# Strategies for mitigating an influenza pandemic.

Neil M. Ferguson[1], Derek A.T. Cummings[2], Christophe Fraser[1], James C. Cajka[3], Philip C. Cooley[3], Donald S. Burke[2]

[1] Department of Infectious Disease Epidemiology, Faculty of Medicine, Imperial College London, St Mary's Campus, Norfolk Place, London W2 1PG, UK.

[2] Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, Maryland 21205, USA.

[3] RTI International Inc., P.O. 12194, 3040 Cornwallis Rd., Research Triangle Park, North Carolina 27709, USA

**Keywords**: influenza, pandemic, H5N1, mathematical model, antivirals, neuraminidase inhibitor

## SUPPLEMENTARY INFORMATION

# *Contents*

**Contents** ..... **2**

**1. Population data** ..... **3**
Population density data ..... 3
Household size and age structure data ..... 3
School size data and school allocation model ..... 4
Workplace data and allocation model ..... 5
Commuting distance data ..... 5
Air travel data ..... 6

**2. Model details** ..... **7**
Details of the transmission model ..... 7
US air travel model ..... 8
Model of infection seeding ..... 10
Interventions ..... 11
Numbers of realisations & computational resources ..... 13

**3. Natural history and transmission parameters** ..... **13**
Estimation of $R_0$ from inter-pandemic influenza household data ..... 14
Estimation of $R_0$ for influenza from past pandemics ..... 15
Assumptions about population behaviour during a pandemic ..... 18
Transmissibility scenarios ..... 19

**4. Impact of combination policies** ..... **20**

**5. Sensitivity analyses** ..... **23**
'H5N1-like' infection characteristics ..... 23
Within-place group structure and targeting ..... 24
School closure policies ..... 24
Natural history parameters ..... 26
Proportion of school/workplace transmission ..... 28
Proportion of infections which become clinical cases ..... 28
Behavior of symptomatic individuals ..... 30
Timing of start of mitigation policies ..... 31
Effect of kernel shape on US pandemic dynamics ..... 32
'Containment' policies applied in GB or the US ..... 33
Internal travel restrictions ..... 34

**Bibliography** ..... **36**

# *1. Population data*

### Population density data

The 2003 Landscan[1] dataset prepared by Oakridge National Laboratory was used as the source of population density information for the continental US and GB. This dataset has a 30 arc second resolution, equating to approximately <1km for US and GB latitudes. The dataset is a model of population density which is constructed from census, remote sensing, land use and transport network data. It is a model of instantaneous population density (i.e. where people are at an instant of time) rather than residential population density, but is very comparable to high resolution census data for GB and US, and has the advantage of being rasterized – while census datasets are defined for populations within with irregularly shaped and variable size administrative units (e.g. census tracts).

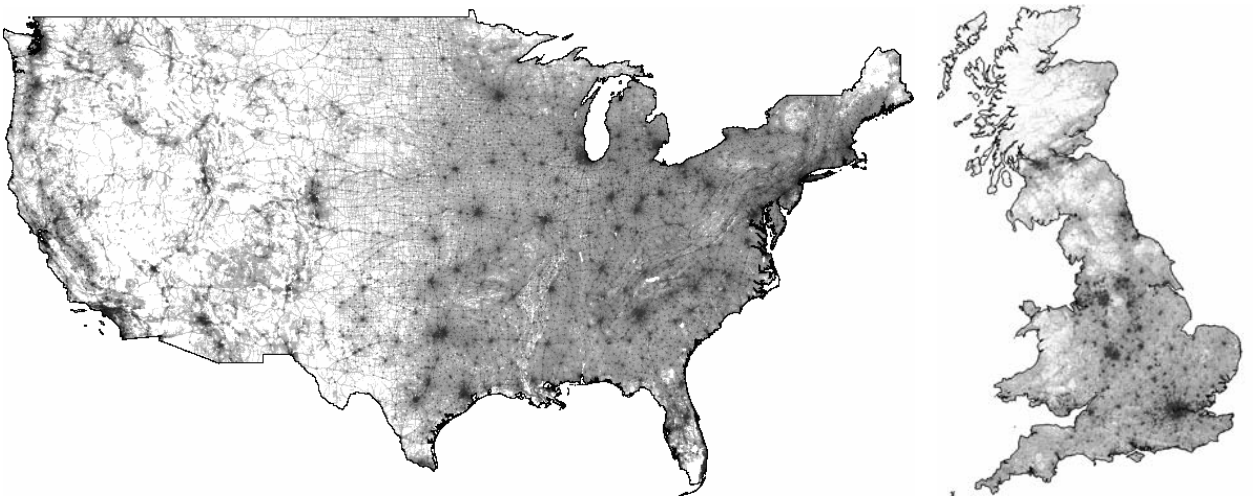

**Figure SI1**: Landscan 2003 [1] population density dataset, with population density plotted on a logarithmic scale (lighter areas have lower density), for USA (left), and GB (right).

For the USA, we modelled transmission the continental US (excl. Hawaii and Alaska), and (allowing for the 3 million per year growth rate) assumed a total population size of 300 million [2]. For the UK, we modelled Great Britain alone, and assumed a population size of 58.1 million [3].

### Household size and age structure data

Census data was used for age[2,3] and household size distribution data[4,5] for both the US and UK. We used a heuristic model[6] to generate the ages of individuals in households which maintained generational age gaps with households while matching marginal household size and age distributions (Figure SI2). We assumed household size and age distributions did not vary geographically.

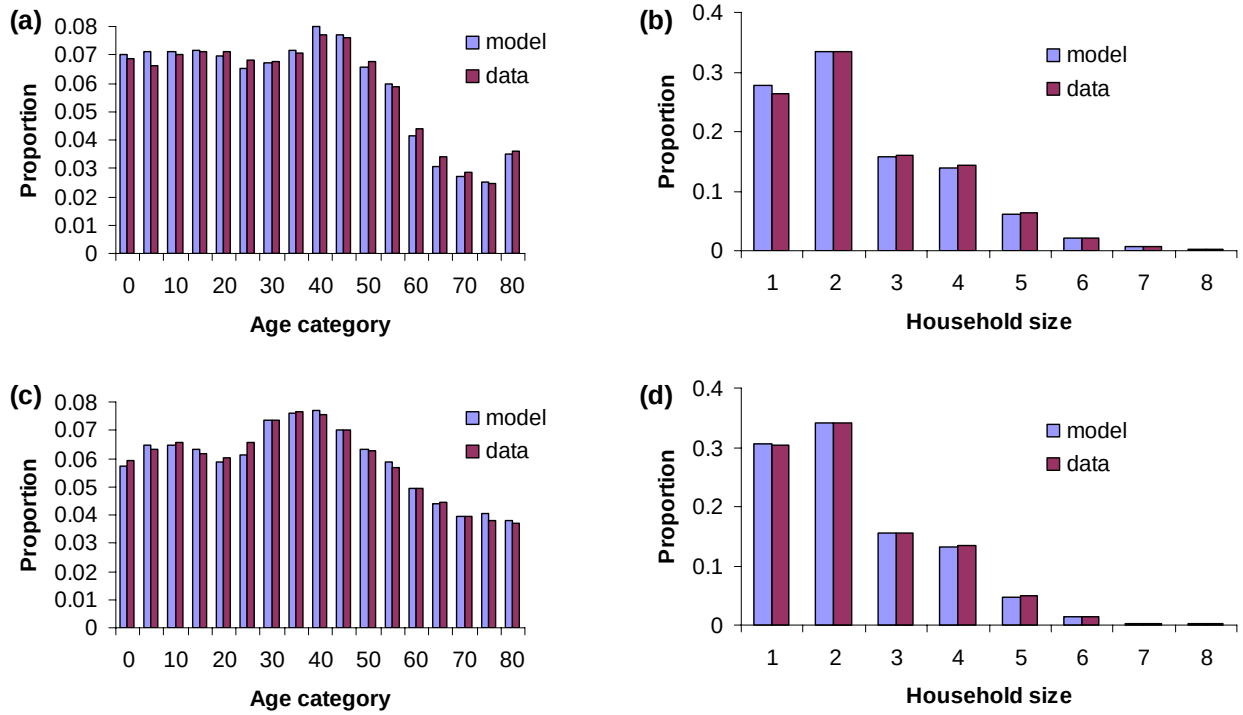


**Figure SI2:** (a) Comparison of model generated population age distribution with data for US (see text for sources). (b) as (a) but for household size distribution. (c) as (a) but for UK. (d) as (b) but for UK.

## School size data and school allocation model

For GB, we used data[7] on the average size of primary (218 pupils in 2004) and secondary schools (949 pupils in 2004) and on average class sizes and staff-student ratios and applied the methods used in past work[6] to generate a synthetic population of schools distributed in space with a density proportional to local population density. These methods allocate children to schools using a free selection algorithm; each child 'picks' a school at random from the nearest 3 (for primary) or 6 (for secondary) schools of the appropriate type, with a probability weighted by the distance kernel derived from UK census data on commuting behaviour (see below). Restricting school choice to the nearest 3 or 6 schools avoids unrealistic long tails to the distribution of distances travelled to school, albeit at the cost of slightly underestimating mean distances travelled to school by approximately 30% [8]. Staff student ratios were used to determine the proportion of adults to schools rather than other workplaces.

