Simple Models of Influenza Progression Within a Heterogeneous Population

Richard C. Larson

Center for Engineering Systems Fundamentals, Engineering Systems Division, and Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, rclarson@mit.edu

The focus of this “OR framing paper” is to introduce the operations research (OR) community to the need for new mathematical modeling of an influenza pandemic and its control. By reviewing relevant history and literature, one key concern that emerges relates to how a population’s heterogeneity may affect disease progression. Another is to explore within a modeling framework “social distancing” as a disease progression control method, where social distancing refers to steps aimed at reducing the frequency and intensity of daily human-to-human contacts. To depict social contact behavior of a heterogeneous population susceptible to infection, a nonhomogeneous probabilistic mixing model is developed. Partitioning the population of susceptibles into subgroups, based on frequency of daily human contacts and infection propensities, a stylistic difference equation model is then developed depicting the day-to-day evolution of the disease. This simple model is then used to develop a preliminary set of results. Two key findings are (1) early exponential growth of the disease may be dominated by susceptibles with high human contact frequencies and may not be indicative of the general population’s susceptibility to the disease, and (2) social distancing may be an effective nonmedical way to limit and perhaps even eradicate the disease. Much more decision-focused research needs to be done before any of these preliminary findings may be used in practice.

Subject classifications: health care: epidemiology; philosophy of modeling: community; probability: applications.

Area of review: OR Forum.

History: Received August 2006; revision received November 2006; accepted January 2007.

The threat of an impending pandemic of virulent influenza is being viewed seriously by virtually all nations of the world. While no one knows whether the current H5N1 influenza virus will mutate to become efficiently transmittable from human to human, most “experts” say that an influenza pandemic is inevitable. Like earthquakes, it is not a matter of “if,” but “when.” Because literally hundreds of millions of lives may be at stake, it is essential for the scientific community to explore to the limits of our abilities the consequences of alternative actions that may be taken in the presence of a pandemic situation to mitigate its deadly effects.

The purpose of this paper is to introduce the operations research (OR) community to the “flu pandemic” problem area and to encourage others to undertake decision-oriented research. In doing this, we review some brief history of the 1918–1919 “Spanish Flu” and Hong Kong’s experience with the 2003 SARS epidemic. We discuss the overall modeling approach of the more traditional epidemiology community, as they wrestle with flu modeling. We then create a family of simple models of influenza spread through a population. But primarily, we view this as a “problem framing” paper, suggesting ways in which the decision-focused, resource allocation-focused OR community may contribute to understanding and control of an infectious disease that potentially could kill hundreds of millions of people. The OR community has previously contributed in substantial ways to our collective knowledge of massive health threats to the population, as in emergency response to a smallpox attack (Kaplan et al. 2002) and as in response to an anthrax attack (Wein et al. 2003). In November–December 2006, the INFORMS journal Interfaces devoted an entire issue to applications in homeland security, including preparedness and response to health-related events (Lowe 2006).

Many researchers have written about mathematical models of influenza spread. But the majority of the existing models make rather sweeping assumptions about the population they are modeling. These often include one or more of the following assumptions:

(1) All members of the susceptible population are identical from a modeling point of view;

(2) There is a fundamental input constant for the model, the “basic reproductive ratio $R_0$,” that characterizes the mean number of new influenza infections created by each newly infected person;

(3) Local groups involving at least one infected person are characterized by homogenous mixing, meaning that each susceptible member of the group is equally likely to become infected from any infected person; and

(4) Social behavior does not change during the course of the pandemic.

Our approach is somewhat different from the usual in that we focus on a heterogeneous population. In our model, the susceptible population is heterogeneous in a variety of ways. We allow for socially active people who interact with many other people on a given day, and we allow
for relatively inactive people who interact with few others. We provide for highly susceptible people who are more likely to become infected once exposed to the virus, and we include those who are less susceptible. In a similar vein, we allow for highly contagious as well as less contagious infected persons. The reasoning behind these assumptions is that heterogeneity across the population in these attributes may affect in first-order ways the manner in which the disease propagates and, consequently, the manner in which we should address mitigation measures.

As a final new assumption, we allow social behavior to influence how the illness propagates over time throughout the population. We do not agree with implicit modeling assumptions behind the so-called basic reproductive ratio $R_0$, the currently popular input parameter to many influenza-modeling efforts. We do believe that each “exposure” will transform a susceptible healthy person into an infected person with some given probability, but the net growth of illness over time depends critically on the frequency of such exposures. That frequency depends on social behavior. Socially active persons have many more contacts per day than others, and thus—all else being equal—are more likely to become infected and then to spread the infection to others. We allow social behavior to change as the pandemic evolves. In particular, we model in a very aggregate way the “social distancing” and “hygienic steps” that are observed in pandemic situations, these occurring both by individual choice and by government fiat. Social distancing, including such measures as telecommuting, closing of schools, and even mandated minimum physical distances between co-workers, has the effect of reducing the frequency and intensity of human-to-human contacts that could spread the disease. Hygienic steps, including covering one’s mouth when coughing or sneezing, wearing masks, and frequently washing hands tend to reduce the likelihood of becoming infected because of contact with an infected person.

Our modeling framework utilizes difference equations to model the day-to-day progression of the pandemic through the population. We do not explicitly incorporate geography or other spatial attributes. Rather, ours is a stylized model whose primary purpose is to discover the contributions made by a population’s heterogeneity and social behavior in the evolution of the disease. The hope is that understanding the effects of heterogeneity and social distancing with hygienic steps—to first order—will help both governments and individuals develop better control measures for fighting an influenza pandemic, when it next occurs. We also hope, as stated above, to encourage others from the OR community to work in this vitally important area.

1. Social Distancing and the Limitations of Binary Decisions

Our notion is that an influenza pandemic is inevitable and that usual isolation controls and vaccines will not stop the progression of the disease. Judging by past attempts to control influenza pandemics, international controls at airports and seaports will not be effective. That is because a typical person infected with the influenza virus on “day 0” usually has a latent period of two to three days with no symptoms and with no ability to infect others, followed by a period of roughly 24 hours when he is infectious but asymptomatic, followed by an infectious symptomatic period. After that, the person is either dead or recovered and immune to recurrence of that viral infection. It is the 24-hour infectious but asymptomatic period that makes border controls of limited value. Testing for a person’s body temperature or other symptoms will show an apparently healthy person, yet that person is infected and may be infecting others as he is now “shedding virus” simply by breathing. Regarding vaccines, past experience has shown that six months is the minimum time needed between the occurrence of virus mutation so that it becomes human-to-human efficiently transmittable to production of a tested and effective vaccine. Within six months of the virus mutation, the virus may have engulfed the Earth in a pandemic. Even after six months, production capacities would limit the vaccine to a small fraction of the planet’s population. So, international controls will not work and vaccines will not be available when needed, if past performance is any predictor of future behavior.

If “the flu is among us,” then governments and individuals have behavioral options to consider. The quarantining of a population and the isolation of infected patients represent extreme governmental actions, types of “binary decisions” where the two decisions are “do nothing” or “quarantine or isolate.” Ideally, one wants to look at the full spectrum of life style, work style, and governmental policies between “do nothing” and “quarantine or isolate.” Our working hypothesis is that a wide range of voluntary and also government-suggested policies less than “quarantine or isolate” could have a major positive effect on limiting the progression of the flu through the population. Many of these policies have been labeled “social distancing.” Others relate to hygienic behaviors.

