

## Assessing the impact of airline travel on the geographic spread of pandemic influenza

Rebecca F. Grais<sup>1</sup>, J. Hugh Ellis<sup>1</sup> & Gregory E. Glass<sup>2</sup>

<sup>1</sup>Department of Geography and Environmental Engineering; <sup>2</sup>W. Harry Feinstone Department of Molecular Microbiology and Immunology Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA

Accepted in revised form 16 May 2003

**Abstract.** The objective of this research is to explore what would happen if the Hong Kong influenza pandemic strain of 1968–1969 returned in 2000. We report the results of a series of simulations of an SEIR epidemic model coupled with air transportation data for 52 global cities. Preliminary results suggest that if the 1968–1969 pandemic strain returned, it would spread concurrently to cities in both

the northern and southern hemispheres thereby exhibiting less of the characteristic seasonal swing. In addition, after recognition of pandemic onset in the focal city, the time lag for public health intervention is very short. These findings highlight the importance of coordinated global surveillance and pandemic planning.

**Key words:** Air travel, Geographic spread, Pandemic influenza, Prevention and control

### Introduction

Since the last substantial influenza pandemic over 30 years ago, international air travel has increased significantly. During the 1968–1969 influenza pandemic, over 160 million persons traveled internationally on commercial flights. By the late 1990s, over 620 million persons travel internationally on commercial flights [1]. Increased travel volumes potentially enhance the speed and spread of a future pandemic. In addition, the global shift of populations to urban centers has increased the size and population density of cities thereby favoring the spread of infectious diseases such as influenza [2].

In 1985, Rvachev and Longini demonstrated through retrospective modeling that the 1968–1969 pandemic diffused through a network of global cities interconnected by air travel [3]. Model parameters were estimated from incidence data for Hong Kong (the first city to report the new strain of influenza (A2/Hong Kong/1968)). World Health Organization (WHO) morbidity and mortality reports were used to validate the model. Transportation data consisted of the average annual daily number of airline passengers traveling between 52 cities around the world (Table 1). The Rvachev and Longini model has subsequently been applied to forecast the spread of influenza within countries and by other models of transportation [4–6].

The goal of this research is to explore how the 1968–1969 pandemic would spread around the world if the strain returned with contemporary air travel volumes. The model does not address transmission of

influenza on the airplane proper, but rather the geographic spread of influenza cases. We updated the model of Rvachev and Longini to examine the spatial and temporal dynamics of disease spread for 1968–1969 pandemic in the year 2000. Simulation results may provide insight into how the next pandemic strain of influenza will spread around the world today.

### Materials and methods

The following is a brief overview of the model following the notation of Rvachev and Longini. (For detailed model formulation see references: [3, 7, 8]). The population ( $n_i$ ) of city  $i$  is divided into four mutually exclusive disease states: susceptible ( $x_i(t)$ ), latent ( $u_i(\tau, t)$ ), infectious ( $y_i(\tau, t)$ ), and removed ( $z_i(t)$ ) due to recovery or death. Two time indices are employed: calendar time ( $t$ ) and a shifted time index ( $\tau$ ) used to describe the progress of infection within individuals once they have been infected.

Contact ( $\lambda$ ) between susceptible and infectious individuals sufficient for infection determines whether an individual may become infected. Newly infected individuals are calculated by the standard mass action formulation as the product of the number of susceptibles, number of infectious persons and the contact rate at time  $t$  in city  $i$ . Once infected, individuals remain latent for an uncertain upper-bounded period ( $f(\tau_1)$ ) and progress to the infectious state for uncertain upper-bounded period ( $g(\tau_2)$ ). The probabilities of transition from the latent to infectious state