For the US, we used a database of all US schools[9], geocoded the school addresses, and used the resulting georeferenced dataset to populate the model directly. The dataset included information on numbers of students by school year, which was used to specify an age-specific student capacity for each school. In this case, a free-selection algorithm could not be used to assign pupils to schools, as schools had predetermined numbers of pupils of different ages. Instead, we used the worker to workplace algorithm [6], but with a very local selection kernel of form $f(d) \sim 1/[1+d/a]^b$, where $a$=14 km and $b$=11. This gave an average distance from household to school of 7.7km (4.8 miles), in agreement with detailed survey data [10] and broadly in line with other less specific data on average trip distance travelled from home to school or church [11].

## Workplace data and allocation model

US workplaces obey a power law distribution for the number of employees[12]. However, as the model incorporates workplaces rather than firms, we used establishment size data [13]. Data on numbers of workplaces by size is provided in size groupings: we fitted this categorical distribution to a discrete power-law distribution using maximum likelihood methods by integrating over each size category. The resulting workplace size distribution has Zipf-like form (namely an offset truncated power distribution)

$P(n>m)=\left(\left[(1+m_{\max}/a)/(1+m/a)\right]^{c}-1\right)\Big/\left[(1+m_{\max}/a)^{c}-1\right]$ for $m \le m_{\max}$ with parameter estimates $a$= 5.36, $c$=1.34 and $m_{\max}$=5920 (see figure SI3). Independent data were not available for UK establishments, so US parameters were assumed.

The algorithm to allocate individuals to workplaces used was as detailed in previous work[6].

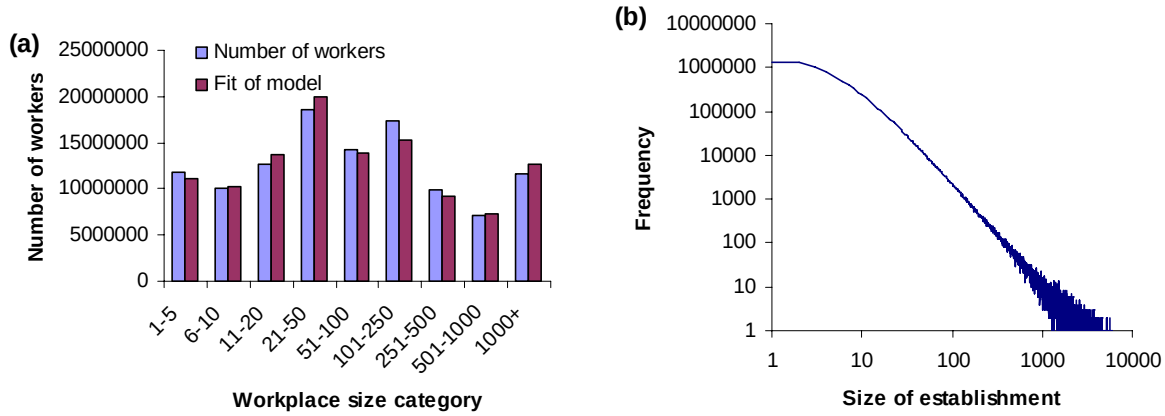


**Figure SI3:** (a) Data on numbers of workers in US working in workplaces of different size categories (see text for source), compared with maximum likelihood fit of offset power model for workplace size distribution. (b) Best fit modelled distribution of workplace sizes.

## Commuting distance data

For GB, data on origin-destination flows for travel-to-work derived from 1991 census were used[3]. The resolution of this data was ward level, there being approximately 10,000 wards in GB. Data from the 2001 census was unusable due to the method of disclosure control adopted by the UK Office for National Statistics, which in essence removed all useful information on flows beyond 50km.

For the US, comparable data is obtainable from the STP64 dataset[14], which has a census tract resolution, there being some 65,000 census tracts in the US.

In both cases, data on the centroid locations of the administrative units used (wards or tracts) was used to construct distributions of the distance travelled to work, as shown in Figure SI4. These distributions were not found to vary significantly between regions in the GB or states in the US.

For both datasets, no distance information is available for individuals who live and work in the same tract or ward, other than the approximate size of the tract/ward. This biases the estimates of the distribution of commuting distances for short (<~15km) distances. We therefore redistributed those individuals (~10% for the US data ~3% for GB, due to the smaller size of wards) who report living and working in the same census tract/ward according to the estimated population-weighted distribution of tract radii. However, short distance estimates shown in Figure SI4 are

still likely to be subject to some biases/error – for instance, due to the fact that only distances between tract/ward centroids are known, not precise journey distances.

The parameters and functional form of the choice kernel used in the worker-to-workplace assignment algorithm were then adjusted until the model fit the empirical distributions well. In the case of GB a 2 parameter offset-power law kernel fit well: $f(d) \sim 1/[1+(d/a)^b]$, where $a$=4 km and $b$=3. For the US, the change in slope of the tail of the distribution (on a log scale) meant that a 4 parameter function was needed: $f(d) \sim 1/[1+d/a]^b + k/[1+d/a]^c$ where $a$=35km, $b$=6.5, $k$=0.0004 and $c$=2.2. Figure SI4 illustrates how well the model fitted the data.

The extent to which the GB and US distributions differ is perhaps surprising – the GB distribution is asymptotically close to power law (in common with many other travel distance datasets we have collated), but the US dataset shows a much larger modal distance (other than 0, which is the most common distance in each dataset, as the datasets include home workers) of 5 (rather than 1km), but then a steeper decline between 20 and 100km.

One important distinction of the datasets is the census question form used: for GB, people were asked to give the address of their workplace, while in the US, people were asked to identify the place they spent most time working at in the previous week. The former of this latter question means that the US dataset contains data on occasional work-related trips as well as on journeys to individuals' typical workplace. As work trips lasting most of a week can be expected to be involve longer distances than typical travel to work, it is tempting to attribute the 'kink' in the tail of the US distance distribution to such occasional work-related travel. While such a conclusion cannot be robustly justified (since, by comparison, there are significant number of 200+km commutes in GB dataset), Figure SI4c shows how well the distance distribution generated by a 2 parameter kernel $f(d) \sim 1/[1+d/a]^b$ (where $a$=35km, $b$=6.5) fits the US data up to about 200km. We present some sensitivity analyses in Section 4 which examine the impact of assuming the model distribution shown Figure SI4c rather than Figure SI4b.

Some additional support for the suggestion that the US STP64 data may not well represent true commuting or other travel behaviour also comes from recent work examining the movement of paper money in the US[15], where pure power law behaviour was seen in the distribution of distances notes moved over fixed time intervals. However, other recent work[16] indicates commuting data is predictive of patterns of seasonal influenza spread, and is well represented by a gravity model.

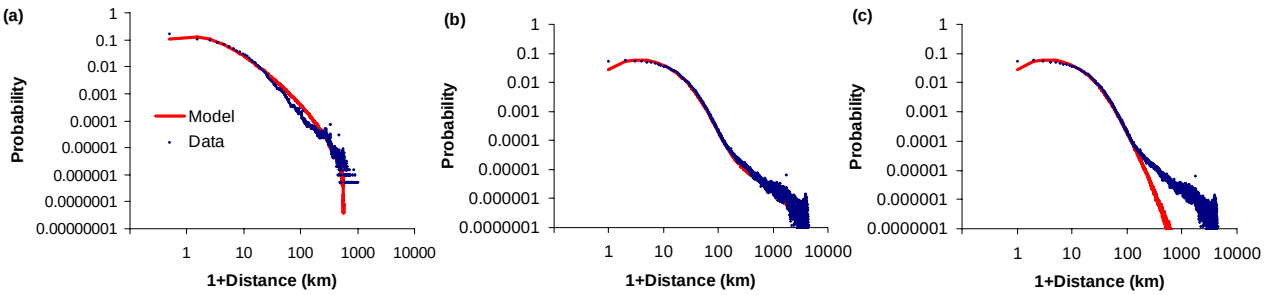


**Figure SI4:** (a) Comparison of model generated distance-to-work distribution with data for GB (see text for sources). (b) as (a) but for US. (c) as (b) but with fit to short-distance behaviour only.

## Air travel data

We used the Origin-Destination Survey of Airline Passenger Traffic dataset[17] collected by the Bureau of Transportation Statistics of the US Department of Transportation to parameterize air-travel in our model (Reference below). This survey is a 10% sample of itineraries reported by

U.S. air carriers that include at least one scheduled service with large aircraft. Air traffic on these carriers account for 90% of domestic air-travel. These data include the origin and destination of each leg or flight included in an itinerary. Itineraries consist of at least one leg, but vary in the number of legs included. These data are aggregated by quarter of travel. A major deficiency of this dataset is that no temporal data is included in the itineraries. This makes it impossible to identify the true destination of each itinerary (*i.e.* the location people intended to reach, not flight connection locations) or how much time passengers spend at each destination in their journey. Itineraries with an even number of legs beginning and starting at the same location are intuitively easier to interpret.

We assumed that from an even number of legs, the true destination was the destination airport of leg $n/2$ where $n$ is the total number of flights taken. The probability of flight from a particular origin arriving at a particular destination is estimated as the proportion of these even numbered flights from the origin arriving at each destination. Round trip flights accounted for 88% of the itineraries in this dataset. Of these, 80% consist of an even number of legs. We compared the probability distribution of destinations conditional on origins calculated using just even numbered flights and several other inclusion criteria and found no systematic differences in these distributions. Data from all quarters of 2004 were used. Overall, there were an average of 438,000 outbound journeys per day from round-trip itineraries with even numbers of legs. This was scaled by 1/0.88/0.8=1.42 to correct for the censoring of the dataset to round-trip itineraries with even numbers of legs. The even legged, roundtrip data includes 561 airports as either an origin, destination or both. We did not specifically estimate or address seasonality in air travel for our model.

As part of the air travel model, we also needed data on nights spent away for long distance air and car travel [18]. These data (derived from the 2001 National Household Travel Survey) are plotted in Figure SI5.

