CHAPTER 5

Using Mathematical Modeling to Integrate Disease Surveillance and Global Air Transportation Data

Julien Arino

Department of Mathematics, University of Manitoba, Winnipeg, MB, Canada

Kamran Khan

Centre for Research on Inner City Health, St Michael's Hospital, Toronto, ON, Canada

5.1 INTRODUCTION

Because of the relationship between the movement of populations and the spread of infectious diseases, it is important to understand and model mobility. Note that we focus here on the mobility of human populations; consideration of the movement of animal or vector populations is also critical but is beyond the scope of this work.

Populations are increasingly mobile. Simplifying the situation to the extreme, mobility takes two major forms: migration and travel. Migration is mobility in the long term, where an individual changes their place of residence. Travel is a shorter term mobility, where an individual usually keeps the same place of residence. An intermediate form of consequence to public health is the case of migrant workers, both within and between countries.

The main fluxes of immigration form a gradient from poorer to richer countries. For example, from 2002 to 2011, four countries each contributed more than 100,000 new permanent residents to Canada: China, India, Pakistan and the Philippines,

Analyzing and Modeling Spatial and Temporal Dynamics of Infectious Diseases, First Edition. Edited by Dongmei Chen, Bernard Moulin, and Jianhong Wu.

making up 37% of the almost 2.5 million new Canadian permanent residents in that period (Research and Evaluation Branch CIC 2011). Because health care systems vary considerably, migrants present specific challenges to public health systems in their destination countries, for instance, because of different immunization schedules or practices, prevalence of diseases such as tuberculosis.

However, migration fluxes have become secondary in volume to travel fluxes. For instance, in 2010, Canada saw 115,271 temporary residents (work visas, students, etc.) make their initial entry into the country and had 280,691 new permanent residents. The same year, 19,360,480 airline trips originated in the rest of the world and terminated in Canada. These trips include not only those of new immigrants, whether temporary or permanent, but also trips of residents of Canada abroad and tourist or business visits to Canada. Note in particular that because travel has become easier and cheaper, there is a good amount of post-migration flux, with immigrants returning for visits to their country of origin much more frequently than used to be the case.

Therefore, public health issues related to mobility cannot be considered any longer as a problem that a country has to deal with only at the time of first entry of a migrant. Also, the continual flow of individuals between countries should be taken into account. This is true in particular concerning emerging and reemerging diseases.

Indeed, perhaps the most important teaching of the 2003 SARS epidemic concerns the potentially disastrous consequences of the globalization and acceleration of travel on global public health security. SARS was exemplary of the ability of an emerging disease to spread very fast over large distances. SARS also illustrated the ever increasing role of commercial aviation in the spread of emerging and reemerging infections: of the documented 137 SARS cases that are known to have crossed state boundaries, 129 traveled by plane.

Further confirmation of the role of travel came in 2009, with the H1N1 influenza pandemic (pH1N1). In Khan et al. (2009), the relationship between the number of passengers inbound from Mexico in a two month period and the likelihood of importation of cases of pH1N1 was studied. It was found that cities connected to an airport that had received more than 1,400 passengers from Mexico in March and April 2008 (used as a proxy for the 2009 travel data, which was not available at the time) were at a greatly elevated risk of importing the disease.

Because of the increasingly interconnected nature of public health issues, traditional surveillance has been complemented in recent years by internet trawling surveillance systems such as Global Public Health Intelligence Network (GPHIN) or HealthMap. These systems continuously monitor internet news sources in a variety of languages to generate alerts concerning public health threats. However, these systems have the drawback that they generate a very high number of alerts.

We discuss here a method for prioritizing these alerts in terms of the risk they represent to a given public health entity, using mathematical modeling and information about the global air transportation network. This is work carried out in the context of the BioDiaspora Project and follows ideas proposed in Khan et al. (2012).

5.2 THE NETWORK

The BioDiaspora Project focuses on air travel, although it also documents "ground conditions" in order to assess risk. See Arino et al. (2011); Khan et al. (2009) for more detail about the air transportation data. Here, we only mention that the data used is from IATA (the International Air Transport Association) and details most trips taken worldwide from 2005 to 2012, including up to 5 intermediate stops.

The data has a resolution of one month. As a consequence, it is important to take time into account since travel volumes vary widely depending on the period of the year. So, in all considerations that follow, it should be understood that graphs evolve with time.

