6) period of the mosquito (vector) activity in those European countries. An

interesting development was building a model due to limited data for the Chikungunya virus. The authors came out with quantitatively defined results - a specified number of the median Dengue viremic person-days across countries (most of that in Italy). The most potentially dangerous traffic came from Asia and the most potentially vulnerable season was quarter 3. For the Chikungunya fever disease, exposure was much lower and all of them were estimated to arrive from India to Italy. The paper gave a good example of concrete outcomes that could be used by epidemic control authorities.

Hartemink et al. (2009) investigated the bluetongue virus basic reproduction number. The bluetongue virus is a dangerous infection agent. It had a history of outbreaks in Europe, such as Belgium-Netherlands-Luxemburg-Germany outbreak as of August, 2006. That particular strain (BTV-8) was traced to come from Africa, though the virus was present in other hot climate places like Asia and America. Usually the virus affected cattle, but it can bite humans. The vector for that agent could be a number of European resident mosquito species. The basic reproduction number ( $R_0$ ) was the expected number of secondary infection cases due to the single infected person who entered a healthy population. That model could be used in assessing the potential of and consequences after the arrival of an infected person from overseas to the host country. It draws an insight on the spread-out of the infection within the population. The study developed maps of  $R_0$  for better utility in health care and epidemics control practice. Such maps help identify areas of higher epidemic risk and pre-plan for epidemic outbreaks. The method incorporated effects of the environment on the agent proliferation into the model. The technique was based on integrating vector-abundance data along with statistical analysis to forecast abundance from satellite visualization. Biologically mechanistic modeling to underlie  $R_0$  was also employed in the model. The test of the model was performed in three dimensions with the bluetongue virus in the Netherlands. First, the authors developed a map of situation for 09/2006. Second, they developed species-specific basic reproduction number maps as a result of satellite-backed forecasts. Third, they plotted monthly maps for the year around. The novelty of their paper and approach was in combining the mathematical modeling with GIS (Geographic Information System). The value of GIS for epidemics control in the case of vector-related infections (transmitted by mosquitoes and other biological species) was in incorporating climate parameters, geographic details, human-influenced land use parameters, other anti-epidemics, and industrial factors into the method then available to epidemic control authorities. The methodology developed by Hartemink et al. (2009) seemed to be one of the most applicable and comprehensive for present times. The method merged various knowledge sources and control techniques, and its application generated cooperation of federal and local authorities, epidemics control and land use agencies, etc.

Townsend-Peterson et al. (2009) addressed the spread-out of bird-moderated viral diseases over the last 20 years (such as avian flu, H5N1, etc.). A part of that problem was migratory movement of birds. Those movements could cause quick and extensive outbreaks of epidemics along the migration route. The authors researched into seasonal (over-wintering and breeding periods) dispersion of birds across North America (392 species researched), which could play a role in disseminating the virus. On the grounds of information available, the authors built a model to show forecasted patterns and parameters of bird-transmitted diseases due to the birds' migratory movements. Due to fragmented and incomplete information from bird migration maps, the methodology employed was ecological niche modelling (ENM). That technique used environmental information to educate interpolations between registered locations where birds showed up (Townsend-Peterson, 2009). The inputs for the model were birds – species and numbers – and environment types. The evolutionary computing environment calculated, in ENM, a forecasted spread-out of bird species. That forecast then was analyzed and corrected on the basis of colony-forming probabilities to come out with a more realistic data set (Townsend-Peterson, 2009). The bird species selection was dominated by their epidemiological potential threat as vectors for pathogenic viruses. Environment parameters were selected as those contributing to appearance/lack of nesting birds in that particular neighbourhood. The whole data set was divided on three groups for validation purposes: 1) training data (25%), 2) intrinsic test data (25%), and 3) extrinsic test data (50% for model validation). As a result, the model achieved better explanatory capacity on bird-powered spread-out of epidemics in North America. New bird migration maps were introduced for use by epidemics control authorities. Another interesting, but to some extent, hardly believable, use of the model was suggested to be bio-terrorism investigation – theoretically, migrating birds can be targeted as a means of transporting viruses to the host country from over-wintering locations. The work in that direction could yield substantial applicable results when more information is available for modelling the migratory trajectories of birds.