Social distancing has roots over centuries, often as a type of group evolutionary survival mechanism. In rural India in the nineteenth and early twentieth centuries, subsistence farm families who lived closely together in villages but who worked separate land plots outside of the villages, left the villages and lived separately on their land whenever they heard from a trusted messenger that “a plague” was “in the vicinity.” They returned to their village homes once the signal was given that the risk of plague had subsided, the duration of the distancing typically being about two weeks.2 While this policy seemed to work well for rural subsistence farmers, we may well ask, “What is the analogue to the movement to the land in our highly-networked interconnected Western style of life?” We are not self-sufficient and we rely on others to provide virtually all essential services and products for living. Given all the
interconnected networks upon which we rely, is social distancing itself, in the simple ways in which we can do it, sufficient to control the evolution and penetration of a flu pandemic?

Some severe forms of social distancing are being integrated into state-created pandemic flu response plans. Many state plans mention closing of schools, restricting mass gatherings, limiting travel, and quarantine as types of anticipated “social distancing” steps. Closing events and facilities is another type of binary decision, namely, either “open for business as usual” or “closed and shuttered.” Many state plans are being created hurriedly and without helpful research supporting the selected policies.

In 1918, during the “Spanish flu pandemic,” various forms of nonbinary social distancing were tried. In addition to “the binary” closing of schools, churches, social events, “saloons,” and vaudeville acts, the cities of Boston, New York, and Washington, DC negotiated staggered work hours with key employers. The staggered work hours were thought to reduce crowding on public transportation, and thus reduce the risks of becoming infected with the flu. In London, England, times between shows in theatres were stretched to two hours, to allow clean air into the building. St. Louis and Atlanta were especially vigilant in implementing their social distancing, early and often. But St. Louis had its limits. City officials banned dancing in hotels and coffee houses but not in dance halls! In effect, each city in 1918 in its reactions to the pandemic served as a type of local “quasi-experimental” response to the flu.

A key issue with social distancing measures is the timing of their implementation. Once the flu is among us, closing schools and other public facilities may be too little, too late. Then, there is the decision as to when to undo social-distancing measures. In 1918, several cities that unwound social-distancing policies too early experienced another wave of the flu, apparently caused by increased social mixing.

The critical need for performing research in this area is illustrated, for instance, by a letter written by Congressman Burt Gordon, the ranking member of the Science Committee of the U.S. House of Representatives, to Michael O. Leavitt, Secretary of Health and Human Services (HHS) (Gordon 2005). There he criticizes the national flu response plan to be too much medical science and too little social science. Among other things, he says:

I was most concerned by the incompleteness of the plan in the areas of risk communications, social distancing, and analysis of collective behavior in times of crisis. …Social distancing, effective communication, and other public health measures will be our only realistic line of defense.

The HHS plan recognizes the importance of addressing the social dimensions of a pandemic but does nothing more than list them as considerations and needs. There is no indication of how HHS intends to address these issues; there is no indication that they expect anything more than for the states and localities to magically know how to address the items listed. Agreed, many of the social issues are the providence of the state and local government. However, the states and local governments look to the Federal government for guidance and best practices. Such guidance is lacking in the current Pandemic Plan.

I am concerned that the social science research that is generally being utilized for disaster planning is 50 years old—apparently the Department of Homeland Security still relies on a civil defense model of risk communication and response. The Civil Defense model (e.g. “duck and cover”—hide under your desk should a nuclear bomb be dropped) has long been discredited by social scientists but, as we understand it, that is the science on which the failed “duct tape” for anthrax/bioterrorism campaign was based. . . .

Unfortunately, the fields of biology and social science do not often overlap and the pandemic plan clearly reflects authorship by the former and not the latter. . . .

Research such as this should be the underpinnings of determining how social distancing and isolation techniques would be tolerated and adhered to by the public. Widespread understanding and engagement by the public through a transparent process will be critical to ensuring tolerance and cooperation. (Emphasis added.)

2. The Limitations of R0

A critical parameter used in the aggregate modeling of pandemics is $R_0$, the basic reproductive ratio, defined to be the average number of new infections generated by an infected person in a population of 100% susceptibles. Early in the pandemic, it is the average transmission rate from one person to others. Or, in early generations, it is the average growth factor in numbers newly infected from “generation to generation” of infected individuals. An $R_0$ of 2, for instance, would indicate that an average infected person transmits the disease to two others. Any $R_0$ less than 1.0 virtually guarantees that the flu will die out of the population. Any $R_0$ greater than 1.0 will cause a near-term exponential increase in the number of people infected with the flu. Mills et al. have estimated $R_0$ of each of 45 U.S. cities in the 1918–1919 Spanish Flu pandemic, finding typical values of $R_0$ approximately equal to 2 (Mills et al. 2004). Perhaps coincidently, a recent highly publicized mathematical simulation used a fixed $R_0$ of 2.0. We quote from the news release (Knox 2005):

A federally financed study used supercomputers to predict what might happen if a virulent and easily spread new strain of flu entered the United States. . . .(The researchers) assumed the pandemic virus would leak into the country despite efforts to screen travelers for flu symptoms. If each infected person spread the virus to two others, large outbreaks of flu would occur all over the country within about two months after the virus began to spread. The national epidemic would peak around day 85, with about 4.5 million people falling ill that day. In the end, 122 million Americans may have gotten sick, more than four times the toll in a usual flu season.
The concept of using $R_0$ seems to be accepted in an unchallenged way in the epidemiology community. From Heesterbeek (2002, p. 189), an expert of the history of $R_0$, we hear,

The basic reproductive ratio (or number) $R_0$ is arguably the most important quantity on the study of epidemics and notably in comparing population dynamical effects of control strategies. The quantity is defined as the expected number of new cases of an infection caused by a typical infected individual in a population consisting of susceptibles only. In the last 10–15 years $R_0$ is an ingredient in almost all papers that use some mathematical modeling in studying the spread of infectious agents.

As described by Heesterbeek (2002), $R_0$ was created in Germany by demographers in the 1880s and formalized in 1925 to model the progression of a country’s population. The original $R_0$ was defined to be the average number of female offspring born to one female over her entire life. For the year 1879, this number for Germany was estimated by Richard Bochk to be 1.06 (Heesterbeek, p. 191). The time scale was decades and the system was in approximate equilibrium. With an influenza epidemic, the time scale is in days and weeks and nothing approximating equilibrium exists. To the contrary, the system is characterized by markedly changing parameter values as society copes daily with the influenza’s progression. Over the last three decades, epidemiologists have adopted the $R_0$ concept and applied it to a variety of diseases, some of which (e.g., malaria) existed in a type of quasi-equilibrium similar to that of population demographics. But the original demographic motivation and near steady-state environment supporting $R_0$ simply do not exist in a dynamic influenza epidemic situation.