**Table 1.** 1968 and 2000 Global city populations and climate zones

City	Zone	Population ( $10^3$ )		City	Zone	Population ( $10^3$ )	
		1968	2000			1968	2000
London	N	7379	11100	Kinshasa	E	1404	3000
Paris	N	8197	9775	Johannesburg	S	1433	3650
Rome	N	2800	3175	Casablanca	N	1506	2475
Berlin	N	3249	5061	Mexico City	E	3026	14100
Madrid	N	2936	4650	Bogota	E	2818	4260
Warsaw	N	1356	2323	Havana	S	1755	2125
Budapest	N	2039	2565	Caracas	E	1035	3600
Sofia	N	910	1205	Lima	E	2541	4344
Stockholm	N	973	1450	Santiago	S	2556	4100
Hong Kong	E	3900	5396	Buenos Aires	S	2972	10750
Tokyo	N	11410	23620	Rio de Janeiro	E	4316	10150
Peking	N	7570	10500	Sao Paulo	E	5187	15175
Shanghai	N	7000	9300	Honolulu	E	628	762
Singapore	E	2017	3025	Sydney	S	2780	3365
Manila	E	1582	5474	Melbourne	S	2425	2833
Bangkok	E	2027	6450	Perth	S	725	995
Jakarta	E	4915	8600	Wellington	S	136	350
Calcutta	E	3141	11100	Montreal	N	1214	2921
Bombay	E	5970	9950	New York	N	11572	16472
Delhi	N	3647	7200	Los Angeles	N	7000	9764
Madras	E	2470	4475	Washington	N	2836	3221
Seoul	N	5536	15850	Houston	N	1231	2755
Teheran	N	3400	6400	Chicago	N	7800	7717
Karachi	N	3469	5300	San Francisco	N	3700	4054
Cairo	N	5384	9300	Atlanta	N	1258	1962
Lagos	E	901	3800	Capetown	S	691	1790

( $\gamma(\tau)$ ) and from the infectious to the removed state ( $\delta(\tau)$ ) are computed from the probabilities that prescribe the length of time an individual spends in the latent and infectious states.

#### Travel between cities

Global cities are directly and/or indirectly connected through a symmetric ( $n \times n$ ) air travel matrix. The elements of the matrix ( $\sigma_{ij}$ ) are defined as the daily passenger flow from city  $i$  to  $j$  (i.e., the average number of individuals that travel from city  $i$  to city  $j$  in 24 hours). The probability of travel is calculated by dividing the average daily number of travelers from city  $i$  to  $j$  divided by the population of city  $i$  ( $\sigma_{ij}/n_i$ ). Similarly, the probability of travel from city  $j$  to city  $i$  is calculated by dividing the average daily number of travelers from city  $j$  to  $i$  divided by the population of city  $j$ .

We assume that susceptible and well latent individuals travel and infectious individuals do not. Susceptible and latent individuals are assumed to travel in proportion to their representation in each city at time  $t$ . A transportation operator ( $\Omega_i$ ) is applied to the susceptible and latent state equations (Equations 2 and 3) to account for travel between cities. Epidemics within each city can occur and in-

dividuals traveling though the transportation network can create inter-city epidemics.

#### Computational algorithm

The model consists of a separate but identical set of difference equations defining the disease states for each city. The initial number of susceptible individuals is assumed to be a fraction ( $\alpha$ ) of the city population. The initial number of latent individuals is pre-specified for the initial city of release and is zero for all other cities. There are assumed to be no initial infectious individuals in any city. The daily incidence ( $w_i(t)$ ) is calculated by stepping through the following equations for all cities. Because only a fraction of influenza cases are reported, forecast daily incidence is multiplied by a reporting rate ( $\rho$ ).

$$u_i(0, t) = x_i(t) \frac{\lambda}{n_i} \sum_{\tau=\tau_1}^{\tau_2} y(\tau, t) \quad (1)$$

$$\begin{aligned} \Omega_i[x_i(t)] &= x_i(t) \\ &+ \sum_{j=1}^N \left[ x_j(t) \frac{\sigma_{ji}}{n_j} - x_i(t) \frac{\sigma_{ij}}{n_i} \right], \quad i = 1, 2, \dots, n \end{aligned} \quad (2)$$

$$\begin{aligned} \Omega_i[u_i(\tau, t)] \\ = u_i(\tau, t) + \sum_{j=1}^N \left[ u_j(\tau, t) \frac{\sigma_{ji}}{n_j} - u_i(\tau, t) \frac{\sigma_{ij}}{n_i} \right], \\ i = 1, 2, \dots, n; \quad \tau = 1, 2, \dots, \tau_1 - 1 \end{aligned} \quad (3)$$

$$x_i(t+1) = \Omega_i[x_i(t)] - u_i(0, t) \quad (4)$$

$$u_i(\tau+1, t+1) = [1 - \gamma(\tau)]\Omega_i[u_i(\tau, t)], \quad \tau = 0, 1, \dots, \tau_1 - 1 \quad (5)$$