Koelle et al. (2009) addressed the issue of pathogen mutation influence on the epidemics development, spread-out, and dynamics. That aspect is of contemporary focal attention. For instance, it deals with the evolution of the flu viruses, such as H5N1, H1N1, etc. Rapid genetic changes in viruses could substantially change models of disease emergence, progress, and dying-out which could result in significant variations and shortages in resource preparation, allocation, and logistics. As a result, quality of service could degrade and even health threats could emerge. Usually, models for rapidly-mutating pathogens faced two problems. Generated complexity of the model masked the models' numerical predictive capacity and opportunity to statistically estimate parameters. Then, detailing level limited generalizability of those models if a pathogen demonstrated variations in its properties. The model of Koelle et al. (2009) addressed those two problems and attempted to overcome them. The principal difference between the previous (multi-strain and SIRS) models and this one was simulating the tempo of the antigenic change instead of genetic changes in viruses. SIRS models had to deal with a complicated assumption of change in the host's immunity over the course of time. Multi-strain models had to incorporate parameters of many strains. The benefit of the Koelle et al. approach was the reduction in computational complexity. The second benefit was the avoidance of sticking to a particular virogenetic hypothesis. The new model fit them all.

Testing of the antigenic tempo model was performed in the influenza virus. Numerical results obtained matched findings from traditional models of the same viral epidemics in general. The difference was the substantially lower computational complexity of the new model. Hence, it was more reproducible in industry and public service by epidemics control practitioners.

For academia and epidemics researchers, the approach of antigenic modelling allowed for incorporation of spatial influences (such as geographic spread, etc.), extrinsic factors (such as humidity, temperature, etc.), immunologic hypotheses, and even some statistical inferences.

Handel et al. (2009) addressed the issue of infections overlap. SIR linear model type was used. An influenza epidemic often provoked the outbreak of complementary viral and bacterial infections, which inflated morbidity and mortality during the epidemics. The authors developed a model to estimate consequences and management strategies for those superposition outbreaks. Dimensions of the infections' interaction were spread-out of the epidemics' morbidity/mortality rates. The findings from the model showed that antibacterial interventions during viral epidemics could decrease the number of cases (incidences) and mortality. The value of the findings was in providing a quantitative framework to elaborate scenarios and support allocation of resources for those interventions (both prophylactic and therapeutic). Also, the authors supported the notion of developing preparedness for antibacterial interventions as a part of anti-epidemic plans. Collection of data was the major problem associated with the model development. Insufficient data available caused deviations in the model's outcome. The study made a first step in that area of research. Further models might wish to address heterogeneity of population issue and asymptomatic infections.

Boys and Giles (2007) advanced the use of SEIR (susceptible – exposed – infected – removed) stochastic compartmental models for forecasting epidemics outbreak dynamics with only partially known data (if only removal time is available). The model worked in the following way: those compartments (or stages) were common for population members in many epidemics. The statistical and modelling component was included through probabilities of transfer from one compartment to another. Those probabilities were usually derived from rates of contact, getting infected, etc. The authors stipulated that constant removal parameters were not applicable to many outbreaks. Thus, they developed a model with time-dependent function for the removal time/rate. The technique used was the reversible jump MCMC (Markov Chain Monte Carlo) algorithm. Gamma distribution was used. It allowed to incorporate Bayesian inference in the model through the use of model parameters associated with the step function. Validation of the model was performed on a smallpox outbreak and a respiratory infection epidemic. According to the authors, an important finding was the need to introduce time dependence by contrasting the predictive distributions of removal times. Then they might be compared to actual times. Should those estimated parameters be derived accurately enough, it would enable the model to predict epidemics' course and plan according to epidemics control and health management strategies. An interesting solution was to use the MCMC model and the reversible jump MCMC algorithm jointly to fit the model to actual data and make it more precise and applicable. Interestingly, that only one parameter available – the removal times of the infective individuals – enabled the model to develop posterior distributions for both the number and the position for removal rate changes. Boys and Giles



(2007) indicated that further research could yield an application of the same MCMC - the reversible jump MCMC joint use technique for studying and predicting infection and other rates.