From a modeler’s perspective, $R_0$ has additional limitations. It does not link policy to outcome. As is widely recognized, $R_0$ varies considerably across the population. It is a random variable, not a constant, and its accurate modeling should reflect that. But more fundamentally, an individual’s $R_0$ will depend critically on life-style and work-style choices. For instance, consider two individuals, $A$ and $B$, who have identical lives except at work, 9:00 A.M. to 5:00 P.M. each day. One could consider these individuals as virtual clones whose lives in a statistical social networks sense are identical outside of work hours. But person $A$ sells subway tokens to 500 commuters per day, and person $B$ is a lone security guard at a lighthouse at the end of a peninsula. Suppose that on a given day, each had been infected with the flu virus two or three days earlier, and each is currently asymptomatic but unknowingly shedding virus at identical shedding rates. Person $A$ has the potential to infect up to 500 people during work hours, and Person $B$ can infect no others during his work hours. This is an example of how $R_0$ is person specific and situation specific. Illustrating this idea in a social distancing situation, consider a suburban commuter who usually travels from home to work by public transportation but instead—in the presence of the flu—stays at home and telecommutes. His/her transmission rate (or $R_0$) should decrease dramatically. A toddler usually in day-care, who instead is kept at home with a parent or other care provider, should have a lower $R_0$. The key point is that $R_0$ depends on many factors, a large number of which are under our control. A better model is to consider the rate of virus shedding, the duration of that shedding, and then the numbers of people in proximity, with some informative descriptor of that proximity. In decision-oriented modeling, the equivalent of $R_0$, appropriately defined, should be viewed as an outcome measure of the model, not a fixed input parameter.4

Some epidemics are said to be started by index patients, so-called “patient zeros.” Among these are Mary Mallan (AKA Typhoid Mary), who is said to have infected 47 people; Gaetan Dugas, a Canadian flight steward, who allegedly infected many men with the HIV/AIDS virus; and Liu Jianlun, who allegedly started the SARS epidemic in Hong Kong by infecting people staying on the same floor of the Hotel Metropole of Kowloon, and his brother-in-law (who subsequently infected up to 79 others). Whether or not these are true patient zeros, they certainly point to extreme heterogeneity in the population of those infected, demonstrating that one individual—a “super-spreader”—may infect scores of others, far from any “average” disease reproductive ratio. Especially at the beginning of an epidemic, such behaviors by super-spreaders can radically affect any computation of $R_0$.

Even in demography, where quasi steady-state operation supports use of the $R_0$ concept, human behavior demonstrates that $R_0$ is far from an immutable constant. In Germany today, more than a century after the first estimate of Germany’s $R_0$, the current $R_0$ is estimated to be about 0.70.5

3. Hong Kong SARS 2003

The potential for social distancing and supportive actions in “personal hygiene” is shown by statistics on respiratory diseases in Hong Kong during the 2003 SARS epidemic (March 11–June 2). SARS is a severe respiratory disease spread by infectious respiratory droplets and through exposure to objects or substances contaminated with the virus. To reduce their risk of SARS infection, Hong Kong residents during this 2.5-month period radically altered their social behavior, both in terms of hygienic behavior and social distancing. For example, surveys indicate that 76% of the population wore masks, 65% washed their hands after contact with potentially contaminated objects, 78% covered their mouths when sneezing or coughing and used diluted bleach for household cleaning (>50%) (Lo et al. 2005). In addition, various social-distancing steps were taken, including closings of schools, libraries, swimming pools, and a major sports center. There was a sharp reduction in many discretionary activities such as attending social events, shopping, and going to restaurants.
As is well known, the SARS epidemic was stopped in Hong Kong and in other locations where it occurred. But one may reasonably ask, “What were the beneficial effects of the population’s hygienic steps and social distancing?” To answer that question independently of the success in eliminating SARS, researchers have examined the incidence of other acute respiratory viral diseases during the key months April and May 2003, and compared disease incidence to seasonal norms. These other respiratory diseases are seasonal influenza, parainfluenza, respiratory syncytial virus (RSV), and adenovirus. Compared to a five-year previous base period, the incidence of these other respiratory viral infections in medical laboratory specimens decreased an average of about 90% during April and May 2003 (Lo et al. 2005). Because the analysis was retrospective with no statistical controls, we must be cautious in extrapolating these numbers to other environments. We do not know the relative effectiveness of the social-distancing steps versus the hygienic steps. Influenza is not SARS, and various steps taken to eliminate SARS are not available for influenza with its much shorter incubation time and its infectiousness under asymptomatic conditions. But, the behavioral modifications induced by SARS reduced incidence of seasonal flu—as measured in laboratory specimens—by about 90%. This represents compelling evidence that behavioral modifications can dramatically reduce the spread of respiratory infections. It should be clear that any modeling analysis that ignores behavioral changes, many of which occur over very few days, removes at the beginning of the investigation perhaps our greatest single set of disease-progression control strategies.

The SARS epidemic caused great concern in the epidemiology modeling community. Two remarkably timely modeling reports were published in late May 2003, while the epidemic was still on going (Lipsitch et al. 2003, Riley et al. 2003). They may well represent the state-of-the-art in what we might call $R_0$-oriented mathematical modeling of epidemics. Reading these papers, done by recognized experts in their fields, one sees the “hoops that need to be jumped through” to cope with disease progression realities that are not represented in the concept of a fixed $R_0$. These realities include marked heterogeneity in the population, super-spreaders of the disease, lack of steady-state or equilibrium conditions, and social controls (what we call social distancing). From Riley et al. (2003), we hear: “These results indicate (that the equivalent $R_0$) is now below the self-sustaining threshold of 1 in Hong Kong and thus that epidemic is currently under control— in the sense that infection rates are declining. They also suggest that reductions in population contact rates (both in the community and in hospitals) played the predominant role in achieving control” (emphasis added). From Lipsitch et al. (2003), we hear: “Future work should certainly focus on quantifying transmission and other epidemiological parameters in a variety of circumstances…to construct more detailed models of transmission that realistically incorporate the effects of heterogeneities in specific settings” (emphasis added). The two analyses that were published in May 2003 contributed greatly to our understanding of SARS and its progression, as the authors went far beyond the usual $R_0$ concept to explain what was happening in the population. But they themselves argue for models that explicitly include population contact rates and control thereof and for population heterogeneity.

### 4. Other Related Literature

There is a growing literature on flu pandemic modeling, and covering it all would require a separate paper. Here we cite illustrative recent works. At present, there are two popular ways to model the progression of an infectious disease through a population. One is via modeling with differential equations (Cahill et al. 2005) and the other is via detailed simulation modeling building on transportation planning models and social networks (Eubank 2005, Eubank et al. 2004, Ferguson et al. 2005, Germann et al. 2006, Christley et al. 2005). The differential equation models assume that members of the population are in one of a finite number of states, typically susceptible (S), infected (I), and recovered (R); the models are called S-I-R models. Isham (2005) has written an excellent overview of recent S-I-R modeling work that has attempted to introduce heterogeneity and spatial spread.

Increasingly, the social network modeling seeks to address some of the issues addressed herein related to social behavior and risk of infection. But the methods used and assumptions made are distinctly different from ours. Longini et al. (2005) use a massive simulation model with social networks to address the possible spread of influenza in Southeast Asia. Their goals are similar to ours. Meyers et al. (2003) use social network theory to explore how human contacts of caregivers with patients spread the incidence of pneumonia. Pourbohroul et al. (2005) use “contact network epidemiology” to study control strategies for SARS. Watts et al. (2005) link local homogeneous mixing to a more complex nested spatial transmission model to depict via a “metapopulation model” complex evolution of the disease over space and time. Wu et al. (2006), building from the Hong Kong SARS experience, use simulation modeling to demonstrate the effectiveness of even partial voluntary “household-based” social controls such as self-isolation and quarantine, predicting a significant reduction in disease spread as a result. Finally, there is emerging a small group of researchers who are critiquing commonly used modeling approaches. An example is Wearing et al. (2005), who use basic modeling principles to critique the typical use of common parameters found in epidemiological models.