$$\begin{aligned} y_i(\tau+1, t+1) \\ = \begin{cases} \gamma(\tau)\Omega_i[u_i(\tau, t)] + [1 - \delta(\tau)]y_i(\tau, t), & \tau = 0, 1, \dots, \tau_1 \\ [1 - \delta(\tau)]y_i(\tau, t), & \tau = \tau_1 + 1, \tau_1 + 2, \dots, \tau_2 - 1 \end{cases} \end{aligned} \quad (6)$$

$$w_i(t+1) = \rho \sum_{\tau=0}^{\tau_1} \gamma(\tau)\Omega_i[u_i(\tau, t)] \quad (7)$$

### Seasonality

Like many respiratory diseases, influenza follows a seasonal pattern exhibiting a low summer and high winter incidence [9]. This tendency to peak in colder months leads to a characteristic oscillation of influenza activity between the northern and southern hemispheres at approximately 6-month intervals. In the equatorial region, the disease occurs year round. To incorporate seasonality, the contact parameter is scaled to mimic seasonal changes.

Past research has assumed either a constant contact parameter ( $\lambda$ ) or used two values of the parameter to differentiate between the influenza season and non-influenza season [3]. This research refines the seasonal scaling to a monthly level. Global cities are divided into three zones: northern hemisphere (N), southern hemisphere (S) and equatorial (E) (Table 1). A monthly scaling factor is applied to the contact parameter in northern and southern hemispheric cities to differentiate the influenza season from the non-influenza season (Table 2). No scaling factor is applied to equatorial cities. The factors were determined by linear interpolation assuming the contact parameter is 10% of its full value during the non-influenza season. A reduction of 90% was chosen based its use in past research. A linear relationship was chosen for ease of computation.

### Data

We used the same 52 global cities and 1968–1969 transportation matrix as Rvachev and Longini to allow for comparison between the simulated 1968–1969 and 2000–2001 pandemics. Global city populations for 2000 were taken from the US Department of State's International Population Database. A com-

**Table 2.** Monthly seasonal scaling factors applied to the contact parameter ( $\lambda$ )

Month	Southern hemisphere	Northern hemisphere
January	0.10	1.0
February	0.25	0.85
March	0.55	0.70
April	0.70	0.55
May	0.85	0.25
June	1.0	0.10
July	1.0	0.10
August	0.85	0.25
September	0.70	0.55
October	0.55	0.70
November	0.25	0.85
December	0.10	1.0

plete list of cities, their populations in 1968 and 2000, and climate zone is provided in Table 1. The 2000–2001 transportation matrix was compiled from several sources. We obtained permission from the US Department of Transportation to use the Origin-Destination database, T100 and T100(f) databases. These databases provide average annual daily passenger counts for travel between foreign ports to the US. Passenger flows between foreign ports were obtained from the Official Airline Guide, the International Civil Aviation Organization, International Air Transport Association, Back Aviation Solutions, Inc. and air travel volumes in the published literature [1, 3, 10–12]. A complete list of passenger statistics for 2000 is presented in the Appendix A. See reference [3] for 1968 passenger counts.

To reflect the seasonal fluctuation in air travel (higher volume in the summer months), we applied a scaling factor to the transportation matrix. The scaling factors were estimated from published research on passenger air travel trends [1]. The full value was used for the first quarter, 0.90 for the second quarter, 1.2 for the third quarter and 1.1 for the fourth quarter. Ideally, measured quarterly matrices would be employed. A key limitation to the passenger statistics is that only direct flights are considered. Some flights are likely to be segments in a longer journey and as a result, the number of passengers traveling between certain city pairs may be inflated. In other cases, the number of passengers traveling may be underestimated because direct service between two cities is not available. Further, because a symmetrical matrix is used, individuals essentially have the same probability of leaving a city as they do returning to their origin. Because average international trip lengths are longer than the latency period for influenza, this simplifying assumption is of less concern. The lack of specific passenger travel statistics remains a serious obstacle to this research.

The initial values for all disease parameters are those used by Rvachev and Longini [3]. The authors

used weekly reported incidence of influenza in Hong Kong between June 13, 1968 and July 27, 1968 (ascending limb of the epidemic curve) to estimate parameter values. They estimated the contact parameter to be 1.055 and the susceptible fraction to be 0.61 using a least squares estimation procedure ( $R_{1-\alpha} = 1.9$ ). Their parameter estimation procedure is well documented in the literature [3, 7]. We used the same state probabilities and reporting rate (30%). The pandemic was assumed to originate in Hong Kong on June 13, 2000 with a time horizon of 1 year.