Schwartz (1992) addressed an interesting issue - epidemics in children and predictability of small amplitude epidemics (which still could be dangerous and costly). The general method that was used employed SEIR modelling with seasonal forcing. That method was capable to forecast the outbreak of seasonal and unexpected epidemics based on children's disease parameters. The problem with that technique was a must of a large population size for that model. Should the population size be smaller, the solutions lost in quality. Schwartz (1992) experimented with relaxing assumptions of homogeneity in the arrival of susceptibility into the model. The finding was that the model predicts reasonably well "stable long period oscillatory epidemics having small amplitude". The problem could come from the least expected direction: the disease with which the population of children were immunized against could outbreak time-to-time and cause major health threat and costs. It was hardly possible to predict such events or forecast their spread and population affect. Schwartz (1992) used the coupled population model (i.e., the population size is identified by the sort of epidemic outbreak it produces). He showed that softening the susceptibility input rate into the small population homogeneity assumption could explain long period outbreaks not limited by population. The magnitude of those outbreaks, according to the model, was the same as "a small amplitude period 1 outbreak". In other words, adjustment in the rate of susceptibility introduction into the population could create a model explaining small amplitude outbreaks. That unveiled the course of hard-to-predict sub-type of epidemics. The model enabled the health care authority to better forecast and proactively manage the outbreaks.

Roberts (1982) developed a model for studying epidemics of parasite diseases with two hosts. An example of that could be larval cestodes in sheep. The author argued that the intermediate host, after contact with a parasite, could be infective or immune, or stay susceptible if no certainty. More complexity came from a possible outcome upon recovery - yes/no immunity - and the immunity duration could be short. Based on those assumptions, the authors used the SIRS model along with differential equation methods. That type of model could be useful for epidemics control practitioners. The model enabled the forecasting and understanding of dynamics and outcomes of parasite diseases. That alleviated anti-parasite plan development and resource justification and allocation. The significance of the paper extended over the limits of health care to agriculture.

The course of the outbreak and probability of extinction of an epidemic at the end of a large outbreak was researched by van Herwaarden (1997). The author developed an asymptotic model for identifying a probability value for extinction of an infection at the end of a large outbreak. The model used was a stochastic SIR model (where S stands for "susceptible", I stands for "infected", and R stands for "removed"). The researcher built local asymptotic expansions for plotting deterministic trajectories of the systems. The method applied was the Fokker-Plank equation for the stochastic system, used for asymptotical solution of a boundary value problem. The paper addressed the important issue of the population exposed ("susceptible") renewal rate. That scenario was common in case of active anti-epidemic measures. Thus, the model enabled to assess probability of success in eradicating the disease after an outbreak had happened and strong efforts made to combat it. If that rate was large, the outbreak would end up endemic. If the rate was not large, the probability of extinction was high. However, the study addressed the intermediate scenario too. That poly-scenario approach made the paper practically useful in epidemics control agencies and health care authorities' practice. It will help justify expenditures for active anti-epidemic efforts and determine health care economics of epidemics control.

Davies et al. (2003) addressed another important issue in modelling disease outbreak and dynamics. They focused on consequences of decision-making in planning for and combating epidemics - how to flexibly keep those expenditures minimal. For that, the authors argued, the policy-maker should be aware of alternative options in treatment that produce various outcomes. Based on scenario analysis, the decision-maker could choose the most efficient option. For so doing, the modeller might incorporate disease dynamics and various response options. That made a model more complicated theoretically, structurally, and computationally. Different modelling techniques would produce different results if applied to such dynamic and unstable problem conditions. The way to incorporate treatment plans and epidemics dynamics, according to Davies et al., was the time for transition from one state to another. However, that data could contain a substantial extent measurement error. The authors advised to check for guidelines and procedures to compare compatibility of inputs. On the basis of their own model developed for