Many who model epidemics know the limitations of an assumption of population homogeneity. Examples of good work using heterogeneity include Wallinga et al. (2006), who examine age-specific heterogeneity with regard to spreading respiratory infections, and Lipsitch (2005), who
demonstrates the effects of behavioral heterogeneity of a population spreading sexually-transmitted diseases.

Regarding social distancing as an important control measure to delay and reduce incidence of influenza infection, the World Health Organization is now promoting social distancing as a first-order policy (World Health Organization Writing Group 2006). Recently, researchers at the Harvard School of Public Health have surveyed Americans who have expressed surprising willingness to engage in social-distancing measures if and when pandemic influenza is among us (Blendon et al. 2006). Finally, using a social-networks construct, Glass et al. (2006) develop a stylized model of a small town to demonstrate the positive efforts of social distancing to delay and avoid pandemic influenza.

The need for modeling for framing decisions has recently been emphasized by the Institute of Medicine’s Committee on Modeling Community Containment for Pandemic Influenza (2006).

5. Model 1: Active and Inactive Persons Only

In our first model, we allow for two types of persons in the susceptible population: highly socially active persons and persons whose social activity is low. We assume that “social activity” refers to frequency of human contacts, not necessarily in purely social situations but more often in professional and day-to-day living situations. A retail store clerk has many human contacts per day and is thus socially active. A novelist working from an office at home has fewer social contacts on a typical day and is thus characterized by a low level of social activity. We assume that social contacts occur as a homogenous Poisson process, with rate parameters dependent on the level of social activity. In particular, define

\[
\lambda_H = \text{Poisson rate of social contacts per day of a high-activity person.}
\]

\[
\lambda_L = \text{Poisson rate of social contacts per day of a low-activity person.}
\]

\[
n_h = \text{initial population of high-activity persons.}
\]

\[
n_l = \text{initial population of low-activity persons.}
\]

\[
n_h + n_l = \text{total population.}
\]

We like to think of human interactions, generated by Poisson processes, in terms of some physical model that we can visualize. Suppose that each time a person interacts with another, she leaves a slip on the ground, labeled H or L, depending on whether she is high activity or low activity, respectively. Each interaction provides two slips. At end of day we can sample interactions by randomly “picking up slips.” There are on average \(n_h\lambda_H\) interactions of high-activity people during the day and \(n_l\lambda_L\) interactions of low-activity people during the day. The mean total number of interactions during a day is \((n_h\lambda_H + n_l\lambda_L)/2\). This result clearly generalizes to any number of activity-level categories.

To obtain some comfort and familiarity with this heterogeneous mixing model, consider a random person, R, high activity or low activity. The next interaction of R with another will be with a high-activity person with probability \(n_h\lambda_H/(n_h\lambda_H + n_l\lambda_L)\). The next interaction of R will be with a low-activity person with probability \(n_l\lambda_L/(n_h\lambda_H + n_l\lambda_L)\).

Consider a randomly selected interaction pair, with persons R1 and R2. The next R1-R2 interaction is likely to be high activity with another high activity with probability \([n_h\lambda_H/(n_h\lambda_H + n_l\lambda_L)]^2\). Think of this as picking up two paired slips of paper from the ground. The next interaction is likely to be an inactive person with another inactive person with probability \([n_l\lambda_L/(n_h\lambda_H + n_l\lambda_L)]^2\). The next interaction is likely to involve both a high-activity person and a low-activity person with probability \(2[n_h\lambda_H/(n_h\lambda_H + n_l\lambda_L)][n_l\lambda_L/(n_h\lambda_H + n_l\lambda_L)]\).

These results also generalize to more than two activity-level categories.

5.1. Day 1

We now model the time progression of the influenza infection. We are assuming that members of the population can become infected only through contact with others who are infected and infectious. Our stylized model is ignoring the process of becoming infected by touching a contaminated object or by breathing air that has been contaminated by someone who has left the area and so would not be counted as a human-to-human contact.

We assume that a susceptible person infected on day \(i\) is infectious but asymptomatic only on day \(i + 1\), the last day he or she interacts with the population in the usual way, \(i = 1, 2, 3, \ldots \). We assume that such a person is symptomatic on day \(i + 2\) and is removed from the population at the start of that day. Thus, an infectious person, always viewed as asymptomatic, can only infect others on the day after he becomes infected. Once removed from the general population, an infected person will eventually either die or recover and return to the population—no longer susceptible due to acquired immunity. We start the model on day \(i = 1\) and iterate it through as many days as desired. The simplifying assumptions related to time progression allow us to examine “generations of the disease” as it progresses through the population. Here, day \(i\) of the pandemic is equivalent to “generation \(i\)” of the disease. The model can easily be generalized to allow for a two to three day latency period prior to development of asymptomatic infectiousness. It could also be generalized to allow for a probabilistic distribution of time that an infected person is infectious and asymptomatic.

Define

\[
n_{h1}(1) = \text{day 1 population of high-activity infectious asymptomatic persons.}
\]

\[
n_{l1}(1) = \text{day 1 population of low-activity infectious asymptomatic persons.}
\]

\[
n_{h1}(1) = n_h - n_{h1}(1) = \text{day 1 population of high-activity susceptible persons.}
\]

\[
n_{l1}(1) = n_l - n_{l1}(1) = \text{day 1 population of low-activity susceptible persons.}
\]
\( p = \) probability that a susceptible person becomes infected, given contact with an infectious individual.

The constant parameter \( p \) may be deemed too simplistic to depict the likelihood of infection, given “contact.” This is because there are many different kinds of contact, ranging from paying a toll to a tollbooth employee, to sitting next to someone in an airplane, to a couple sleeping together, to a caregiver attending an infirmed loved one. Our interpretation of \( p \) is that it depicts the weighted average likelihood of infection, given contact with an infectious person, over the range of possible types of social contacts, the weights corresponding to the respective relative frequencies of the various kinds of contact.