Model validation was conducted by comparing results of simulations employing the same parameter values and air transportation statistics as Rvachev and Longini. We successfully completely reproduced their published results for temporal and spatial spread of the 1968–1969 influenza pandemic.

## Results

### *Spatial dynamics*

In 1968 influenza cases spread around the world beginning with cities closest to Hong Kong, followed by cities in northern latitudes and finally to cities in southern latitudes. Substituting 2000 travel patterns for 1968 travel patterns forecast a somewhat different pattern. The disease spread exhibits the characteristic seasonal hemispheric swing, but only in cities where

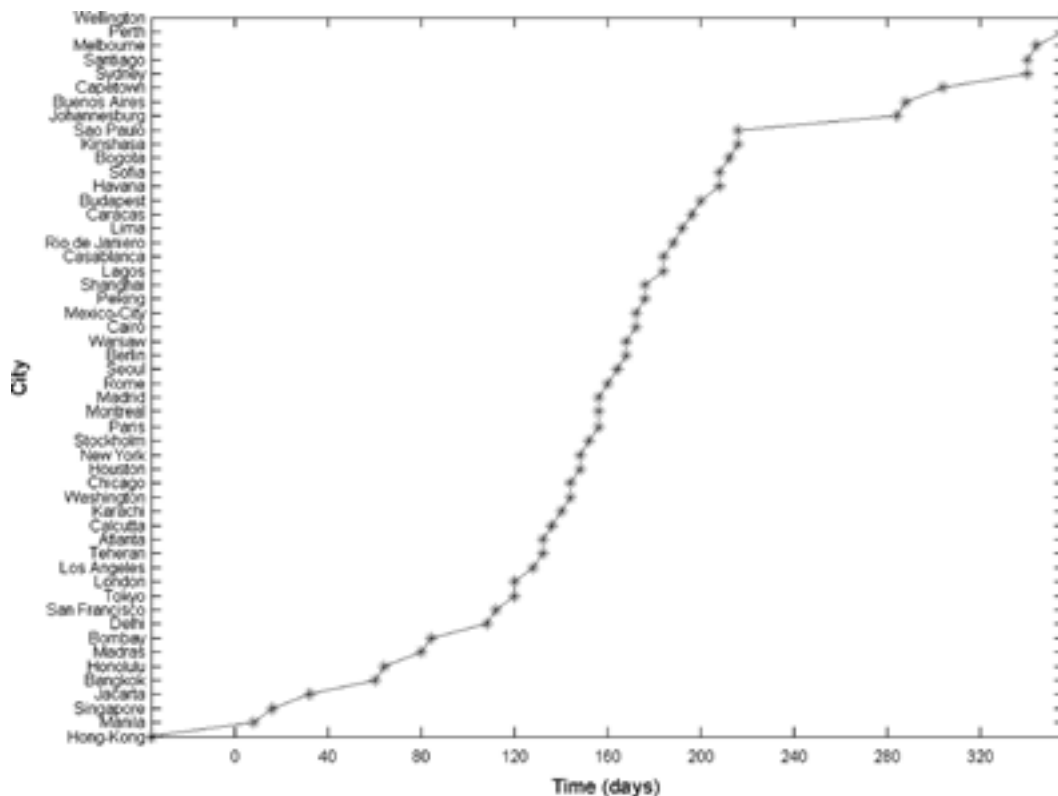
there is relatively less air travel. As may be expected, the disease progresses to cities closest to Hong Kong first, but unlike the 1968 pandemic, also to cities with significant air travel volumes. For example, using 1968 travel volumes the model forecast Sydney, Australia as the 48th city to report cases. In the 2000 model, Sydney is forecast to be among the first cities to report cases along with Singapore, Johannesburg, Melbourne, Perth and Wellington. We would expect the pandemic to reach southern hemispheric cities such as Melbourne, Perth and Wellington after the pandemic spread through the northern and equatorial latitudes. Figure 1 gives the temporal sequence of epidemic peaks with 1968 travel volumes and Figure 2 with 2000 travel volumes.