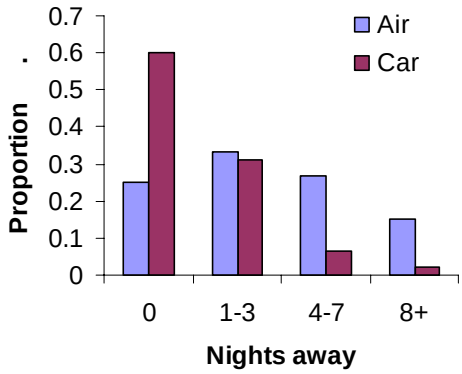


**Figure SI5:** Reported distribution of nights spent away from home for long distance journeys by air and person vehicle (car). See text for source.

## *2. Model details*

### Details of the transmission model

The model is a stochastic, spatially structured individual-based simulation, details of which have been given in past work [6]. In any time-step of $\Delta T$=0.25 days, a susceptible individual $i$ has probability $1-\exp(-\lambda_i \Delta T)$ of being infected, where $\lambda_i$ is the instantaneous infection risk for individual $i$. (an expression for $\lambda_i$ is given in the Supplementary Information for [6]). Infection

risk comes from 3 sources: (1) household, (2) place, and (3) random contacts in the community. When varying $R_0$, the transmission coefficients for these 3 components were varied by the same factor.

Community transmission depends explicitly on distance, as it is intended to represent random contacts associated with movements and travel. The probability that individual *i* infects individual *k* is weighted by a kernel function $f\left(d_{i,k}\right)$, where $d_{i,k}$ is the distance between individuals *i* and *k*. This represents a form of gravity model[19,20]. We assume the same kernel for community transmission as is used for assigning workers to workplaces (see section on Travel Data above).

For the GB model, Euclidean coordinates and distances between points were used in evaluating kernel functions, while polar (longitude/latitude) coordinates great circle distances were used for the US model.

Compared with past work[6], the GB modelling presented here used the same model, but with 2 (primary and secondary) types of schools being modelled rather than the 3 which had to be incorporated in the Thai model. In the current work we also refined the modelling of within-place (i.e. school, workplace) transmission to incorporate a simple group structure. Individuals were randomly assigned to school classes (of a size determined from data) or workgroups (assumed to have an average size of 10). Workgroups can be viewed as the set of work colleagues an individual mostly closely interacts with. As a baseline, we assumed 75% of an individual's contacts were within-group, and 25% with individuals picked randomly from the entire population of the school or workplace. This additional structure does not affect baseline epidemic dynamics or the effectiveness of control measures, except for those which target the school class or workgroup specifically. See sensitivity analyses below.

In modelling the US, we used a single place type for schools, but schools had age-specific capacities determined by the school-level dataset used (see above).

We did not model disease-related mortality here as we were principally interested in policies aimed at containing spread rather than minimising mortality/morbidity – though of course, successful containment can be expected to dramatically lower both the national and global health burden any emergent pandemic. We quantify the effectiveness of interventions by their effect on clinical attack rates. It should be noted that which social distance measures (e.g. school closure) can only affect attack rates by reducing transmission, while antivirals and vaccine can have a direct effect (in reducing the chance that an infected person will be ill enough to be a clinical case) as well as an indirect effect (on transmission).

**US air travel model**

For the US model, we also added an explicit simulation of air travel within the country. Airports were modelled explicitly, and airport to airport passenger flows simulated to match available data (see above). Incorporating explicit long-distance travel in the model was achieved by defining a 3rd place type, 'hotels'. Given the 3 million hotel rooms occupied per night in the US [21], 30,000 hotels of average occupancy 100 (and peak occupancy of twice that) were distributed according to population density across the US. Passengers for air travel were selected randomly from the population, with the origin airport being selected at random, with probability of selection being proportional to the product of airport daily passenger volume and the workplace selection kernel used to assign workers to workplaces (see above). Destination airport was chosen at random to match the observed origin-destination flows for the origin airport. Destination hotel was picked at random using a kernel depending on distance between destination airport and hotel as $f(d) \sim 1/[1+d/a]^b$, where $a$=70 km and $b$=5. No data were found to justify this arbitrary parameter

choice, however the resulting kernel is relatively flat out to 50km, and is intended to reflect the relatively long distances many people travel from their destination airport and final destination. If the first hotel picked is at capacity, then additional hotels are selected until one is found with space.

To represent mixing of air travellers with people local to an area, it was assumed that 50% of the occupants of hotels were local. Individuals were picked at random from the local community, with a distance weighting given by the workplace selection kernel.

Residency times in hotels for air travellers and local (assumed car) travellers (which might represent both staff and travellers at hotels). were derived from data on long-distance travel in the US[18] (see above). Where the reported nights away was zero, we assumed that travellers spent 0.5 days mixing in the hotel. People in hotels were assumed to have no household, school or workplace contacts while away. Transmission in hotels was assumed to occur at random between all occupants, at the same rate as workplace transmission. Inclusion of air travel more than doubled the computational requirements of the US model.

Many aspects of this air-travel and associated local mixing model are arbitrary, and as such, subject to criticism. For instance, air travellers were not modelled as staying in households or mixing in destination workplaces to minimise computational complexity. However, perhaps the most important test of whether additional detail needed to be included (or different assumptions tested) is the impact air travel makes to transmission dynamics. As Figure SI6 shows, the inclusion of air travel only makes a difference (a delay of 1 week in the peak of the epidemic) to epidemic dynamics (with or without interventions) if the US epidemic is seeded in a single location. This is because even if air travel is absent, there is still a high volume of long-distance travel by other means.

If even as few as 100 seed infections are distributed across the US, rather than being introduced only in 1 place, then the effect of air travel on baseline dynamics is much reduced, being only visible in slightly increasing the synchrony of regional epidemics and hence slightly increasing the national peak case incidence (though not affecting cumulative attack rate or overall epidemic timing). If realistic seeding (1000s of cases being introduced across the US throughout the epidemic) is assumed, then air travel makes no appreciable difference to epidemic dynamics or the effect of interventions. Only if border restrictions were able to reduce imported cases to under 20 would any additional delay be expected from shutting down internal plane travel.

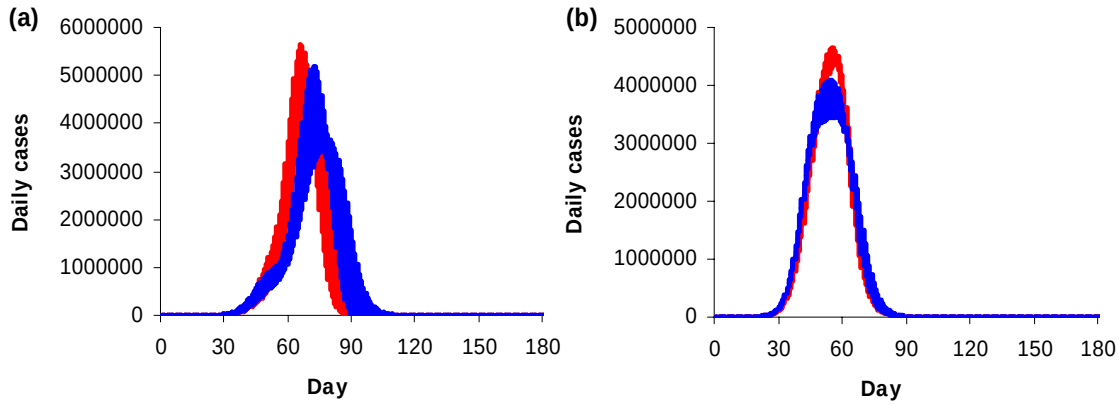


**Figure SI6**: (a) Epidemic curve for $R_0$=2.0 for US when the epidemic is seeded in a single random location on Day 0 with 10 infections. Averages of 10 realizations are shown. The vertical thickness of the curve represents twice the standard deviation seen between runs in the incidence at that time point.. (b) As (a) but with 100 initial infections randomly distributed across the US (proportional to population density).

**Model of infection seeding**

Optimally one would model the seeding of infection in a single country (such as the US or GB), with detailed predictions of numbers of infected people entering the country from different foreign countries derived from a global transmission model. However, building a truly global model is a daunting undertaking. The best that has been done so far are a range of so-called 'patch' models which have modeled spread of epidemics in different cities and the transport of infection between those cities through air travel. That class of models have been used to examine global spread of the 1968 pandemic, and travel patterns were found to be somewhat predictive of the order in which cities were affected[22]. However, in that case the source was known, at least approximately (namely Hong Kong), and overall travel volumes were much less than today. Given the huge uncertainty about where the next pandemic model will start, the slight predictive advantage given by use of a patch model disappears.

For that reason, we use a much simpler 'global' model here – one which just assumes homogenous mixing of the world population – a so-called SEIR model[23] – a much simpler model than the detailed simulation used for within country spread. It gives reasonable estimates for the total numbers of infected individuals that might be expected to enter each country during a pandemic (as it matches expected attack rates), and is qualitatively reasonable in terms of how it distributes those infections over time. However, it cannot be viewed as quantitatively accurate in terms of the projected time course of infections expected to enter GB or the US (which depend on the details of which region is first affected and detailed connectivity patterns between countries).

We assume $R_0$ =1.6 for the SEIR model (to give a cumulative infection attack rate of 64%) and a duration of infectiousness adjusted to give the same real time growth rate as the simulation model with the same $R_0$ – namely $r \approx 0.2$/day. The lower $R_0$ used in the global model is an admittedly ad-hoc adjustment to account for the spatial heterogeneity of the global population, the social structure of populations and the possible impact of travel restrictions [24]. Data from the 1968 pandemic [22] would suggest that international spread, while fast, may not occur so fast as to completely synchronise all epidemics around the world.