Consider a random person, \( X \). The probability that the next interaction of \( X \) is with an infected person is equal to

\[
\beta = [\lambda_H n_H(1) + \lambda_L n_L(1)]/(\lambda_H n_H + \lambda_L n_L).
\]

We now derive

\[ p_H^i(1) = \text{probability that a random susceptible high-activity person becomes infected on day } 1. \]

Our target person has a Poisson number of interactions during the day, with mean \( \lambda_H \). Given that the Poisson random variable takes on experimental value \( i \) on day 1, the number of infected interactions is equivalent to a sequence of \( i \) independent Bernoulli coin flips, with probability \( \beta \) that any given flip will correspond to interaction with an infected person. If \( j \) of the \( i \) interactions are with infected persons, then the conditional probability that our target susceptible person becomes infected is one minus the probability that none of the interactions with infected persons leads to infection. Placing this into equations, we can write

\[
p_H^i(1) = \sum_{i=0}^{\infty} \frac{\lambda_H^i}{i!} e^{-\lambda_H} \left[ \frac{1}{(1 - \beta)^{i-j}} \right] \left( \frac{1 - p}{1 - \beta} \right)^j.
\]

\[ p_H^i(1) = 1 - \sum_{i=0}^{\infty} \frac{\lambda_H^i}{i!} e^{-\lambda_H} \left[ \frac{1}{(1 - \beta)^{i-j}} \right] \left( \frac{1 - p}{1 - \beta} \right)^j,
\]

\[ p_H^i(1) = 1 - \sum_{i=0}^{\infty} \frac{\lambda_H^i}{i!} e^{-\lambda_H} \left[ 1 - \beta \right] \left[ \frac{1 - p}{1 - \beta} \right]^j.
\]

\[ p_H^i(1) = 1 - \sum_{i=0}^{\infty} \frac{\lambda_H^i}{i!} e^{-\lambda_H} \left[ 1 - p \right]^j,
\]

\[ p_H^i(1) = 1 - p(\beta, \text{ Poisson})^j \approx 1 - e^{-\lambda_H \beta p},
\]

Here, the terms \( p^i_H(z \mid \beta, \text{ binomial}) \) and \( p^i_H(z \mid \lambda_H, \text{ Poisson}) \) represent the discrete or \( z \)-transforms of the binomial and Poisson probability mass functions, respectively, where in general for any probability mass function \( \Pr(N = n) = p_n \), its discrete transform is defined as

\[ p^i_H(z) = \sum_{n=0}^{\infty} p_n z^n \quad \text{for } |z| \leq 1. \]

Summarizing, the key result is the probability that a random susceptible high-activity person becomes infected on day 1 is equal to

\[ p_H^i(1) = 1 - e^{-\lambda_H \beta p}. \]

Equivalently, for low-activity persons, we obtain the probability that a random susceptible low-activity person becomes infected on day 1,

\[ p_L^i(1) = 1 - e^{-\lambda_L \beta p}. \]

We can expand Equation (1) in a Taylor series,

\[ p_H^i(1) = 1 - e^{-\lambda_H \beta p}
\]

\[ = 1 - \{ 1 - \lambda_H \beta p + (1/2)(\lambda_H \beta p)^2 + \cdots \}. \]

If we have \( \lambda_H \beta p \ll 1 \), then we can simplify,

\[ p_H^i(1) = \lambda_H \beta p + \text{higher-order terms}. \]

When \( \lambda_H \beta p \ll 1 \), the probability \( p_H^i(1) \) that a susceptible high-activity person becomes infected on day 1 is given by the product of the infection probability per contact, \( p \), the mean number of contacts per day, and the conditional probability that a random contact is with an infectious asymptomatic person, i.e.,

\[ p_H^i(1) \approx p \lambda_H [\lambda_H n_H(1) + \lambda_L n_L(1)]/(\lambda_H n_H + \lambda_L n_L), \]

\[ p_H^i(1) \approx p \lambda_H \beta. \]

Equivalently, we have

\[ p_H^i(1) \approx p \lambda_L [\lambda_H n_H(1) + \lambda_L n_L(1)]/(\lambda_H n_H + \lambda_L n_L), \]

\[ p_H^i(1) \approx p \lambda_L \beta. \]

We must remember that Equations (1) and (2) are exact results and that Equations (3) and (4) are approximations that apply only when \( \lambda_H \beta p \ll 1 \) and \( \lambda_L \beta p \ll 1 \), respectively.

Note the product \( p \lambda_H p \lambda_L \) in Equations (1) and (2).

In terms of reducing incidence of the disease, this says that reducing \( \lambda \) by 50% is equivalent to a vaccine that is 50% effective (i.e., one that reduces \( p \) by 50%). Thus, social activity is just as important as a vaccine in spreading or reducing the spread of the infection.

Now we examine the concept of \( R_0(1) \) (basic reproductive ratio) within the context of our simple model. \( R_0(1) \) for day 1 is simply the ratio of the number who are newly
infected on day 1 and thus infectious on day 2 to the number who were already infectious on day 1, i.e.,

$$R_0(1) = \left[ p_H^0(1)[n_H - n_L^1(1)] + p_L^0(1)[n_L - n_L^1(1)] \right]/\left[ n_H^1(1) + n_L^1(1) \right].$$  \tag{5}$$

Recalling from Equations (1) and (2) that $p_H^0(1) = 1 - e^{-\lambda_H t}$ and $p_L^0(1) = 1 - e^{-\lambda_L t}$, we have

$$R_0(1) = [(1 - e^{-\lambda_H t}) [n_H - n_H^1(1)] + (1 - e^{-\lambda_L t})$$

$$\cdot [n_L - n_L^1(1)]]/[n_H^1(1) + n_L^1(1)].$$  \tag{6}$$

We now work through an illustrative numerical example to compute $R_0(1)$ for day 1. Suppose that we have a town of 20,000 residents with the following as input data:

$p = 0.1,$

$n_H = 10,000, \quad n_L = 10,000,$

$\lambda_H = 50, \quad n_H^1(1) = 5,$

$\lambda_L = 5, \quad n_L^1(1) = 5.$

Then, we obtain

$$\beta = [\lambda_H n_H^1(1) + \lambda_L n_L^1(1)]/(\lambda_H n_H + \lambda_L n_L)$$

$$= [50(5) + 25]/(50[10,000] + 5[10,000])$$

$$= 275/550,000 \approx 0.0005,$$

$$R_0(1) = [(1 - e^{-\lambda_H t}) [n_H - n_H^1(1)]$$

$$+ (1 - e^{-\lambda_L t}) [n_L - n_L^1(1)]]/[n_H^1(1) + n_L^1(1)],$$

$$R_0(1) = [(1 - e^{-50(0.000501)}10,000 - 5]$$

$$\cdot [10,000 - 5]/[5 + 5],$$

$$R_0(1) \approx [(0.0025)(9,995) + (0.000025)(9,995)]/10,$$

$$R_0(1) \approx [24.99 + 2.50]/10 \approx 2.75.$$ 

In other words, an “average” infectious person on day 1 newly infects an average of 2.75 persons on day 1, who themselves becomes infectious on day 2. But note that approximately 91% of this factor is due to high-activity infectious people on day 1. That is, the five high-activity infectious day 1 people created 25 new infections. Their basic reproductive ratio is 5.0. The five low-activity people created only 2.5 new infections. Their basic reproductive ratio is 0.5. While the average is 2.75, one can see how misleading an average such as this can be.

In this numerical example, because $\lambda_H \beta p \ll 1$ and $\lambda_L \beta p \ll 1$, the linear approximation works, and because the day 1 infected populations are infinitesimal fractions of the total population (i.e., $n_H^1(1) \ll n_H$ and $n_L^1(1) \ll n_L$), Equation (5) simplifies to

$$R_0(1) \approx \frac{p_H^0(1) + \lambda_L n_L^1(1)}{n_H^1(1) + n_L^1(1)}.$$  \tag{7}$$

This equation is intuitively appealing. It states that the ratio of the number of new day 1 infections to the number of those who are infectious on day 1 is the product of the per-contact infection probability $p$ and the mean number of contacts by infected persons in day 1 with the susceptible population, divided by the total number of infectious persons on day 1. Using Equation (5), we (again) obtain

$$R_0(1) \approx \frac{0.1[50(5) + 5(5)]}{[5 + 5]} = 27.5/10 = 2.75.$$ 

Note that on day 2, we will have 27.5 newly infected individuals who, now as infectious people, will infect susceptible people on day 2, who in turn will infect others on day 3, etc. Recalling our earlier discussion, a random interaction on day 1 by an individual $R$ with another will be with a high-activity person with probability

$$n_H \lambda_H/(n_H \lambda_H + n_L \lambda_L)$$

$$= (10,000)(50)/(10,000)(50) + (10,000)50$$

$$= 0.909/550,000 = 0.0909.$$ 

Thus, in a mean-value sense, 0.909(27.5) = 25 of the 27.5 newly infected people are highly socially active, with rate parameter $\lambda_H$. Low social activity levels characterize only 2.5 of the newly infected. This is troublesome, as the high-activity people are much more efficient at spreading the virus.