### *Temporal dynamics*

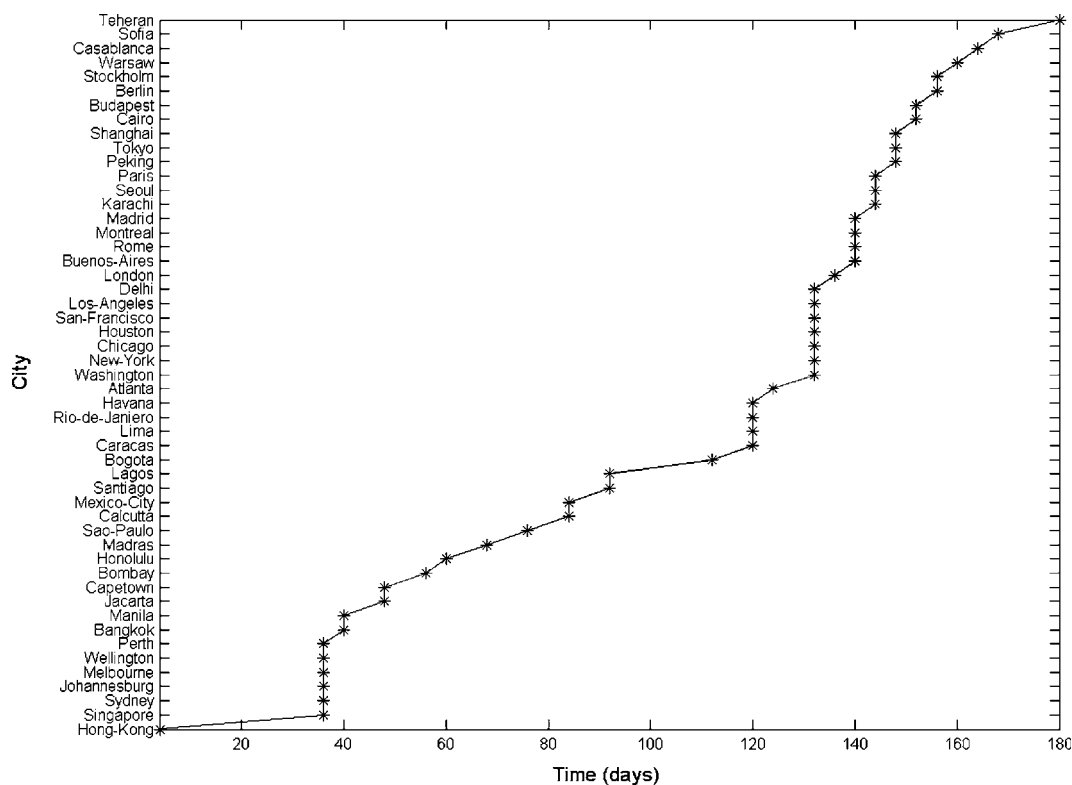
When 2000 travel volumes are used in the analyses, the pandemic reaches northern hemispheric cities an average of 111 days earlier than it was forecast in 1968. Figure 3 shows the course of the 1968–1969 pandemic over the time horizon. Figure 4 shows the time compression evident in the 2000 forecast where all cities report cases significantly earlier.

### *Impact*

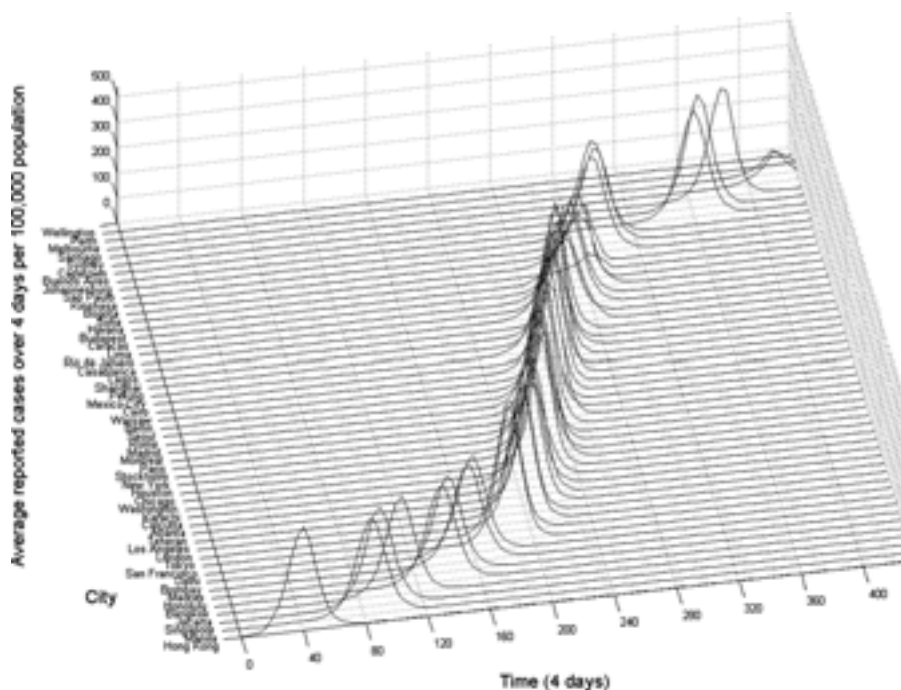
The forecast number of reported cases for the 1968 pandemic strain simulated with 2000 travel patterns is