The simple model is used to calculate the incidence of infection through time, $I(t)$. The expected number of imported infections per day, $M(t)$, is then given by

$$M(t) = \frac{nL}{365N} I(t),$$

where $n$ is annual number of inbound passengers from overseas to the UK or US, $L$=1.5 days is the average latent period of the disease (we assumed symptomatic individuals – namely clinical cases – would not travel) and $N=6.5\times10^9$ is the global population size. For the UK[25], $n$=92 million, while for the US, $n$=73 million (sum of inbound foreign travellers from I-92 data plus outbound US resident travellers, under the assumption US residents return to the US – see ref [26]). The UK figures are higher than the US (despite a 5 fold difference in population size) because of the much larger number of foreign trips made by GB residents compared with US residents.

Imported cases were assumed to select destinations in the US proportional to population density. This is a crude assumption, as foreign visitors are disproportionately likely to visit large cities, but detailed national data for the UK and US on visitor numbers to each local area were not available. In addition, around 1/3 of inbound travel is due to residents returning home from overseas trips. Such returning travelers might be expected to have more random destinations in the US than foreign visitors.

Figure SI7 shows the temporal profile of the average number of infections entering GB and the US per day resulting from the global model.

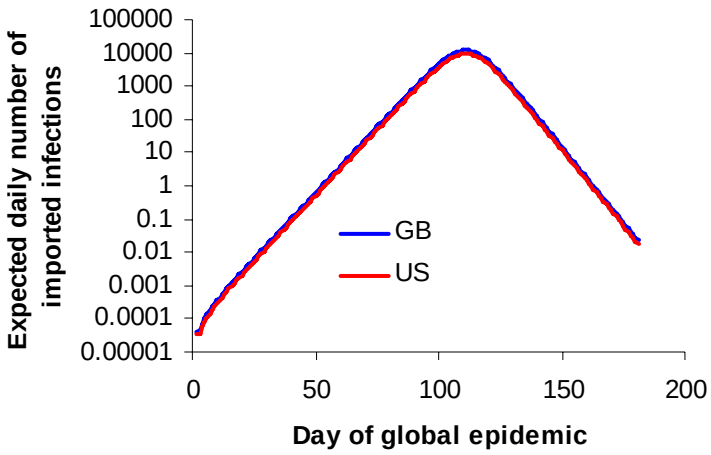


**Figure SI7**: Average number of infected individuals expected to enter the US and GB per day (plotted on a log scale) during a global pandemic, as derived from a simple global model (see text). Actual numbers entering are sampled from a poisson distribution each day, with mean given by the plotted values. The probability of importing infection in the first 30 days is very small.

## Interventions

We examine a large number of different interventions, so it is helpful to state the baseline assumptions made regarding effect and efficacy:

1. **Case detection**: We generally assume that 90% of clinical cases are detected – in the sense of being noted by healthcare systems and potentially receiving some type of response.

2. **Border controls**: The number of infected individuals entering the country from outside is reduced by a fixed proportion (90%, 99% or 99.9%) from a certain time (usually day 30) in the global pandemic.

3. **Area quarantine**: Affected zones are defined by drawing a 20km ring around the household location of each new detected case. Overlapping rings form part of the same affected zone. Area quarantine is the reduction of travel (by a fixed amount – either 75%, 90% or 100%) from an affected zone to an unaffected area and vice-versa for 7 days. We assume it takes 2 days from detection of a case for an existing zone to be extended to include a new case or for a new zone to be established.

4. **Blanket travel restrictions:** All travel further than either 20km or 50km from the home is reduced by a fixed fraction (75%, 90% or 100%) for the duration of the epidemic. It is unclear how such a policy might be enforced in practice.

5. **Antiviral treatment and prophylaxis**: We use recent statistically rigorous estimates of antiviral efficacy[27], as in past work[6]. Treatment of a symptomatic individual is assumed to reduce infectiousness by 60% from the point in time drug is first taken. The overall drop in transmission occurring as a result of treatment therefore depends on when treatment is initiated, but time delays from the onset of infectiousness to the recognition of symptoms and then start of treatment will

inevitably mean that even very prompt treatment can only give reductions of around 25%. Prophylaxis of uninfected individuals is assumed to reduce susceptibility to infection by 30%, reduce infectiousness if infection occurs by 60%, and reduce the probability of clinically recognizable symptoms by 65%. Unless otherwise stated, a 1 day delay from report of symptoms to treatment initiation (and associate close contact group prophylaxis) is assumed. By default, we assume 100% of detected cases receive treatment (i.e. 90% of all cases). If household prophylaxis is being used, 100% household members of treated cases are assumed to be prophylaxed. If social prophylaxis is being used, it is assumed that 90% of the school classmates or workgroup colleagues of 90% of treated cases are reached (giving 80% coverage overall of the school/workplace mixing group of treated cases). Treatment courses are assumed to last 5 days, and prophylaxis 10 days, unless otherwise stated.

6. **Case isolation**: This is assumed to reduce the household, workplace/school and community contact rates of the isolated case by a fixed factor – 90% unless otherwise stated. By default, isolation is assumed to occur 1 day after report of symptoms, and to last for 7 days. In practice, this level of effectiveness of isolation would almost certainly require cases to be removed from the home.

7. **Household quarantine**: This is assumed to reduce the workplace/school and community contact rates of all individuals in a household containing a clinical case and which complies with quarantine by 75% for a period of 14 days from 1 day after the index case reports symptoms. Within-household contact rates are assumed to double during that period. While these parameter values seem plausible, no data exist to provide validation. We assume either 50% of 70% of households of detected cases comply with a quarantine policy.

8. **Reactive school/workplace closure**: Once a school or workplace has accumulated a specified threshold number of pupil/staff member as cases, it closes the day after diagnosis of the case causing the threshold to be breached. Closure is for a fixed period of time, typically assumed to be 3 weeks. Once a school reopens, it can close again if new cases accumulate to the level required for the closure threshold to be breached again. Typically we assume a closure threshold of 1 detected case. With this level of threshold, schools are typically closed for 95% of the time over the peak 12 weeks of the pandemic – irrespective of how long each individual closure period is. Closure of a school/workplace is assumed to eliminate transmission in that location, but to increase household contact rates and community contact rates by 50% and 25% respectively for individuals with a closed school or workplace.

9. **Vaccine**: a vaccine produced from (and therefore precisely matched to) the pandemic virus is assumed to reduce susceptibility of all those receiving it by 70%, and the infectiousness and probability of becoming a clinical case of those who receive it and still get infected by 30% and 50% respectively. Protection is assumed to start 2 weeks after being vaccinated. Multi-dose vaccines are not modelled. A pre-pandemic (and hence perhaps poorly matched) vaccine is assumed to only reduce susceptibility to infection by 30%, but still reduce the infectiousness and probability of becoming a clinical case of those who receive it and still get infected by 30% and 50% respectively. Vaccine coverage is assumed

to be 90% during a pandemic, with those not being vaccinated assumed to be a random cross-section of society.

**Numbers of realisations & computational resources**

It is essential to undertake sufficient realisation to ensure ensemble behaviour of a stochastic is well characterised for any one set of parameter values. For our past work which examined extinction probabilities, this necessitates very large numbers of model realizations being generated. In the current work, only the timing of the initial introduction of virus into a country is potentially highly variable – once case incidence reaches a few hundred cases per day, dynamics are much closer to deterministic. Hence outcome variables such as the cumulative clinical attack rate, or number of antiviral courses needed, varied by less than 0.1% between realisations. Peak daily attack rates and the timing of the first case in a country (and to a lesser extent the timing of the peak of a national epidemic) were more variable (see Figure 1d of main text), but averages could be reasonably well determined with between 5 and 20 realisations (though it should be noted that the standard deviations plotted in the error bars of Figures 2-4 in the main text are less precisely determined) . This is fortunate, as the models used were very computationally intensive – given the population sizes modelled (300m for US) and that each realization ran an epidemic to completion.

The simulation was written in C using the OpenMP 2.0 SMP libraries. The US simulation used some 55GB of RAM, and each realization ran in 1-2 hours on 8 CPU Opteron 854 based servers, and in 2-5 hours on 16 CPUs of the NCSA SGI Altix 3700 system ('cobalt'). The longer run times on the latter system were mainly due to increased memory latency on that larger system. Overall, approximately 20,000 CPU hours were used to generate the US results, and approximately 8000 hours to generate the GB results.

## *3. Natural history and transmission parameters*

Past modelling studies have largely assumed distributions of the incubation and infectious periods first reported in two historical papers[22,28]. However, the primary data for these estimates (in particular, those used to estimate the duration of infectiousness) are lacking. Previous research has assumed constant infectiousness from the end of infection latency to recovery from infection, while we model infectiousness varying continuously through infection. In our previous study we therefore derived new estimates of the incubation period and development of infectiousness over time for human influenza from available primary data [6]. The resulting incubation period distribution and infectiousness function (which incorporates the incubation period and a fitted profile of infectiousness post symptom onset) are reproduced in Figure SI8. They result in an estimated generation time for human influenza of $T_g$=2.6 days. Figure SI8b also shows how the profile of infectiousness we estimate mirrors observed temporal variation in viral shedding.