5.2. Day 2

We now move to day 2, requiring the following quantities:

$$n_H^2(2) = \quad \text{day 2 population of high-activity infectious persons} = p_H^0(1)[n_H - n_H^1(1)].$$

$$n_L^2(2) = \quad \text{day 2 population of low-activity infectious persons} = p_L^0(1)[n_L - n_L^1(1)],$$

$$n_H^2(2) = n_H - n_H^1(1) - n_H^1(2) = \quad \text{day 2 population of high-activity susceptible persons},$$

$$n_L^2(2) = n_L - n_L^1(1) - n_L^1(2) = \quad \text{day 2 population of low-activity susceptible persons.}$$

The iteration continues as on day 1. Let us continue with our numerical example:

$$n_H^2(2) = 25,$$

$$n_L^2(2) = 2.5,$$

$$\beta(2) = [\lambda_H n_H^2(2) + \lambda_L n_L^2(2)]/(\lambda_H n_H(2) + \lambda_L n_L(2))$$

$$= [50(25) + 5(2.5)]/[50(9,995) + 5(9,995)]$$

$$= 0.002293,$$

$$n_H^2(2) = n_H - n_H^1(1) - n_H^1(2) = 10,000 - 5 - 25 = 9,970,$$

$$n_L^2(2) = n_L - n_L^1(1) - n_L^1(2) = 10,000 - 5 - 2.5 = 9,992.5,$$

$$R_0(2) = [(1 - e^{-\lambda_H t})[n_H - n_H^1(1) - n_H^1(2)]$$

$$\cdot (1 - e^{-\lambda_L t})[n_L - n_L^1(1) - n_L^1(2)]$$

$$\cdot (n_H^1(2) + n_L^1(2))^{-1},$$
\[ R_0(2) = [(1 - e^{-50 (0.002293)^0.1}) (10,000 - 5 - 25] + (1 - e^{-5 (0.002293)^0.1}) [10,000 - 5 - 2.5)] \cdot [25 + 2.5]^{-1}, \]
\[ R_0(2) \approx [(0.01114)(9.970) + (0.00114)(9.992.5)]/27.5, \]
\[ R_0(2) \approx [113.65 + 11.45]/27.5 \approx 4.55. \]

This last number is quite troubling, as we see the basic reproductive ratio growing over time due to the highly socially active subpopulation.

5.3. Day 3

\[ n_H^I(3) = 113.6, \]
\[ n_L^I(3) = 11.45, \]
\[ n_H^R(3) = n_H - n_H^I(1) - n_H^I(2) - n_H^I(3) \approx 9.856, \]
\[ n_L^R(3) = n_L - n_L^I(1) - n_L^I(2) - n_L^I(3) \approx 9.981, \]
\[ R_0(3) \approx 4.43. \]

In this first model, we see that the population’s basic reproductive ratio is quite sensitive to time since first infections. It changes substantially on a day-to-day or generation-to-generation basis. The day 1 average basic reproductive ratio has little to do with disease dynamics after day 1.

In running this model on an Excel spreadsheet, we find that \( R_0(i) \) has already peaked on day 2, and over the next five days yields values of 3.93, 2.49, 0.88, 0.48, and 0.40, respectively. The plummeting values of \( R_0(i) \) are due to the high-activity susceptible population members becoming infected, then being infectious, and then leaving the circulating population. The residual susceptible population has far less social activity and thus an effective reproductive number well below 1.0. Note that the gross reproductive number varies markedly generation-to-generation, even in the absence of deliberate social controls such as social distancing, due only to the heterogeneity in the population. In this model run, virtually all of the high-activity susceptibles become infected within seven generations. Eventually, about 70% of the low-activity susceptibles become infected. Most of this latter infectivity is due to high-activity infectious people having contact with low-activity people. Once the high-activity people have left the circulating population, the gross reproductive number is well below 1.0 and very few additional low-activity susceptibles become infected as the disease dies away geometrically over time.

5.4. Generalization of Model 1

We now generalize this first model. Instead of only two activity levels, low and high, let us assume an arbitrary number \( J \) of activity levels, with the Poisson rate of interactions for activity level \( j \) being \( \lambda_j, j = 1, 2, \ldots, J \). We now define the following:

Initial inputs (or boundary conditions):

\[ \lambda_j \equiv \text{Poisson daily rate of interactions of a person having activity level } j, j = 1, 2, \ldots, J. \]
\[ n_j^I(0) = \text{initial population of people, all susceptible, in activity level } j, j = 1, 2, \ldots, J. \]
\[ n_j^S(0) = \text{total susceptible population.} \]
\[ n_j^I(1) = \text{initial number of day 1 infectious people in activity level } j, j = 1, 2, \ldots, J. \]
\[ p \equiv \text{probability that a susceptible person becomes infected upon a random interaction with an infectious person.} \]

Moving from day \( (i - 1) \) to day \( i, i = 1, 2, \ldots, \) calculate, in order:

\[ n_j^I(i) = n_j^I(i - 1) - n_j^I(i) = \text{number of susceptible people in activity level } j \text{ on day } i, i = 1, 2, \ldots \]
\[ \beta(i) = \text{probability that on day } i \text{ a random interaction is with an infectious person,} \]
\[ \beta(i) = \frac{\sum_{j=1}^{J} \lambda_j n_j^I(i)}{\sum_{j=1}^{J} \lambda_j n_j^S(i)} \cdot \quad i = 1, 2, \ldots. \]
\[ p_j^S(i) \equiv \text{probability that a susceptible person in activity level } j \text{ becomes infected on day } i. \]
\[ p_j^S(i) = 1 - e^{-\lambda_j \beta(i) p}. \]
\[ p_j^S(i) \equiv \lambda_j \beta(i) p \text{ \quad if } \lambda_j \beta(i) p \ll 1. \]
\[ R_0(i) = \frac{\sum_{j=1}^{J} p_j^S(i) n_j^I(i)}{\sum_{j=1}^{J} n_j^I(i)} \]
\[ n_j^S(i + 1) = p_j^S(i) n_j^S(i). \]

GO TO NEXT DAY BY INCREASING INDEX \( i \) BY ONE.


Model 1 assumes the same social contact rate per day, even as members of the susceptible population become sick and leave the circulating population. We may want to alter this assumption to account for an anticipated reduced amount of social activity. For example, if a salesman usually has 40 contacts per day when everyone is healthy, we might expect a reduction of 20% in contacts per day if 20% of the population is sick and out of circulation. An alternative to model 1 is one in which \( \lambda \), the average number of daily contacts per person, is proportional to the remaining population in circulation, both susceptible and infectious.