Connections between airports are represented by an $N \times N$ matrix of volumes detailing, for any pair $i, j = 1, \ldots, N$, the volume v_{ji} of travel from airport i to airport j. We denote $\mathcal{V}^I$ as this matrix. Corresponding to this matrix, $\mathcal{G}^I(t)$ is the graph obtained from the IATA data.

5.3 AIRPORT CATCHMENT AREAS

For the model, it is necessary to have an estimate of the population situated within the so-called *catchment area* of this airport, that is, that uses this airport for its international transportation needs. Because of the nature of the transport data, we use airport catchment areas (ACAs) as the units of the analysis.

Since airports are located throughout the world, it is unrealistic to gather information about ACAs manually, in particular, concerning their population. In order to gather this information automatically, we use a *weighted Dirichlet tessellation* of the plane. This proceeds as follows (see, e.g., Ash and Bolker (1986)). Let $\mathcal{P}$ be a finite set of points on a sphere, the *sources*. For each pair of points $P, Q \in \mathcal{P}$, define

$$H_{PQ} = \left\{ X : \frac{|X - P|}{\sigma(P)} \leq \frac{|X - Q|}{\sigma(Q)} \right\}$$

where $\sigma(P) > 0$, and

$$K_{PQ} := H_{PQ} \cap H_{QP} = \left\{ X : \frac{|X - P|}{\sigma(P)} = \frac{|X - Q|}{\sigma(Q)} \right\}.$$

For each $P \in \mathcal{P}$, let $R_p = \bigcap_{Q \neq P} H_{PQ}$ and $R = \{R_P, P \in \mathcal{P}\}$. Then $R(\mathcal{P})$ is the Dirichlet (or weighted Voronoi) tessellation of the sphere. If the weight function $\sigma(P) = 1$ for all P, then in the plane, the regions are polygons and the result is often called a Voronoi diagram (Thiessen polygons in the geographical literature).

Limitations of the classic weight function $\sigma(P) = 1$ are that the importance of the airports under consideration is not taken into account. Using weights equal to the

volume of trips out of airports overemphasizes major airports, so we use a Holling type 2 function of the form

$$\sigma(v_i) = v_{\max}(t)\frac{v_i(t)}{v_i(t) + v_{\text{med}}(t)},$$

where $v_{\max}(t)$, $v_{\text{med}}(t)$ and $v_i(t)$ are the volume out of the busiest airport, median volume, and volume out of the airport i under consideration, respectively, from the IATA database. The tessellation is computed for every month in the database, since the relative importance of airports varies monthly.

Note that the results obtained using this method are not meant to represent the exact location where people using the airports live but rather, provide an estimation of the population relying on a given airport for long distance travel.

5.4 MODELING

Because of the nature of the travel data, we consider airports and their catchment areas as the units of analysis. We describe the model in three steps: (1) the epidemiology in airport catchment areas; (2) the description of transport; and (3) the integration of both.

5.4.1 The Model in Airport Catchment Areas

The model in each ACA $i = 1, \ldots, N$ is an SLIAR model, which has individuals in one of the epidemiological states susceptible, latent, symptomatically and asymptomatically infectious and recovered, with numbers at time t in airport catchment area i denoted $S_i(t)$, $L_i(t)$, $I_i(t)$, $A_i(t)$, and $R_i(t)$, respectively. When this does not lead to ambiguities, the dependence of state variables and those parameters that are time-dependent on t is not indicated.

We describe briefly the model here; see Arino et al. (2006) for details about the deterministic system. The model used is an *epidemic* model, in that it considers one epidemic event in a population and neglects birth and death. Indeed, simulations are performed for a short time frame of one to several weeks, and variations of the total population during this duration are negligible compared with variations in the number of individuals in the different epidemiological compartments. The flow of individuals between the different compartments is assumed to happen as illustrated in Figure 5.1.