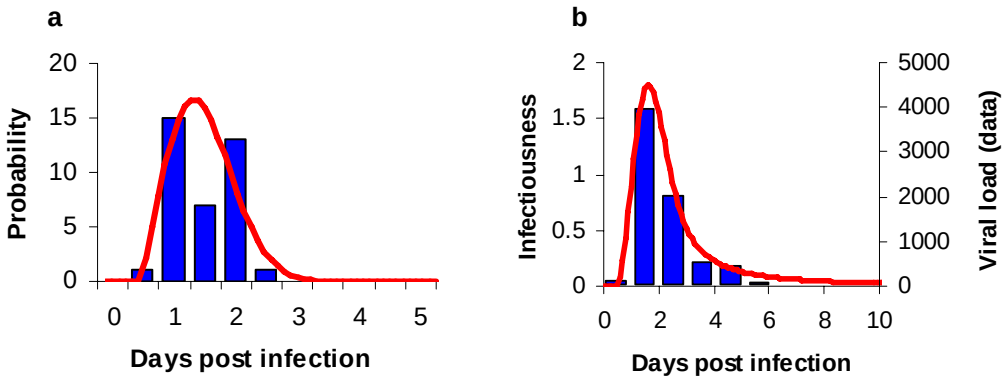


**Figure SI8**: a. Incubation period data (blue)[29] and best fit to offset Weibull distribution (red). b. Mean viral shedding data for experimental influenza infection (blue)[30] compared with the profile of average infectiousness over time from the model (red). Note that the infectiousness profile was not fitted to shedding data. See [6] for more details.

## Estimation of $R_0$ from inter-pandemic influenza household data

In implementing the GB and US models, we discovered an error in the previous implementation[6] of community disease spread when a high proportion of individuals in the population were already infected. This meant that the cumulative attack rates reported in the earlier study[6] (namely 50% infected at $R_0$=1.8) were underestimates (the true figure for Thai parameters is 68%). It should be emphasised that as the error only affected spread once infection rates were saturating (*i.e.* late in an epidemic), none of the detailed results on containment policy effectiveness in the earlier study are changed (even when the slightly altered household transmission parameter discussed below were used).

However, the error did necessitate re-estimation of within- and between- household transmission rates. In addition to fitting to the household data set presented in ref [31], we also optimised the values used in the simulation to fit a subset of household infection data from the Tecumseh study reported by Longini[32]. The data subset used was the infection rates in a season in households where all household members had a 'low' antibody titre to influenza at the start of the season.

Fitting to the latter dataset enables the ratio of within-household to between-household transmission to be estimated, since the dataset reports the proportion of households where no infection was detected. In fitting to this dataset, we used the estimates of population immunity reported in that paper which correspond (crudely) to assuming that 27% of the population is immune at any one time. We assumed immunity was clustered in households and therefore initialized simulations of that dataset by making all members of 27% of households (randomly selected) immune, and assuming complete susceptibility in the remainder.

The best fit of the simulation model to the 2 household datasets gave a within-household $\beta$ of 0.49 for non-severe cases (assumed to be 50% of all infections) and we assumed a 2-fold higher infectiousness for severe cases (namely $\beta$=0.98 – well within the 95% credible interval given in table SI1). Within-household transmission was estimated to be 31% of all transmission in the US and 29% in GB (due to smaller household sizes) and the best estimate of $R_0$ for inter-pandemic influenza from this household dataset was 1.7 (reduced from the previous estimate of 1.8[6]). The fits are not shown here – they are indistinguishable from those shown in the Supplementary Information to ref [6]. The reason for the reduced estimate of $R_0$ is the need to match overall proportions of households affected in the Longini dataset – as cumulative attack rates were previously underestimated, so $R_0$ was overestimated. In the Thailand context, these revised estimates increase the proportion of transmission occurring in the household from 33% used

previously to 37% (for $R_0 = 1.8$). This has a negligible impact on the results of our previous paper[6].

It is worth commenting that the differences between US/GB and Thailand in household sizes and assumed proportion of transmission occurring in schools+workplaces result in significant differences in the total proportion of the population expected to be infected during a pandemic. For $R_0 = 1.8$, 61% are infected in GB or US and 68% in Thailand. For $R_0 = 2.0$, 68% are infected in GB or US, and 76% for Thailand. This dependence of infection (and therefore clinical) attack rate on unknown transmission parameters (i.e. the proportion of transmission occurring in schools/workplaces versus the community as a whole) means that reliable prediction of likely attack rates in a future pandemic is difficult.

We assume no cross-protective immunity between one subtype and another in this work, so the entire population is completely immune to any new pandemic strain. We similarly assume that the 'low-titre' households from the Tecumseh are completely susceptible. If these assumptions are incorrect, then the 'fundamental' $R_0$ of influenza might be larger than estimated, since we would have really estimated the rate of transmission measured in the presence of some immunity. However, such a bias might not affect conclusions about control policies for any new pandemic. If heterosubtypic cross-immunity was significant even in 1918, we might also expect it to be so for any future pandemic, given the 1918 virus is now thought to have an entirely avian origin[33].

Lastly, we did not incorporate the effects of seasonality in this model. In some policy scenarios slowing of the epidemic might - in synergy with seasonality - delay a proportion of cases into the next year.

**Estimation of $R_0$ for influenza from past pandemics**

$R_0$ is defined as the average number of secondary infections caused by a *typical* primary infection in an infinite and entirely susceptible population. The phrase *typical* is important here – in a population with heterogeneous mixing, not all people are equally likely to be infected, so a typical infected individual is not a randomly selected member of the population. For compartmental epidemic models, the solution to this averaging problem is well known: $R_0$ is the dominant eigenvalue of the so-called next-generation operator of the model[34]. For models with network or household or other small-group structure (so-called models with multiple levels of mixing), the definition is less clear. While $R_0$ and other threshold quantities (e.g. $R_*$) have been defined for simple models with multiple levels of mixing, no simple analytical expressions exist for more complex models of the type used here (with age and spatial structure).

We therefore parameterise the model using a related quantity, $R_{rand}$, defined as the average number of secondary infections generated by a randomly selected individual in the population when the rest of the population is susceptible. This quantity is not truly $R_0$ in the sense defined above, as a 'typical' infection is not equivalent to a randomly selected individual. However, the quantity is readily calculated, and it is possible to calculate an overall scaling to the transmission coefficients of the model which results in any selected value of $R_{rand}$.

Testing how far this definition of $R_0$ differs from the 'true' definition is challenging. Verification might be obtained by explicitly tracking the average number of infections generated in each generation of infection in the model – the true $R_0$ value should then be the ratio of infections in one generation to those in the previous generation in the regime where that ratio equilibriates. However, for the population sizes modelled, it was often difficult to identify a region in which the

ratio truly equlibriated. That said, in our previous work[6], we verified that for the baseline household, place and community transmission proportions assumed, $R_0 \approx R_{rand}$.

Another verification is that the realtime growth rate seen is the exponential phase of epidemic growth obeys the correct characteristic equation, namely:

$$1 = R_0 \int_0^\infty \rho(\tau) \mathrm{e}^{-r\tau}\, d\tau \qquad [A]$$

where *r* is the observed real-time growth rate and, $\rho(\tau)$ is the relative infectiousness of an individual time $\tau$ since infection, normalised to be a PDF. In our case, $\rho(\tau)$ (as shown in Figure SI8b) is given by the convolution of the incubation period distribution and the profile of infectiousness as a function of time since the end of incubation. In reality we used the discrete-time equivalent of the above characteristic equation to correct for the fact that the simulation uses a 0.25 day timestep (though the magnitude of this correction was small), namely:

$$1 = R_0\, \mathrm{e}^{-r\Delta T} \sum_{m=0}^{\infty} \rho(m\Delta T) \mathrm{e}^{-rm\Delta T} \qquad [B]$$

where $\Delta T$ is the timestep used (0.25 days).

For SIR epidemic processes with no latent period and an exponentially distributed infectious period, then it is straightforward to show from the characteristic equation [A] that $r = (R_0 - 1)/T_G$ where $T_G$ is the generation time, here equal to the mean infectious period. Solving [B] numerically for the functional form of $\rho(\tau)$ shown in Figure SI8b, $R_0$ was also found to be approximately proportional to *r* for our system, at least in the range $1.5 \le R_0 \le 2.0$, but in this case the relationship was given by $r = (R_0 - 1)/T_{eff}$, where $T_{eff}$ =3.2 days. Note that for our system, $T_G$, as defined by the first moment of $\rho(\tau)$, was 2.6 days – so $T_{eff} > T_G$ (this is a result of the strongly peaked nature of the $\rho(\tau)$ function, not because we use a discrete time formulation).

For the US and GB model, we compared $R_{rand}$ values with $R_0$ values calculated by explicitly tracking infections per generation and with $R_0$ values calculated from the real time growth rate of simulated epidemics using equation [B]. For both measures of $R_0$, we found $R_0 \approx R_{rand}$ +0.2 – unlike the near equality found for the Thai parameter set. This difference appeared to result from the lower level of household transmission in the US/GB models (a result of smaller household sizes) and the assumed higher level of school/workplace transmission.

In previous work[6], in evaluating the range of reproduction numbers seen in the 1918-19 pandemic from mortality time-series for GB and US cities, we used the expression $R_0 = 1 + rT_G$ with $T_G$=2.6 days, however. Then we calculated the maximum *R* value to be approximately 1.8. Using the more accurate expression $R_0 = 1 + rT_{eff}$ with $T_{eff}$ =3.2 days slightly increases these maximum estimates to just below 2 (see Figure SI9). It should be stated that this value represents the absolute maximum growth rate seen in only a handful of locations in the highest growth weeks of the first and second waves of the 1918-19 pandemic. In many ways, $R_0$ =1.8 is a more reasonable estimate of the 'typical' initial growth rate in 1918-19. The initial growth rate seen in the largest GB cities, London and Birmingham was 1.7, which was also the population-size-weighted average initial growth rate.