**Figure 1.** Temporal progression of forecast 1968–1969 pandemic (depicts the temporal progression of the 1968–1969 pandemic ordered from earliest to latest epidemic peak day).



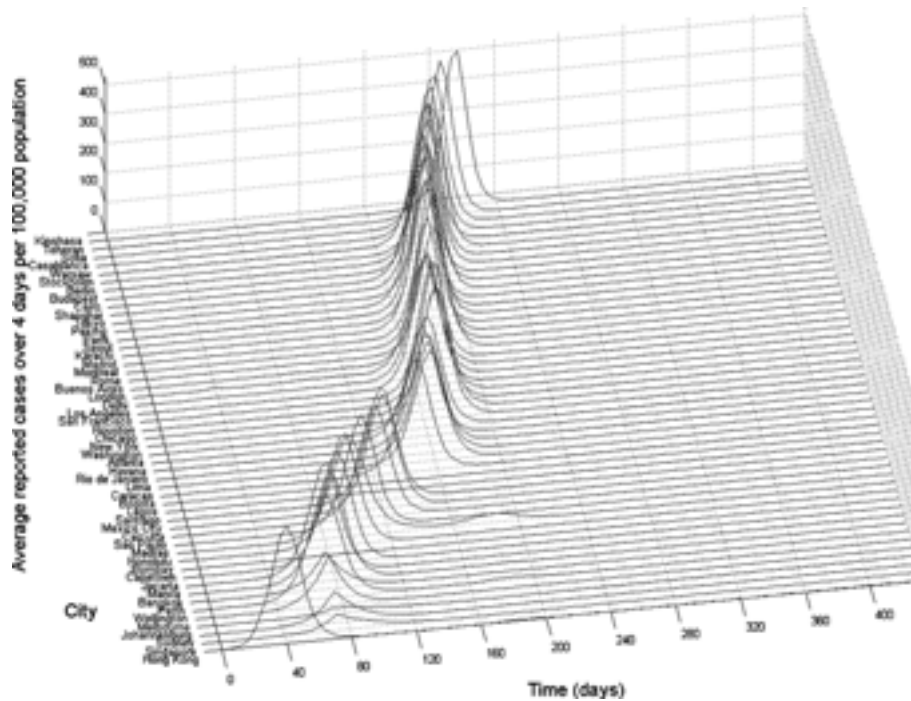
**Figure 2.** Temporal progression of forecast 2000–2001 pandemic (depicts the temporal progression of the forecast 1968–1969 pandemic if it occurred with 2000 air travel volumes. Cities are ordered from earliest to latest epidemic peak day).



**Figure 3.** Forecast average 4-day incidence per 100,000 persons for the 1968–1969 pandemic for 52 global cities ordered from earliest to latest epidemic peak.

markedly greater than that forecast with 1968 travel volumes. The pandemic peak with 2000 travel patterns was forecast to be an average of 176% greater than the 1968 pandemic. Cumulative reported cases were forecast to be an average of 188% greater.

Equatorial cities were forecast to have the greatest increase in both peak and overall magnitude. Southern hemispheric cities were forecast to have a smaller epidemic peak and total epidemic magnitude (Table 3).



**Figure 4.** Forecast average 4-day incidence per 100,000 persons for the 1968–1969 pandemic with 2000 air travel volumes for 52 global cities ordered from earliest to latest epidemic peak.