Susceptible individuals are potentially affected by the disease, if subject to an infecting contact. Such contacts occur at the rate S_iI_i between susceptible and symptomatically infectious individuals and S_iA_i between susceptible and asymptomatically infectious individuals. These contacts result in new infections at the rates $\beta_iS_iI_i$ and $\eta_i\beta_iS_iA_i$ for contacts with symptomatically and asymptomatically infectious individuals, respectively. β_i is the disease transmission coefficient in ACA i and $\eta_i \in [0, 1]$ is the reduction of transmission due to asymptomatic infection (i.e., we assume that asymptomatic infectious individuals are potentially less infectious than symptomatic

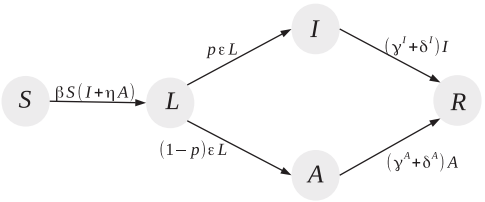

Figure 5.1 Flow diagram of the model in each ACA. Indices of variables and parameters are omitted

ones). This type of incidence is called *mass action* incidence. The disease transmission coefficient represents the probability that infection occurs, given contact. We allow it and all other parameters to vary from location to location, since factors such as hygiene, health care equipment and social distancing play a role in the transmission of the disease and vary widely from place to place.

Latent individuals are susceptibles who have become latently infected because of an infecting contact with an infectious individual. In the case of SARS, estimates of the median of the incubation period (the length of time between infection and the onset of symptoms) were of 4.0 days (95% CI 3.6–4.4) (Lessler et al. 2009), meaning that the inclusion of a class of exposed individuals is necessary. Other diseases have a much shorter incubation period (e.g., the same authors found medians of 1.4 and 0.6 days for influenza A and B, respectively) and might not require the inclusion of a latent period. However, as the system is designed for any emerging or reemerging disease, we always allow the possibility of latency (setting a very small value for the time spent incubating if need be). It is assumed that patients in the latent state do not transmit the disease. The time spent incubating is exponentially distributed with mean $1/\varepsilon_i$ time units.

After incubating, individuals progress either to a symptomatic or an asymptomatic infectious stage, with a proportion p_i becoming symptomatically infectious. Both infectious stages have individuals spreading infection, although it is generally believed that asymptomatically infectious individuals are less infectious to others than symptomatic ones, prompting the use of the *attenuation* coefficient η_i.

Infectious individuals (both symptomatic and asymptomatic) actively spread the infection through contacts with susceptible individuals. Symptomatic and asymptomatic infectious individuals remain infectious for an average $1/(\gamma_i^I + \delta_i^I)$ and $1/(\gamma_i^A + \delta_i^A)$ time units, respectively, with the sojourn time in the infectious classes exponentially distributed. Thus individuals are removed from the I and A classes either by recovery (at rates γ_i^I and γ_i^A, respectively) or by disease-induced death (at rates δ_i^I and δ_i^A, respectively). Note that we distinguish between recoveries and disease-induced death in order to be able to compare with data.

Finally, *removed* individuals are individuals who have ceased to be infectious. Hence, we interpret this class as in Kermack and McKendrick (1927). Individuals in the recovered class play no role in the short-term transmission of the disease, and thus we neglect this class from now on.

5.4.2 Movement Rates

To compute movement rates, we reason using ordinary differential equations (since they are readily converted to their continuous Markov chain equivalents). There are many ways to obtain the movement rates; we show here a method that is extremely simple yet provides a good description of the actual number of trips taken in a short period of time.

Consider two ACAs, say, those of Winnipeg (Manitoba, Canada, IATA code YWG) and Toronto (Ontario, Canada, aggregate IATA code YTO). We want to describe the actual number of trips between the two ACAs. For a short time interval of, say, 1 day, we can neglect other sources of variation of the population in the origin ACA as well as other flows due to travel to and from other ACAs. Thus, the variation of the population in Winnipeg because of trips to Toronto is given by

$$N'_{\mathrm{YWG}}(t) = -m_{\mathrm{YTO,YWG}}(t)N_{\mathrm{YWG}}(t),$$

where $m_{\mathrm{YTO,YWG}}(t)$ is the rate of movement of individuals from Winnipeg to Toronto at time t. Because IATA data is given per month, the rates are computed for each month (but with time units of 1 day). Thus, after 1 day, the population in Winnipeg has changed according to

$$N_{\mathrm{YWG}}(1) = e^{-m_{\mathrm{YTO,YWG}}}N_{\mathrm{YWG}}(0).$$

$N_{\mathrm{YWG}}(1) - N_{\mathrm{YWG}}(0)$ is the loss of population in Winnipeg from trips to Toronto in 1 day. In September 2012, for instance, this was an average of 844 people. Thus, solving for $m_{\mathrm{YTO,YWG}}$, we find

$$m_{\mathrm{YTO,YWG}} = -\ln\left(1 - \frac{844}{N_{\mathrm{YWG}}(0)}\right),$$

where $N_{\mathrm{YWG}}(0)$ is the population of Winnipeg obtained from the catchment area computation of Section 5.3.