As a comparison, we looked for data on real-time epidemic growth rates observed in the 1957-58 pandemic. The only weekly time-series data we could readily obtain was for weekly mortality attributed to Influenza in England and Wales[37]. Examining inter-week growth rates for this time-series gives a maximum $R_0$ value of 1.7 (Figure SI9d). This estimate has to be viewed cautiously

– it was only seen in 1 week of the epidemic and mortality was so much less in 1957-58 than 1918-19 that interpretation and analysis of mortality data is subject to potentially much greater systematic and random error. Nevertheless it is interesting that this value agrees with the $R_0$ value obtained from analysing household data (see below).

The lower maximum value of $R_0$ found for the 1957 compared with the 1918 pandemic may have been due to some heterosubtypic immunity induced by the internal proteins shared between the pre and post 1957 influenza A subtypes (H1N1 and H2N2), but another hypothesis is that reductions in average household sizes between 1918 and 1957 reduced transmission rates[38]. The debate is not an academic one; if household size reduction is responsible then post 1957 reductions in household sizes in the developed world might lead to corresponding reductions in rates of spread of future influenza pandemic viruses. If cross-immunity accounted for the lower $R_0$ of the 1957 virus, then future pandemics might attain the level of transmissibility seen in 1918 if there is little heterosubtypic immunity to the emergent virus.

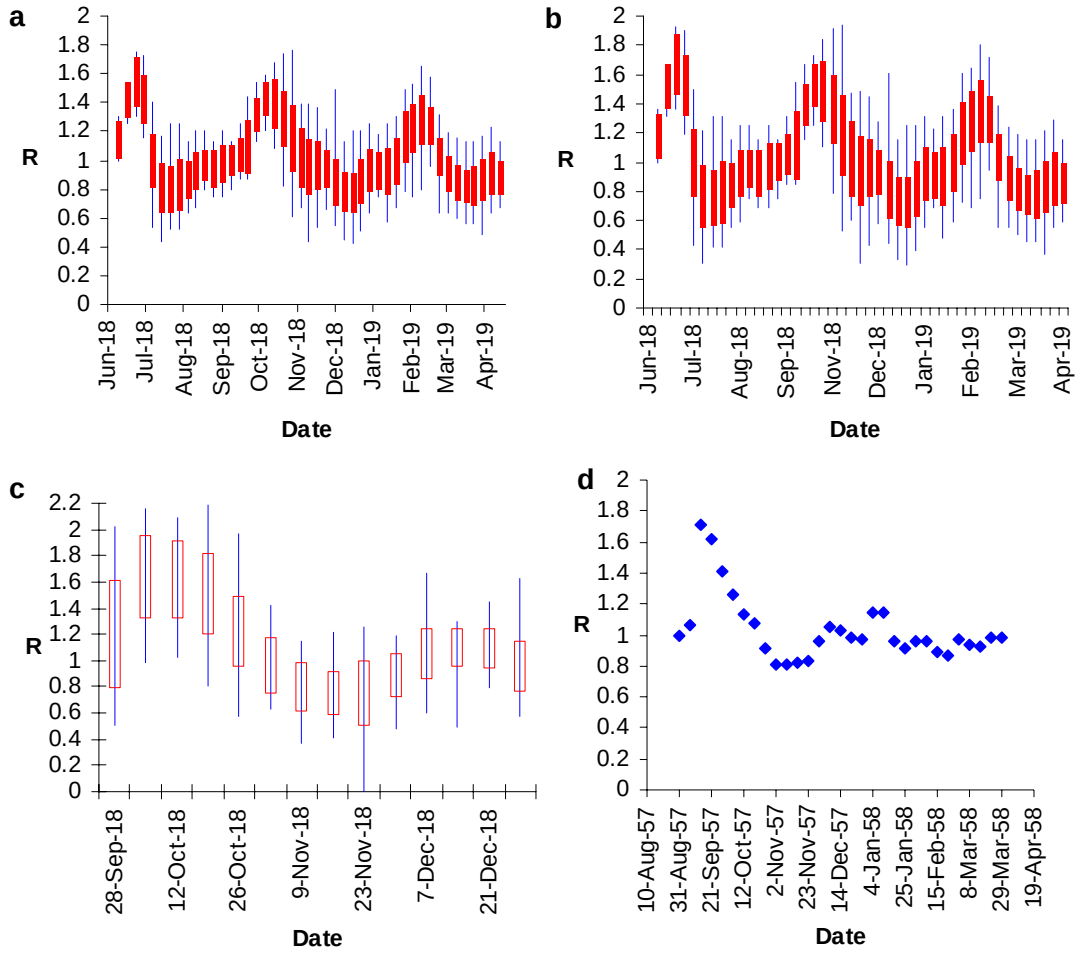


**Figure SI9**: (a) Weekly estimates of the effective reproduction number, *R*, derived (see text above) by calculating the running average exponential growth rate calculated over a 3 week window from mortality data collected in 83 British cities in 1918-19 [35]. Blue line shows minimum and maximum values seen in the data, red bar shows the mean value ± the standard deviation in the observed values. Estimates use the relation $R_0 = 1 + rT_G$ with $T_G$ =2.6 days. (b) As (a) but using the relation $R_0 = 1 + rT_{eff}$ with $T_{eff}$ =3.2 days. The maximum value seen approaches 2.0 for a few cities, but 1.7-1.8 is more typical of the maximum reproduction number in most cities. (c) as (b) but showing *R* estimates from US mortality data from 47 cities (reported in ref [36]). The maximum value seen is 2.2 for 2 cities, 2.0 for 14 cities, but the average maximum value across all cities is 1.8. (d) As (b) but showing single point estimate for *R* derived from overall influenza-attributed mortality during the 1958 pandemic in England and Wales (data from ref [37]).

**Assumptions about population behaviour during a pandemic**

It is also interesting to examine the extent to which *R* values predict the overall shape of past pandemic curves. It is clear that assuming $R_0$ is constant throughout a pandemic poorly matches observed behaviour for the 1918-19 and 1957-58 pandemics (as is also clear from the *R* estimates in Figure SI9). This – together with the currently unexplained phenomenon of multiple pandemic 'waves' - highlights our lack of understanding of key processes underlying the transmission of pandemic influenza in the past. Seasonality in transmission, spontaneous changes in population contact rates and antigenic evolution of the virus are all possible hypotheses for the patterns seen.

If one tries to qualitatively fit the spring and autumn waves of the 1918-19 pandemic with the results of our simulation model assuming constant $R_0$, then Figure SI10a shows that $R_0$ = 1.7 and 1.5 are, respectively, better estimates of the 'overall' $R_0$ value of those 2 waves – though the fit to the decay phase of the pandemic curves is poor for the second wave – probably because of the effect of the end of the first world war and military demobilisation on enhancing population contact rates in mid to late November 1918 (data for individual cities (the same hiatus is not seen in most individual city data). Note that in making this comparison we do not allow for the larger household sizes in 1918 in the simulation. Similarly, Figure SI10b shows that $R_0$ =1.5 fits the main wave of the 1957-58 pandemic better than the maximum value of 1.7. Obviously (given Figure SI9) these fits underestimate the growth rates seen at the start of the pandemic waves.

Assuming a degree of spontaneous behaviour change (such as temporary closure of schools and some workplaces once incidence exceeds some threshold, or reduction in community contact rates in response to high incidence) in the face of a pandemic can give more satisfactory fits to the entire pandemic waves with $R_0$ values which match maximum initial growth rates (results not shown). We know that schools closed, some mass gatherings were cancelled and that healthy adults stayed at home to look after sick family members during past pandemics[39], so such assumptions are not unreasonable. However, since we have no data with which to accurately parameterise spontaneous changes in contact rates, we made the decision to adopt baseline epidemic scenarios which assume no such behavioural changes – and then to examine the impact of school closure and social distancing on those scenarios.

The baseline scenarios used here should therefore be viewed as worst cases in terms of the speed with which the epidemic peaks and therefore the consequent peak height – while early spread may occur at the maximal transmission rates seen in past pandemics, later spread may be slowed by spontaneous behavioural changes which predominantly reduce peak incidence and extend the duration of the epidemic rather than reduce cumulative attack rates.

We do assume clinical disease affects individual behavior. Thus while we assume that clinical cases are twice as intrinsically infectious as non-clinical infections[6], we also assume that from the onset of symptoms they reduce their school, workplace and community contact rates by 90%, 50% and 50% respectively (where applicable). This is intend to represent the effect of sickness-related absenteeism and withdrawal to the home. We were unable to find data from past pandemics to enable these contact rate reductions to be estimated. Household contact rates are assumed to be unchanged.