**Table 3.** Ratio of forecast peak and cumulative cases of the 2000–1968 forecast pandemics

City hemisphere	Peak cases (2000/1968)	Cumulative cases (2000/1968)
Northern	1.7	1.8
Equatorial	2.5	2.7
Southern	0.37	0.54
Average	1.8	1.9

## Discussion

Simulations results suggest two important potential changes in the spatial and temporal spread of a future pandemic strain of influenza. The virus may spread to cities in both the northern and southern hemispheres concurrently thereby exhibiting less of a hemispheric swing and the time lag for intervention action is very short. These findings highlight the importance of coordinated global surveillance and pandemic planning.

Although these analyses are suggestive of potential changes in spatial and temporal spread of pandemic influenza, the results are preliminary. There are multiple simplifying assumptions and limitations to model. Further, there are significant limitations to the air travel data used in this analysis. The results should be interpreted less as an endpoint and more as a starting point for discussion concerning potential differences in the spatial and temporal spread of the next influenza pandemic. Models, such as the one

employed here, highlight areas where future research is needed and unanswered questions about the epidemiology of influenza.

Several key limitations are important to mention. First, the spread of influenza around the world is likely due only in part to global air travel. The virus may spread from continent to continent by other modes of transportation or the virus may ‘over winter’ in a location [9, 13]. It is also likely that during a pandemic, multiple strains would circulate within a population.

Second, we use unique parameter values. Neither national populations nor population sub-groups are uniformly susceptible. Susceptibility, course of infection and transmission of influenza from an infective to a susceptible individual varies by geography, age, economic, social and cultural factors, access to health care and presence of additional potential risk factors. On a global scale, however, homogenous mixing has been shown to be appropriate at the city level and to describe the initial spatial spread of epidemics [14]. Population heterogeneities play an important role in forecasting epidemics on a smaller spatial scale. Modeling the spread of influenza within a city would necessitate inclusion of specific population characteristics.

## Acknowledgement

We wish to thank Dr Ira Longini for his comments and suggestions.



**References**

1. Hanlon J, Hanlon P. Global Airlines: Competition in a Transnational Industry. London: Butterworth-Heinemann, 1999.
2. Wilson M. Travel and the emergence of infectious diseases. *Emerg Inf Dis* 1995; 1(2): 39–46.
3. Rvachev L, Longini I. A mathematical model for the global spread of influenza. *Math Biosci* 1985; 75: 3–22.
4. Flahault A, Deguen G, et al. A mathematical model for the European spread of influenza. *Eur J Epidemiol* 1994; 10: 471–474.
5. Flahault A, Letrait S, et al. Modeling the 1985 influenza epidemic in France. *Stat Med* 1988; 7: 1147–1155.
6. Aguirre A, Gonzalez E. The feasibility of forecasting influenza epidemics in Cuba. *Mem Inst Osw Cruz* 1992; 87(3): 429–432.
7. Longini I. A Mathematical model for predicting the geographic spread of new infectious agents. *Math Biosci* 1988; 90: 367–383.
8. Longini I, Fine P, Thacker S. Predicting the global spread of new infectious agents. *Am J Epi* 1986; 123: 383–391.
9. Thacker S. The persistence of influenza A in human populations. *Am J Epi* 1986; 8: 129–142.
10. RBI. Official Airline Guide. London: Reid Elsevier, 2000.
11. Airbus Industries. The Airbus Global Market Forecast. Cedex, France: Airbus Industries, 2000.
12. Meskill T, Griffiths J. Current Market Outlook 2000: Into the New Century. Seattle, Washington: Boeing Corporation, 2000.
13. Dowell SF. Seasonal variation in host susceptibility and cycles of certain infectious diseases. *Emerg Inf Dis* 2001; 7(3): 369–374.
14. Bonabeau E, Toubiana L, Flahault A. The geographical spread of influenza. *Proc R Soc Lond B* 1998; 265(1413): 2421–2425.

*Address for correspondence:* R.F. Grais, WHO Coordinating Center for Electronic Disease Surveillance, INSERM Unit 444, 27, rue Chaligny, 75571 Paris cedex 12, France  
Phone: +33-1-44-73-8437; Fax: +33-1-44-73-8454  
E-mail: rfreeman\_grais@yahoo.com