This idea gets us to model 2, a generalized model 1 with population-dependent \( \lambda \):

Initial inputs:

\[ \lambda_j \equiv \lambda_j(1) \equiv \text{day 1 Poisson daily rate of interactions of a person having activity level } j, j = 1, 2, \ldots, J. \]
\[ n_j^S(0) = \text{initial population of people, all susceptible, in activity level } j, j = 1, 2, \ldots, J. \]
7. Model 3: Social Distancing and Hygienic Factors

Many, perhaps most models of influenza progression, assume that social behavior of the susceptible population will remain unchanged during the course of the pandemic. History strongly suggests otherwise. As soon as it becomes clear that a serious influenza is in or even near the community, parents pull their children out of school, some professionals start to telecommute, discretionary travel and shopping are reduced—to name a few of such personally selected moves toward social distancing. In addition, governments implement social-distancing controls, closing many public places and events, staggering work hours, etc. Finally, hygienic behaviors are encouraged, such as covering one’s mouth when coughing or sneezing, frequent washing of hands with soap and hot water, wearing face masks, etc. In our stylized simple model, we can depict social distancing and enhanced hygienic behavior by reducing the product, “lambda” times $p_j$, the conditional probability of infection, once exposed, for activity $j$ level susceptibles. We introduce a “social distancing and hygienic factor” $d_j(i)$ for activity level $j$. $d_j(i) \leq 1$. This factor, always less than or equal to 1.0, is used to multiply the product $\lambda_j(i)p_j$ in the equation that computes the probability that a random activity level $j$ susceptible will become infected on day $i$. Here are the details:

**Initial inputs (or boundary conditions):**

\[
\lambda_j(1) = \text{initial (day 1) Poisson daily rate of interactions of a person having activity level } j, \quad j = 1, 2, \ldots, J.
\]
\[
n^s_j = n^s_j(0) = \text{initial population of people, all susceptible, in activity level } j, \quad j = 1, 2, \ldots, J.
\]
\[
n_j = n^s_j(0) = \sum_{j=1}^{J} n^s_j(0) = \text{total susceptible population.}
\]
\[
n_j^r(1) = \text{initial number of day 1 infectious people in activity level } j, \quad j = 1, 2, \ldots, J.
\]
\[
n^t = n^t(0) = \sum_{j=1}^{J} n_j^r(i) = \text{total initial infected population.}
\]
\[
n = n(0) = n^s + n^t = \text{total initial population, infected and susceptibles, in circulation.}
\]
\[
p = \text{probability that a susceptible person becomes infected upon a random interaction with an infectious person.}
\]

**Moving from day $(i-1)$ to day $i$, $i = 1, 2, \ldots$, calculate, in order:**

\[
n_j^r(i) = n_j^r(i-1) - n^r_j(i) = \text{number of susceptible people in activity level } j \text{ on day } i, \quad i = 1, 2, \ldots.
\]
\[
\beta(i) = \text{probability that on day } i \text{ a random interaction is with an infectious person.}
\]
\[
\beta(i) = \frac{\sum_{j=1}^{J} \lambda_j(i)p_j(i)}{\sum_{j=1}^{J} \lambda_j n_j^s(i-1)}, \quad i = 1, 2, \ldots.
\]

**GO TO NEXT DAY BY INCREASING INDEX $i$ BY ONE.**

When operating this modification of the model on the Excel spreadsheet, it turns out that the reduction in social contacts per day due to reduced numbers of susceptibles in the circulating population does have certain first-order effects. As with model 1, nearly all of the high-activity people still become infected, but the number of low-activity susceptibles who eventually become infected is reduced from about 70% to about 48%. The $R_0(i)$ values for days 1 through 6 are 2.75, 4.55, 4.42, 3.90, 2.44, and 0.85, respectively.
Models do numbers high-activity policies. The use of these is a result of another propagation of the pandemic. The number of Susceptibles is reduced by 50-70% through social distancing and other measures. During the peak of the pandemic in day 3, the number of infections is dramatically reduced through these controls. The role of "," lambda-p" factor by 50%, but we start that on day 5. We do nothing with the level 2 and 3 susceptibles. In this case, the peak day or peak generation is extended far into the future, to day 22, but the numbers who do not become infected are impressive: level 1: 28,851; level 2: 52,783; and level 3: 84,333. If we introduce these level 1 policies on day 3 of the pandemic, again leaving alone susceptibles in levels 2 and 3, the numbers who do not become infected are improved only marginally: level 1: from 28,851 to 29,081; level 2: from 52,783 to 52,985; and level 3: from 84,333 to 84,419. To obtain more infection control in this case, we would have to introduce social distancing and hygiene controls for level 2 and perhaps also for level 3.

At least two additional sources of heterogeneity are worth considering in this class of simple models: for fixed levels of social activity (1) the likelihood of becoming infected, and (2) the likelihood of spreading the disease, once infected. The medical community has verified over many years of research that the likelihood of becoming infected—given a certain pattern of contacts with one or more infected persons—depends on age, degree of immune system compromise, and other factors. Similarly, it is widely believed that certain individuals shed virus at higher rates than others and thus are likely to infect more susceptibles than the average virus-shedding person. The online appendix (available as part of the online version that can be found at http://orjourn.informs.org/) contains two additional generalizations of the model, one dealing with heterogeneous susceptibility across the population and the other with heterogeneous infectivity. While adding complexity by increasing the numbers of partitions of the population, the models again demonstrate the key role played within each group by the product of "lambda" (the daily contact rate) and the associated probability of becoming infected and/or of passing the infection to others.

We have carried out extensive analyses with these simple spreadsheet difference equation models. We have been able to replicate other behaviors related to social controls that were found in the 1918–1919 Spanish Flu. Among these is the need to retain controls until the virus has left the population. If controls are relaxed too early, the model shows and history demonstrates that another wave of the flu will return.

9. Frequencies of Daily Human Contacts: Empirical Results

To use confidently a model of the type presented here, one would need data indicating the frequency distribution of
daily human contacts in the region being modeled. The best data we have seen about this are from Yang-chih Fu, who in nine countries and 46 different settings asked the following single question to randomly selected residents:

On an average, about how many people do you have contact with in a typical day, including all those who you say hello, chat, talk or discuss matters with, whether you do it face-to-face, by telephone, by mail or on the Internet, and whether you personally know the person or not? Please give your estimate and select one from the following categories that best matches your estimate:

(1) 0–4 persons, (2) 5–9 persons, (3) 10–19 persons, (4) 20–49 persons, (5) 50–99 persons, (6) over 100 persons. (Fu 2005, 2007)

This question is not precisely suited to our needs, as it includes human contacts that are not face-to-face, such as telephone, mail, and the Internet. So, we need to apply care and caution in the interpretation of the results. For the United States in 2005, Fu asked the single question above via a telephone survey of a representative sample of the working population of the United States, ages 21–64, in urban areas (n = 3,000). The United States results are shown in the following table:

<table>
<thead>
<tr>
<th>Number of daily contacts</th>
<th>Number of respondents</th>
<th>Percent of respondents</th>
<th>Cumulative percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4</td>
<td>410</td>
<td>13.67</td>
<td>13.67</td>
</tr>
<tr>
<td>5–9</td>
<td>426</td>
<td>14.20</td>
<td>27.87</td>
</tr>
<tr>
<td>10–19</td>
<td>685</td>
<td>22.83</td>
<td>50.70</td>
</tr>
<tr>
<td>20–49</td>
<td>792</td>
<td>26.40</td>
<td>77.10</td>
</tr>
<tr>
<td>50–99</td>
<td>349</td>
<td>11.63</td>
<td>88.73</td>
</tr>
<tr>
<td>100+</td>
<td>338</td>
<td>11.27</td>
<td>100.00</td>
</tr>
<tr>
<td>Totals</td>
<td>3,000</td>
<td>100.00</td>
<td></td>
</tr>
</tbody>
</table>

We may ask how many of these contacts were other than face-to-face. The simple answer is that we do not know. But Fu has data from Taiwan indicating a distribution of daily contacts not dissimilar to that for the United States in which he found that about 83% of all daily Taiwanese contacts are face-to-face. The others were 15% telephone, 1.5% the Internet, and 0.5% mail. We do not know if these figures carry over to the United States, but for the purposes of this paper, which uses illustrative data and does not claim to undertake “production runs” of a proven model, Fu’s results are instructive. The results point to a wide range of contacts over members of the population, suggesting significant heterogeneity. Over 11% of the U.S. urban working population has over 100 contacts per day, whereas 14% have four or fewer daily contacts. Even allowing for the confounding factors of telephone, mail, and Internet contacts, the results are striking in their wide dispersion. Any epidemiological model that ignores such significant heterogeneity in the population risks missing valuable insights to guide public policies.

10. Discussion and Future Directions

In this paper, we have suggested a family of simple yet plausible difference equation models to depict the progression of an influenza infection through a population in the presence of population heterogeneity and infection control via social-distancing and hygienic steps. A probabilistic mixing model has been offered to depict nonuniform mixing of subpopulations. Within the modeling assumptions, we derived an expression for the probability that a random susceptible person in a specified category will become infected on a given day. The model is deliberately stylized and simplistic, offered in this form to make several points:

(1) In the presence of population heterogeneity, even without any changes in social behavior, there exists no constant basic “reproductive ratio $R_0$” characterizing the mean number of new influenza infections created by each newly infected person. Because high-activity persons are more likely to become infected early and thus differentially removed from the population of susceptibles at an accelerated rate, the generation-to-generation reproductive ratio changes dramatically compared to standard models using the $R_0$ concept.

(2) Attempts to estimate empirically “first days” values of $R_0$ from earlier pandemics may be biased by the dominant effects of high-activity persons whose group $R_0$ value could markedly exceed any population average.

(3) In the presence of changes in social behavior associated with attempts to reduce the spread of infection, the generation-to-generation reproductive ratio changes even more dramatically. Within the context of attempted social controls, the generation-to-generation reproductive ratio should be viewed as an outcome of the model and not as an input.

(4) In any subpopulation $j$ characterized by frequency $\lambda_j$ of social contacts and probability $p_j$ of contracting the disease, given contact, the product $\lambda_j p_j$ is the primary determining factor associated with disease spread associated with that subpopulation.

(5) When the population can be usefully partitioned into heterogeneous groups, distinguished by frequency of social contacts and infection probabilities, given contact, public policies to control the disease may best be devoted to those groups having highest values of the product $\lambda_j p_j$. These policies can aim to reduce the rate of social contacts $\lambda_j$ or the conditional probability of infection given contact $p_j$, or both.

(6) Spreadsheet analyses with alternative versions of the model have shown the importance of implementing social distancing and related hygienic controls early during the outbreak and keeping those measures in force for an extended period of time. Implementing social distancing and hygienic controls too late, while still helpful, suggests missing opportunities to avoid numerous infections and possible deaths. Withdrawing them too early may increase the generation-to-generation reproductive ratio to
above 1.0, meaning that a second wave of the disease will occur.

The simple difference equation models introduced here are clearly only the beginning of a new decision-oriented approach to modeling the progression of a pandemic of an infectious disease such as severe influenza. Over time, models such as these and their generalizations will have strong spatial components that depict in realistic ways the manner in which infectious persons travel from one population center to another. No doubt, theorems will be proved indicating necessary and sufficient conditions for the epidemic to be contained within specified bounds.

No one knows how or when the next pandemic influenza will emerge and what its intrinsic properties will be. If history can be a guide, the next influenza will have “emergent properties,” meaning that it will mutate during the course of the epidemic and its intrinsic properties will evolve accordingly. Any mathematical model of the disease and its control is bound to be incorrect. We are not seeking multidimensional numerical accuracy, but rather insights on how to limit the spread of the disease. We firmly believe that fresh eyes from the OR community can play a significant role in this quest.

11. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://or.journal.informs.org/.

Endnotes

1. We are aware of efforts worldwide to produce a revolutionary new flu vaccine. At the time of this writing, two promising efforts are by the European biotech firms Acambis and Cytos Biotechnology (http://www.dailymail.co.uk/pages/live/articles/health/healthmain.html?in_article_id=425227&in_page_id=1774 (January 15, 2007)).

2. This policy of Indian farm families was presented to the author by Dr. Nitin Patel, whose father reported that tradition to him. Dr. Patel’s father was born in 1909 and lived in the rural village of Karamsad, state of Gujarat, India. Once as a boy he had to leave the village with his family to avoid “the plague.” Our hypothesis is that the terminology “the plague” related to several different serious and sometimes fatal diseases and did not precisely refer to any specific plague such as the bubonic plague.

3. All of the state plans are accessible at http://www.pandemicflu.gov/plan/states/stateplans.html (January 7, 2007).

4. Technically, \( R_0 \) pertains only to generation zero, when all others in the population are susceptibles. But usually several generations are required until any significant fraction of the population is infected, so \( R_0 \) remains a valid approximation in standard models for early generations. After that, we see in the literature \( R(t) \), which is the generation-to-generation multiplier for generation \( t \). Formally, \( R(0) = R_0 \). Our remarks about the limitations of \( R_0 \) pertain to \( R(t) \) as well. In our modeling work, we retain the subscript zero to emphasize the ties to \( R_0 \), so in our terminology \( R_0(t) = R(t) \).


6. While the number 25 occurs twice in our day 1 numerical example—first, as the number of newly infected people on day 1 who were infected by high-activity infectious people, and second, as the number of high-activity newly infected people on day 1—the sets of individuals involved in these two cases are not in general identical. Some of the high-activity infectious people on day 1 could have infected low-activity people, and some of the low-activity infectious people on day 1 could have infected high-activity people.

Acknowledgments

This paper was motivated by a team effort of colleagues at MIT, seeking to understand the history of influenza pandemics and the key features required in related decision-focused modeling. The author thanks each of the team members: Kelley M. Bailey, Stan Finkelstein, Karima Robert Nigmatulina, Robert Rubin, and Katsunobu Sasana. He also thanks Marc Lipsitch, Clark Abt, John L. Wyatt Jr., Yang-chih Fu, Kimberly M. Thompson, and Laura Zager for helpful suggestions. This paper was supported in part by an IBM Faculty Research Award.

References


For comments and discussion about this paper, readers are referred to the online discussion forum available through the Operations Research website at http://or.pubs.informs.org, or at http://orforum.blog.informs.org/.