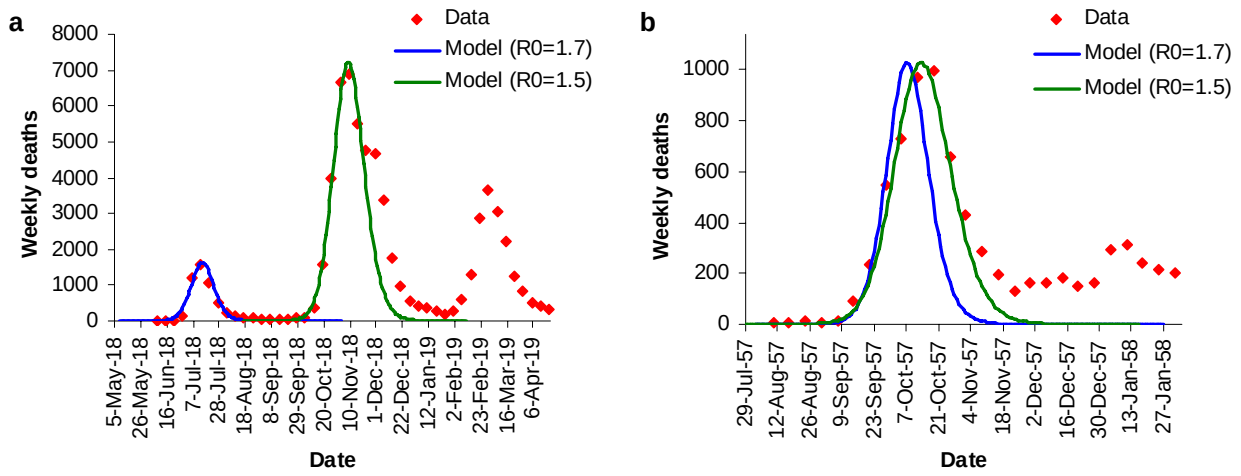


**Figure SI10**: (a) Fit of the GB simulation model with $R_0$ =1.7 and 1.5 to, respectively, the spring and autumn waves of the 1918-19 pandemic in GB. Data shown are total mortality in the 83 cities in the dataset reported in ref [35]. (b) as (a) but with showing how simulations with $R_0$ =1.7 and 1.5 fit the pattern of overall influenza-attributed mortality during the 1957 pandemic in England and Wales[37].

## Transmissibility scenarios

In light of the analysis above, we focus on 2 transmissibility scenarios in the current work: (i) a 'moderate' scenario of $R_0$ =1.7, which might be thought to best represent the maximum transmissibility seen at the start of the 1957 pandemic and typical transmissibility in 1918, and (ii} a 'high' scenario of $R_0$ =2.0, which represents the maximum transmissibility seen in the 1918-19 pandemic. The moderate transmissibility scenario is perhaps more likely, while the high transmissibility scenario is included as a feasible worst case – assuming for both that the next pandemic is caused by a virus with clinical and epidemiological characteristics similar to those seen in past pandemics.

## *4. Impact of combination policies*

The interventions described in the main text are unlikely to be applied in isolation, so here we explore a representative set of combination policies.

Examining these tables (and repeating the comments in the main text), a number of conclusions stand out. Firstly, policies involving immediate vaccination of even a small proportion of the population with poorly matched vaccine have a substantially enhanced impact, due to the effect of such vaccination on reducing the effective reproduction number of the pandemic virus. In addition, the availability of such vaccine stockpiles reduces the size of antiviral stockpile needed for either household or school/workplace prophylaxis policies. Second, household quarantine is potentially a highly effective social distance measure, but only if compliance with the policy is good. Reactive school closure (Tables SI1 and 2) has a limited impact in reducing peak attack rates (but not cumulative attack rates) and enhancing other policies. For the high transmissibility scenario ( $R_0$ =2.0) only policies relying only on 99% border controls can delay spread by enough to enable pandemic vaccine to substantially reduce attack rates (vaccine is assumed to be produced from month 4 of the global epidemic) – and then only if coupled with large-scale social and household prophylaxis.

The impact of interventions (whether singly or in combination) is virtually identical for the US and GB, despite their different physical and population sizes (see Table SI1 and SI2). The only distinctions are in the timing of the pandemic: measured from the start of the global epidemic, the US epidemic takes approximately 15 days longer to peak (relative to the first case in the country) in the US compared with GB, mainly as a result of the 5-fold larger population of the US, though the larger visitor numbers into the UK also contributes somewhat to the faster spread there.

Detailed planning of any combined intervention strategy would be essential – to ensure rapid delivery, but also to ensure that particularly vulnerable populations not explicitly modelled here (such as the institutionalized) might also be protected in a resource-efficient manner.

**Table SI1.** Impact of different combinations of interventions on the impact of the first 220 days of a US influenza pandemic (average of 5-20 realisations shown – SD across realisations is <0.1% for cumulative attack rate and stockpile size, <0.04% for peak attack rate, and 1 to 8 days for peak delay). All antiviral and social distance measures start the day after symptoms in the index case (including case treatment), as does area quarantine. If imposed, area quarantine is assumed to reduce travel between affected and unaffected areas by 90%, with affected areas being any area within 20km of the household location of a case diagnosed in the last 7 days. Blanket movement restrictions are assumed to reduce all travel over 50km by 90% from the start of the national epidemic for its duration. Reactive school closure policy closes schools for 3 weeks after the first detected case. A new case after that period will again close the school. Schools therefore end up being closed for 95% of the worst 12 weeks of the pandemic. Impact is quantified by the cumulative attack rate over the first year (in the absence of pandemic vaccine), peak attack rate and any delay in peak timing induced. Also shown is the cumulative attack rate if a matched pandemic vaccine started to become available (at the rate of 1% coverage per day) 4 months after the start of the global pandemic. For each outcome, results for $R_0$=2.0 and 1.7 are both shown (left and right sub-columns respectively within each outcome column).

| | Antivirals | | | Social distance measures | | | | Travel restrictions | | | | Outcome for $R_0$=2.0/1.7 (high/moderate transmissibility) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Policy number | Case treatment (% treated) | Household prophylaxis (% treated) | Social prophylaxis (% coverage) | Case isolation (% isolated) | Household quarantine (% compliance) | Reactive school closure | Reactive workplace closure (% closed) | Border restrictions (% reduction in imports) | 20km area quarantine (% reduction in travel) | 50km blanket travel reduction (% reduction) | Prevaccination (% vaccinated) | Antiviral stockpile needed (% coverage) | | Cumulative clinical attack rate (%) | | Peak clinical attack rate (%) | | Delay in peak (days) relative to scenario with no intervertions | | Cum. attack rate with vacc from 4 months (%) | |
| 1 | | | | | | | | | | | | 0 | 0 | 34 | 27 | 1.9 | 1.2 | 0 | 0 | 0 | 0 |
| 2 | | | | 50 | 50 | | | | | | | 0 | 0 | 29 | 22 | 1.4 | 0.9 | 4 | 5 | 29 | 21 |
| 3 | | | | 70 | 70 | | | | | | | 0 | 0 | 27 | 20 | 1.3 | 0.7 | 5 | 7 | 27 | 19 |
| 4 | | | | | 50 | ✓ | 10 | | | | | 0 | 0 | 29 | 21 | 1.2 | 0.7 | 13 | 18 | 28 | 17 |
| 5 | | | | | 50 | ✓ | 33 | | | | | 0 | 0 | 27 | 18 | 1.1 | 0.5 | 15 | 21 | 27 | 14 |
| 6 | | | | | 50 | ✓ | 10 | | 90 | | | 0 | 0 | 25 | 15 | 0.9 | 0.3 | 22 | 32 | 23 | 8 |
| 7 | | | | | 50 | ✓ | 10 | 99 | | | | 0 | 0 | 29 | 20 | 1.0 | 0.5 | 39 | 53 | 21 | 6 |
| 8 | | | | | 50 | ✓ | 10 | 99 | | | 10 | 0 | 0 | 25 | 15 | 0.8 | 0.3 | 46 | 65 | 15 | 2 |
| 9 | | | | | 50 | ✓ | 10 | 99 | | | 20 | 0 | 0 | 20 | 9 | 0.6 | 0.2 | 51 | 77 | 10 | 1 |
| 10 | | | | | 50 | ✓ | 10 | 99 | 90 | | | 0 | 0 | 25 | 9 | 0.7 | 0.2 | 54 | 88 | 11 | 1 |
| 11 | | | | | 50 | ✓ | 10 | 99 | | 90 | | 0 | 0 | 27 | 17 | 0.8 | 0.4 | 44 | 60 | 16 | 3 |
| 12 | 50 | | | | 50 | ✓ | 10 | | | | | 13 | 9 | 27 | 19 | 1.1 | 0.5 | 15 | 22 | 27 | 15 |
| 13 | 90 | | | | | ✓ | 10 | | | | | 27 | 19 | 30 | 22 | 1.3 | 0.7 | 10 | 17 | 30 | 18 |
| 14 | 90 | | | | | ✓ | 33 | | | | | 26 | 18 | 29 | 20 | 1.2 | 0.6 | 13 | 19 | 29 | 16 |
| 15 | 50 | | | | 70 | ✓ | 10 | | | | | 13 | 9 | 26 | 17 | 1.0 | 0.4 | 15 | 23 | 25 | 12 |
| 16 | 90 | | | | | ✓ | 10 | | | | 10 | 23 | 15 | 26 | 17 | 1.0 | 0.4 | 15 | 22 | 25 | 12 |
| 17 | 90 | | | | | ✓ | 10 | | | | 20 | 19 | 11 | 21 | 12 | 0.8 | 0.3 | 19 | 27 | 20 | 8 |
| 18 | 90 | | | | 70 | ✓ | 10 | | | | 20 | 15 | 7 | 17 | 8 | 0.5 | 0.13 | 24 | 34 | 15 | 4 |
| 19 | 90 | 90 | | | | ✓ | 10 | | | | | 49 | 32 | 20 | 13 | 0.7 | 0.3 | 17 | 26 | 19 | 8 |
| 20 | 90 | 90 | | | 70 | ✓ | 10 | | | | | 42 | 24 | 17 | 9 | 0.6 | 0.2 | 21 | 35 | 16 | 5 |
| 21 | 90 | 90 | | | | ✓ | 10 | | | | 20 | 35 | 17 | 14 | 7 | 0.4 | 0.11 | 24 | 34 | 12 | 3 |
| 22 | 90 | 90 | | | 70 | ✓ | 10 | | | | 20 | 30 | 12 | 12 | 5 | 0.3 | 0.07 | 27 | 37 | 9 | 2 |
| 23 | 90 | 90 | 80 | | | ✓ | 10 | | | | | 90 | 37 | 11 | 4 | 0.2 | 0.05 | 27 | 38 | 9 | 2 |
| 24 | 90 | 90 | 80 | | 70 | ✓ | 10 | | | | | 75 | 20 | 9 | 2 | 0.2 | 0.02 | 33 | 20 | 6 | 1 |
| 25 | 90 | 90 | 80 | | | ✓ | 10 | | | | 20 | 62 | 15 | 8 | 2 | 0.1 | 0.02 | 36 | 30 | 4 | 1 |
| 26 | 90 | 90 | 80 | | 70 | ✓ | 10 | | | | 20 | 48 | 8 | 6 | 1 | 0.1 | 0.01 | 40 | 2 | 3 | 1 |
| 27 | 90 | 90 | 80 | | | ✓ | 10 | 99 | | | | 81 | 11 | 10 | 1 | 0.2 | 0.03 | 65 | 97 | 4 | 0 |
| 28 | 90 | 90 | 80 | | 70 | ✓ | 10 | 99 | | | | 60 | 3 | 7 | 0 | 0.1 | 0.01 | 76 | 97 | 2 | 0 |
| 29 | 90 | 90 | 80 | 70 | 70 | ✓ | 10 | 99 | | | | 53 | 2 | 6 | 0 | 0.1 | 0.01 | 80 | 96 | 1 | 0 |
| 30 | 90 | 90 | 80 | | | ✓ | 10 | 99 | | | 20 | 42 | 1 | 5 | 0 | 0.1 | 0.00 | 89 | 97 | 0 | 0 |