various pathologies, the authors analyzed how a model type choice affected the results. For example, a *Helicobacter pylori* spread-out (causing stomach ulcer disease) was considered an epidemic, though not like the flu or fever. Meanwhile, the nature and properties of the disease fit epidemics traits. Early models were based on decision trees to address screening decisions. Then Markovian models were developed. The authors developed their own discrete event simulation model and came out with results different from previous models. In that model, individuals with attributes (entities) progress from one event to another in a timely manner (Davies et al., 2003). Those attributes could be diverse in nature to reflect features of individuals. Time transition intervals and passes were developed from sampling data. The attributes affected distribution parameters and brought flexibility into the model. Assumptions of the model dealt with the cohort and incidence patients (a single fixed incidence cohort was used), the base scenario (usually assuming no management measures were taken), system boundaries (what population and environment to include, etc.), and data compatibility. Results included lives saved, life years saved, quality adjusted life years, costs, NPV of savings, efficiency, etc. The main advantage of their technique was the ability to incorporate various patient-type parameters (co-morbidities, etc.) into the model without exponentially enlarging the number of states (contrary to that, previous models had some 5,400 states). Stochastic processes were behind simulation and need of operational researchers in better development and understanding. Then, many values needed were of statistical samples and origin; thus, statistics expertise was required. Also, results came from different points in time. Since systems changed, it was difficult to compare results from previous outbreaks. However, advantages of using discrete simulation models outweighed their disadvantages. The variance between old models' results and the new ones affected screening cost/need decision, length of disease, etc.

Abramson (2004) described an applied set of software (WINPEPI) developed to help epidemiologists in researching and planning epidemic control measures. That software was a good example of building a user-friendly application based on both statistics (mainly) and OR techniques (some of them included). WINPEPI consists of DESCRIBE (used in descriptive epidemiology), COMPARE2 (used for comparing two independent groups or samples in epidemiology, including an option to compare odds ratios for two samples), PAIRSetc (to compare paired and other matched observations), and WHATIS (a utility program for modelling epidemiologic challenges). Those programs together have 75 modules and can be used for alleviating the epidemiology practitioner's life in regard to statistical and modelling issues.

## **FUTURE RESEARCH OPPORTUNITIES**

The methods and models developed are not perfect and still need improvement. It is important to analyze drawbacks and gaps. Nishiura (2007) investigated the inappropriate predictive capacity of models used for forecasting the spread of AIDS (as an example) in the US and Japan. The author identified problems with underlying assumptions used for forecasting. Among those, the author highlights a normal distribution of the epidemic curve in the USA and the coverage ratio of previously diagnosed/undiagnosed cases in Japan. Also, lack of understanding the fundamentals of the disease played a big role in those failures. Reporting system changes enhanced biases too. The back-calculation method was impossible to be applied. Then the author built on those model properties and pieces of information that might be available and accurate. Among those, the author highlighted transmission parameters in vivo and as probabilities in models (sample: population generalizability issue, consistency with assumptions across models to make them comparable, interpretability of models for those involved in health care, and epidemics policy making and management). The research was of rare generalizing (methodological) style and aimed at assessing a spectrum of models from their outcome from a health care point of view. That approach differed from a usually narrow technique-specific stance of most publications. A more conceptual paper was useful for understanding key elements for developing the theory of a model. Nevertheless, the paper was not of use directly for epidemics managers; it contributed to improvements in the model building stage. Those improvements made the paper influential among epidemiology practitioners.

That paper, in particular, and future research sections from all papers mentioned, indicated several directions of future research:

- First, the multidisciplinary approach to the epidemics modelling/forecasting research: The OR/Statistics researchers will receive a complimentary benefit from teaming up with epidemiologists, managers,

physicians, etc. That combination of expertise will help better understand the nature of epidemics and, more exactly, reflect that nature in models.