More generally, trips from X to Y occur at the rate

$$m_{\mathrm{YX}} = -\ln\left(1 - \frac{\Delta_{YX}}{N_{\mathrm{X}}(0)}\right),$$

where Δ_{YX} is the number of trips per day originating in X and terminating in Y and $N_X(0)$ is the population of X obtained using the catchment area computation. Using the population information, travel data and setting diagonal terms so that $\mathcal{M}^N$ has all column sums zero gives the rates of movement between all pairs of ACAs.

Note that we could also have reasoned, for instance, using the volume of passengers received by the Toronto catchment area in a day or the proportion of trips to Toronto in the trips outbound from Winnipeg together with the rate of travel outbound from Winnipeg.

5.4.3 General Model of Infection-Transport

For simulations of the full model, we use continuous time Markov chains. These are indeed readily derived from the deterministic model and have the advantage of allowing discrete (integer) population numbers and incorporating stochasticity. The stochastic process of infection-transport can be derived in several ways, in particular, using infinitesimal probabilities. However, we show here only the most useful one for our purpose: the derivation in terms of times to transitions, since numerical simulations are run using the Gillespie algorithm (Gillespie 1977). Recall that we neglect the dynamics of removed individuals. Suppose that the system is, at time t, in the state

$$(s, l, i, a) = (s_1, \ell_1, i_1, a_1, \dots, s_N, \ell_N, i_N, a_N).$$

Then compute the weight of possible events

$$\begin{aligned}\xi_t = &\sum_{j=1}^{N} \left(\beta_j s_j (i_j + \eta_j a_j) + \varepsilon_j \ell_j + \left(\gamma_j^I + \delta_j^I\right) i_j + \left(\gamma_j^A + \delta_j^A\right) a_j\right) \\ &+ \sum_{j,k=1}^{N} \left(m_{jk}^S s_j + m_{jk}^L \ell_j + m_{jk}^I i_j + m_{jk}^A a_j\right).\end{aligned} \tag{5.1}$$

The next event occurs at time $t + \tau_t$, where τ_t is one realization of an exponentially distributed random variable with parameter ξ_t. At time $t + \tau_t$, the transition $(s, \ell, i, a) \to (s', \ell', i', a')$ occurs, where the new state (s', ℓ', i', a') corresponds to the following events. For simplicity, only the variables that are modified are indicated.

1. A susceptible is infected in ACA j, that is, $(\dots, s_j', \ell_j', \dots) = (\dots, s_j - 1, \ell_j + 1, \dots)$. This occurs with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = \beta_j s_j (i_j + \eta_j a_j)/\xi_t.$$

   Note that the model further allows to identify the origin (symptomatic or asymptomatic infectious individual) of the infection, if needed, by breaking the above probability down in terms of $\beta_j s_j i_j/\xi_t$ and $\beta_j \eta_j s_j a_j/\xi_t$.

2. A latently infected individual in ACA j develops the symptomatic form of the disease, that is, $(\dots, \ell_j', i_j', \dots) = (\dots, \ell_j - 1, i_j + 1, \dots)$. This occurs with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = p_j \varepsilon_j \ell_j/\xi_t.$$

3. A latently infected individual in ACA j develops the asymptomatic form of the disease, that is, $(\dots, \ell_j', a_j', \dots) = (\dots, \ell_j - 1, a_j + 1, \dots)$. This occurs with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = (1 - p_j)\varepsilon_j \ell_j/\xi_t.$$

4. An individual with the symptomatic form of the disease is removed in ACA j, that is, $(\ldots, i'_j, \ldots) = (\ldots, i_j - 1, \ldots)$. Such an event occurs with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = \left(\gamma_j^I + \delta_j^I\right) i_j / \xi_t.$$

   As for new infections, this event can be further broken down in terms of the number of recoveries and disease-induced deaths by considering two separate events with, respective, probabilities $\gamma_j^I i_j/\xi_t$ and $\delta_j^I i_j/\xi_t$.