- Second, the customization of OR/Statistics research for use in epidemiology: Diseases are various and numerous. They all have certain peculiarities that provide variance into the models. Thus, developing disease-customized models will help epidemics control practitioners withstand and better manage outbreaks.
- Third, the use of cutting-edge methods in the epidemics outbreak management research: OR/Statistics is strong in theoretical development of research quantitative methods. Active attempts to try new methods in epidemics control can produce substantial humanitarian, financial, and scholar results.
- Fourth, the resource management in the epidemics outbreak management research: Most of OR/Statistics research focuses on building models of outbreaks. Few studies address the unique nature of resources used in epidemics control and their use. Research in that direction will be warmly welcomed by the health care field.
- Fifth, the involvement of public sector in the epidemics outbreak management research: At present, cross-disciplinary involvement brings together professionals with expertise in specific scientific disciplines - biologists, virusologists, logistics managers, etc. Not enough OR/statistics research has studied and modeled issues facing disease control administrators. Examples are financial models of epidemics - direct expenditures and potential economy losses. Those models would help public administrators better understand epidemics complexity and better communicate the importance of proper proactive measures that save lives.

Thus, there are at least five large opportunity areas in OR/Statistics research toward epidemics control. Developments in that research are both rigorous scholarly and valuable socially. Developments in those research areas will boost OR/Statistics reputation in the society.

## CONCLUSION

Operations Research (OR), powered with statistics and models, is a high potential engine for use in many areas that require evidence-based or model-based decision-making. One of the most promising areas is specifically the infection outbreak management. Given the complexity of and high value at stake in that field, a number of stakeholders and interested participants, such as academia, public policy-makers and practitioners, might be interested in obtaining a review of literature on OR-backed methods and techniques to manage the epidemics outbreak. The paper aimed to contribute the OR in three ways: 1) provide contemporary literature review along with summary of methods/techniques/models used in epidemics outbreak management, 2) improve awareness among the specialist community of OR/statistics use and benefits in their decision-making for epidemics outbreak management, and 3) highlight research opportunities for developing the field by academics and epidemics outbreak management professionals.

Key findings indicated that there were a number of instruments OR could offer to the epidemics control field. The models/techniques could be deployed by practitioners and contribute to the health care. Review of the field would also help OR academics identify prospective research streams and projects to pursue.

## AUTHOR INFORMATION

**Andrey Fendyur** (B.A., MD, M.Sc.) is PhD student in Operations Management at the University of Calgary, Haskayne School of Business (Calgary, Canada). His industry experience includes 14 years in pharmaceutical/biotechnology industry in development and other functions. Research interests: innovation management, new product development and R&D operations management, health care operations, biotechnology and pharmaceutical industry, high-tech industries operations, renewable energy, multivariate analysis and quantitative methods. Member of POMS, INFORMS, ASA and biotechnology/pharmaceutical professional societies.

## REFERENCES

1. Khan & Ashford, 2001. Ready or not – preparedness for bioterrorism. *New England journal of medicine*. Vol. 345, 4, pp.287-289.
2. Pierskall & Brailer, 1994. Application of Operations Research in health care delivery. In: Pollock et al., Eds, *Handbook in OR & MS*, Vol. 6, 1994.
3. Abramson, 2004. WINPEPI (PEPI-for-Windows): computer programs for epidemiologists. *Epidemiologic Perspectives and Innovations*. 2004, 1:6, pp.1-10.
4. Altay & Green 2003. OR/MS research in disaster operations management. *European journal of operational research*. 175(2006) 475-493.
5. Arinaminpathy & McLean, 2009. Logistics of control for an influenza pandemic. *Epidemics*, 1(2009) 83-88.
6. Davies et al., 2003. The evolution of disease prevention and treatment using simulation models. *European Journal of operational research*. 150(2003) 53-66.
7. Nishiura, 2007. Lessons from previous predictions of HIV/AIDS in the United States and Japan: epidemiologic models and policy formulation. *Epidemiologic Perspectives and Innovations* 2007, 4:3, pp.1-7.
8. Roberts, 1982. Stability in a two host epidemic model. *Journal of mathematical biology* (1982) 14:71-75.
9. Handel et al., 2009. Intervention strategies for an influenza pandemic taking into account secondary bacterial infections. *Epidemics*, 1(2009) 185-195.
10. Koelle et al., 2009. Understanding the dynamics of rapidly evolving pathogens through modelling the tempo of antigenic change: Influenza as a case study. *Epidemics* 1(2009) 129-137.
11. Schwartz, 1992. Small amplitude, long period outbreaks in seasonally driven epidemics. *Journal of Mathematical Biology*, 1992(30):473-491.
12. Townsend-Peterson et al., 2009. A prototype forecasting system for bird-borne disease spread in North America based on migratory bird movements. *Epidemics*, 1(2009) 240-249.
13. Hartemink et al., 2009. Mapping the basic reproduction number (R0) for vector-borne diseases: A case study on bluetongue virus. *Epidemics*, 1 (2009) 153-161.
14. van Herwaarden, 1997. Stochastic epidemics: the probability of extinction of an infectious disease at the end of a major outbreak. *Journal of Mathematical Biology*. (1997) 35:793-813.
15. Seyler, 2009. Assessing the risk of importing Dengue and Chikungunya viruses to the European Union. *Epidemics* 1(2009)175-184.
16. Coutinho et al., 2006. Threshold conditions for a non-autonomous epidemic system describing the population dynamics of Dengue. *Bulletin of Mathematical Biology* (2006) 68:2263-2282.
17. Macdonald, 1952. The analysis of equilibrium in malaria. *Tropical Disease Bulletin*, 49, 813-828.
18. Yan et al., 2008. Epidemic threshold and phase transition in scale-free networks with asymmetric infection. *The European Physical Journal B*, 65, 591-594 (2008).
19. Luss & Rosenwein, 1997. Operations Research applications: Opportunities and accomplishments. *European Journal of Operational Research* 97 (1997) 220 – 244.
20. Brandeau, 2008. Resource allocation for epidemic control. In: *Encyclopedia of Optimization*, Second edition. Springer Verlag, 2008.
21. Corbett, van Wassenhove, 1993.
22. Lee & Pierskala, 1988. Mass screening models for contagious diseases with no latent period. *Operations Research* 36:917-928.
23. May & Anderson, 1984. Spatial heterogeneity and design of immunization programs. *Mathematical Bioscience*, 72:83-111.
24. Richter et al., 2000. Optimal resource allocation for epidemic control in multiple independent populations. Technical Report, Department of Industrial Engineering and Engineering Management, Stanford University, Stanford.
25. Tan & Yakowitz, 1996. Machine learning for Markov decision processes with application to an AIDS allocation problem. Technical report, Systems and industrial engineering department, University of Arizona.
26. Kaplan, 1997. Economic evaluation and HIV prevention community planning: a policy analyst's perspective. In: Holtgrave DR (ed) *Handbook HIV Prevention Policy Analysis*. Plenum, New York.

27. Rath et al., 2003. Automated detection of influenza epidemics with Hidden Markov Models. In: *IDA 2003*, LNCS 2810, pp.521-532, 2003. Springer Verlag, Heidelberg, Germany.
28. Quax et al., 2009. Simulating individual-based models of epidemics in hierarchical networks. In: *ICCS 2009*, Part 1, LNCS 5544, pp.725-734, 2009. Springer Verlag, Hamburg, Germany.
29. Panaretos, 2006. Partially observed branching processes for stochastic epidemics. *Journal of Mathematical Biology*, 2007, 54:645-668.
30. Moreno et al., 2002. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B*. 26(2002)521-529.
31. Rao & Kakehashi, 2004. A combination of differential equations and convolution in understanding the spread of an epidemic. *Sadhana*, vol.29, part 3, June 2004, pp.305-313.
32. Boys & Giles (2007). Bayesian inference for stochastic epidemic models with time-inhomogeneous removal rates. *Journal of mathematical biology*, 2007 (55)223-247.
33. Eames et al., 2009. Epidemic prediction and control in weighted networks. *Epidemics*, 1(2009) 70-76.

NOTES