5. An individual with the asymptomatic form of the disease is removed in ACA j, that is, $(\ldots, a'_j, \ldots) = (\ldots, a_j - 1, \ldots)$. Such an event occurs with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = \left(\gamma_j^A + \delta_j^A\right) a_j / \xi_t,$$

   with the event potentially broken down into recoveries and disease-induced deaths if needed, as explained for removal from the I class.

6. An individual currently in the susceptible class travels from ACA j to ACA k (with $k \neq j$), that is, $(\ldots, s'_j, s'_k, \ldots) = (\ldots, s_j - 1, s_k + 1, \ldots)$, with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = m_{kj}^S s_j / \xi_t.$$

7. An individual currently in the latent class travels from ACA j to ACA k (with $k \neq j$), that is, $(\ldots, \ell'_j, \ell'_k, \ldots) = (\ldots, \ell_j - 1, \ell_k + 1, \ldots)$, with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = m_{kj}^L \ell_j / \xi_t.$$

8. An individual currently with a symptomatic infection travels from ACA j to ACA k (with $k \neq j$), that is, $(\ldots, i'_j, i'_k, \ldots) = (\ldots, i_j - 1, i_k + 1, \ldots)$, with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = m_{kj}^I i_j / \xi_t.$$

9. An individual currently with an asymptomatic infection travels from ACA j to ACA k (with $k \neq j$), that is, $(\ldots, a'_j, a'_k, \ldots) = (\ldots, a_j - 1, a_k + 1, \ldots)$, with probability

$$\mathbb{P}_{(s,\ell,i,a)\to(s',\ell',i',a')} = m_{kj}^A a_j / \xi_t.$$

5.4.4 Initial Conditions

Setting initial conditions for the model involves several phases. In a first phase, the susceptible population in each ACA is set at the value obtained from the catchment area analysis of Section 5.3.

The second phase considers what is known about the disease of interest; a certain fraction of this susceptible population may indeed be assigned to the recovered class R because of preexisting immunity in the population. For instance, the WHO estimates (World Health Organization 2013) that the prevalence of immunity to measles varies, depending on countries, from 44% to 99% of the population (from the combined effect of vaccination and immunity acquired from infection), so that if measles were considered, the susceptible population in ACAs would be reduced by the amount corresponding to the prevalence of immunity to measles in the country that the airport belongs to.

The third phase involves setting initial conditions of the number of latently, symptomatically and asymptomatically infected individuals in the places where infection is known to occur. This is the phase in which the simulation system is tied in with the surveillance system.

5.4.5 Parameter Estimation

To choose parameter values, the durations of stages are known from the literature for many diseases. In the case of an outbreak of a disease for which specific parameters are not known, extensive simulations are carried out using parameters in typical ranges.

Because of the short time frame within which it operates, timing is essential in the present model. As a consequence, it is important to be careful when choosing values for the parameters that represent the mean duration of stages. For instance, recall that in Lessler et al. (2009), the median incubation period for SARS was estimated to be 4.0 days. Inherent to the formulation of the model is that the time spent in the latent class L_i for a given individual is an exponentially distributed random variable with mean $1/\varepsilon_i$ and median $\ln 2/\varepsilon_i$. Considering $\ln 2/\varepsilon_i = 4$ days ($\varepsilon_i \simeq 0.17$, i.e., a mean incubation period of 5.77 days) implies that in a cohort of individuals infected on a given day, 25% are still incubating 8 days later and more than 5% are still incubating after 15 days. So we also consider the converse problem: given the data on incubation periods, we determine a 95% "confidence interval" of time spent incubating. Say that, for example, 95% of individuals have become infective after 10 days. Then we find ε_i, the mean of the exponential distribution, by solving for ε the equation $\int_0^{10} \varepsilon_i e^{-\varepsilon_i s}\, ds = 0.95$, giving $\varepsilon_i \simeq 0.3$ (i.e., a much shorter mean incubation time of 3.33 days). We typically perform simulations with parameters in the range given by these two methods.

Estimating β is probably one of the hardest tasks in epidemiological modeling. We use different approaches. Firstly, by running simulations repeatedly and setting values of β leading to realistic spread times. Secondly, during the early stages of an epidemic, a lot of work is conducted to estimate the value of $\mathcal{R}_0$ using various methods. Using this value, the values estimated for the rates of movement and epidemiological parameters, one can estimate values of β from the expression for $\mathcal{R}_0$ deduced from the analysis of the deterministic model. Although the values of $\mathcal{R}_0$ for the deterministic and stochastic models do not usually exactly coincide, the deterministic $\mathcal{R}_0$ provides

a first approximation that is acceptable given the general uncertainty in which the model operates.

5.5 NUMERICAL SIMULATIONS

Simulations are performed using the C programming language, which allows easy implementation of parallel routines and execution in high performance computing (HPC) environments. A large number of independent simulations are performed and a number of characteristics of these simulations are computed: number of realizations where the disease becomes extinct, number of realizations where a given ACA is "hit," that is, imports an infected case, number of realizations with successful invasion, that is, where an imported case infects a local individual, etc.

Alerts can then be ranked by a given public health entity in terms of the proportion of simulations that activate it under one of the criteria above.

5.6 CONCLUSIONS

By incorporating information about how individuals travel on the global air transportation network and using initial conditions emanating from internet surveillance systems, the mathematical model will allow us to classify alerts generated anywhere in the world in terms of the risk they represent to a given public health entity.

The system is currently under development, with one aspect in particular being the object of a lot of work: the speeding up of computations. Indeed, because the time to the next event in the stochastic simulation is exponentially distributed with parameter the total weight of events, the time steps usually take an unreasonably small size. The first method used to circumvent this problem is the so-called τ-leap method (Cao et al. 2006), which allows us to consider "packets of events."

REFERENCES

Arino J., Brauer F., van den Driessche P., Watmough J., and Wu J. (2006). Simple models for containment of a pandemic. *Journal of the Royal Society Interface*, 3(8):453–457.

Arino J., Hu W., Khan K., Kossowsky D., and Sanz L. (2011). Some methodological aspects involved in the study by the Bio.Diaspora Project of the spread of infectious diseases along the global air transportation network. *Canadian Applied Mathematics Quarterly*, 19(2):125–137.

Ash P. F. and Bolker E. D. (1986). Generalized Dirichlet tessellations. *Geometriae Dedicata*, 20(2):209–243.

Cao Y., Gillespie D. T., and Petzold L. R. (2006). Efficient step size selection for the tau-leaping simulation method. *Journal of Chemical Physics*, 124(4):044109.

Gillespie D. T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361.

Kermack W. O. and McKendrick A. G. (1927). A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London, Series A*, 115:700–721

Khan K., Arino J., Calderon F., Chan A., Gardam M., Heidebrecht C., Hu W., Janes D. A., Macdonald M., Sears J., Raposo P., and Wang J. (2009). An analysis of Canada's vulnerability to emerging infectious disease threats via the global airline transportation network. Technical report, The BioDiaspora Project (St. Michael's Hospital, Toronto, Ontario, Canada).

Khan K., Arino J., Hu W., Raposo P., Sears J., Calderon F., Heidebrecht C., Macdonald M., Liauw J., Chan A., and Gardam M. (2009). Spread of a novel influenza A (H1N1) virus via global airline transportation. *New England Journal of Medicine*, 361(2):212–214.

Khan K., McNabb S. J. N., Memish Z. A., Eckhardt R., Hu W., Kossowsky D., Sears J., Arino J., Johansson A., Barbeschi M., McCloskey B., Henry B., Cetron M., and Brownstein J. S. (2012). Infectious disease surveillance and modelling across geographic frontiers and scientific specialties. *Lancet Infectious Diseases*, 12(3):222–230.

Lessler J., Reich N. G., Brookmeyer R., Perl T. M., Nelson K. E., and Cummings D. A. T. (2009). Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infectious Diseases*, 9(5):291–300.

Research and Evaluation Branch CIC. (2011). Immigration overview—permandent and temporary residents. Technical report, Citizenship and Immigration Canada.

World Health Organization [WHO]. (2013). Immunization surveillance, assessment and monitoring—estimates of national immunization coverage. Available at http://www.who.int/immunization/monitoring_surveillance/ (accessed September 30, 2013).