**Table SI2:** As Table SI1, but for GB.

| | Antivirals | | | Social distance measures | | | | Travel restrictions | | | | Outcome for $R_0$=2.0/1.7 (high/moderate transmissibility) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Policy number | Case treatment (% treated) | Household prophylaxis (% treated) | Social prophylaxis (% coverage) | Case isolation (% isolated) | Household quarantine (% compliance) | Reactive school closure | Reactive workplace closure (% closed) | Border restrictions (% reduction in imports) | 20km area quarantine (% reduction in travel) | 50km blanket travel reduction (% reduction) | Prevaccination (% vaccinated) | Antiviral stockpile needed (% coverage) | | Cumulative clinical attack rate (%) | | Peak clinical attack rate (%) | | Delay in peak (days) | | Cum. attack rate with vacc from 4 months (%) | |
| 1 | | | | | | | | | | | | 0 | 0 | 34 | 28 | 2.1 | 1.4 | 0 | 0 | 34 | 28 |
| 2 | | | | 50 | 50 | | | | | | | 0 | 0 | 29 | 23 | 1.5 | 1.0 | 0 | 5 | 29 | 23 |
| 3 | | | | 70 | 70 | | | | | | | 0 | 0 | 27 | 21 | 1.4 | 0.9 | 2 | 6 | 27 | 20 |
| 4 | | | | | 50 | ✓ | 10 | | | | | 0 | 0 | 28 | 20 | 1.1 | 0.6 | 12 | 19 | 28 | 19 |
| 5 | | | | | 50 | ✓ | 33 | | | | | 0 | 0 | 27 | 18 | 1.1 | 0.6 | 12 | 19 | 27 | 18 |
| 6 | | | | | 50 | ✓ | 10 | | 90 | | | 0 | 0 | 28 | 16 | 1.1 | 0.4 | 16 | 23 | 26 | 15 |
| 7 | | | | | 50 | ✓ | 10 | 99 | | | | 0 | 0 | 28 | 19 | 1.1 | 0.5 | 34 | 48 | 26 | 11 |
| 8 | | | | | 50 | ✓ | 10 | 99 | | | 10 | 0 | 0 | 23 | 14 | 0.8 | 0.3 | 44 | 60 | 21 | 5 |
| 9 | | | | | 50 | ✓ | 10 | 99 | | | 20 | 0 | 0 | 19 | 10 | 0.6 | 0.2 | 47 | 68 | 14 | 4 |
| 10 | | | | | 50 | ✓ | 10 | 99 | 90 | | | 0 | 0 | 26 | 15 | 0.9 | 0.3 | 43 | 60 | 22 | 5 |
| 11 | | | | | 50 | ✓ | 10 | 99 | | 90 | | 0 | 0 | 26 | 17 | 0.9 | 0.4 | 43 | 56 | 22 | 7 |
| 12 | 50 | | | | 50 | ✓ | 10 | | | | | 13 | 9 | 27 | 18 | 1.1 | 0.6 | 13 | 20 | 27 | 17 |
| 13 | 90 | | | | | ✓ | 10 | | | | | 26 | 19 | 29 | 21 | 1.2 | 0.7 | 10 | 15 | 29 | 20 |
| 14 | 90 | | | | | ✓ | 33 | | | | | 25 | 17 | 28 | 19 | 1.2 | 0.6 | 10 | 15 | 28 | 18 |
| 15 | 50 | | | | 70 | ✓ | 10 | | | | | 13 | 8 | 25 | 17 | 1.0 | 0.5 | 15 | 22 | 25 | 15 |
| 16 | 90 | | | | | ✓ | 10 | | | | 10 | 22 | 14 | 25 | 16 | 1.0 | 0.4 | 16 | 22 | 24 | 14 |
| 17 | 90 | | | | | ✓ | 10 | | | | 20 | 18 | 10 | 20 | 11 | 0.7 | 0.3 | 20 | 22 | 20 | 10 |
| 18 | 90 | | | | 70 | ✓ | 10 | | | | 20 | 14 | 7 | 16 | 8 | 0.5 | 0.2 | 24 | 20 | 15 | 6 |
| 19 | 90 | 90 | | | | ✓ | 10 | | | | | 46 | 30 | 20 | 12 | 0.7 | 0.3 | 18 | 22 | 20 | 11 |
| 20 | 90 | 90 | | | 70 | ✓ | 10 | | | | | 40 | 23 | 17 | 9 | 0.6 | 0.2 | 19 | 23 | 17 | 8 |
| 21 | 90 | 90 | | | | ✓ | 10 | | | | 20 | 32 | 17 | 14 | 7 | 0.4 | 0.14 | 25 | 20 | 13 | 5 |
| 22 | 90 | 90 | | | 70 | ✓ | 10 | | | | 20 | 27 | 13 | 12 | 5 | 0.3 | 0.10 | 27 | 17 | 10 | 4 |
| 23 | 90 | 90 | 80 | | | ✓ | 10 | | | | | 92 | 49 | 11 | 5 | 0.2 | 0.09 | 25 | 14 | 9 | 4 |
| 24 | 90 | 90 | 80 | | 70 | ✓ | 10 | | | | | 78 | 36 | 9 | 3 | 0.2 | 0.07 | 25 | 12 | 7 | 3 |
| 25 | 90 | 90 | 80 | | | ✓ | 10 | | | | 20 | 64 | 27 | 8 | 3 | 0.1 | 0.06 | 27 | 12 | 6 | 2 |
| 26 | 90 | 90 | 80 | | 70 | ✓ | 10 | | | | 20 | 53 | 21 | 6 | 2 | 0.1 | 0.05 | 25 | 10 | 5 | 2 |
| 27 | 90 | 90 | 80 | | | ✓ | 10 | 99 | | | | 85 | 36 | 9 | 3 | 0.1 | 0.04 | 52 | 67 | 6 | 1 |
| 28 | 90 | 90 | 80 | | 70 | ✓ | 10 | 99 | | | | 68 | 12 | 7 | 1 | 0.1 | 0.02 | 68 | 111 | 4 | 0 |
| 29 | 90 | 90 | 80 | 70 | 70 | ✓ | 10 | 99 | | | | 64 | 10 | 7 | 1 | 0.1 | 0.01 | 69 | 115 | 3 | 0 |
| 30 | 90 | 90 | 80 | | | ✓ | 10 | 99 | | | 20 | 53 | 4 | 6 | 0 | 0.1 | 0.01 | 78 | 157 | 2 | 0 |

## 5. Sensitivity analyses

**'H5N1-like' infection characteristics**

In everything above, we have assumed that any future pandemic virus will cause infections with similar biological, clinical and epidemiological characteristics to past human influenza viruses. However, the human infections with the avian H5N1 virus seen to date have been much more severe and protracted than typical human influenza[40,41]. We therefore model how conclusions would change if a transmissible virus arose with current H5N1-like biological characteristics.

We assume a 3 day incubation period and 12 days of constant infectiousness (giving $T_g$ =9 days), and that 75% of those infected become clinical cases, 90% of which withdraw to the home after the onset of clinical symptoms. This increases the proportion of transmission occurring in the household to 49%, with 24% occurring in schools/workplaces, and 27% through random community contacts (for $R_0$ =1.8). With these parameter, a combined case isolation, treatment and household prophylaxis strategy (see Figure SI11) would reduce clinical cases by 2/3 for $R_0$ =1.8 (with a 46% stockpile) – slightly better than for 'human-like' virus parameters. The improvement is largely due to the greater proportion of transmission occurring in the household.

Much more importantly, the global pandemic (even in the absence of controls) would be much slower – for $R_0$ =1.8, the first cases would not be expected into GB or the US until month 7 of the global pandemic or later, and the GB epidemic would not peak until month 16. Even if $R_0$ were as high as 2.8, the global rate of spread would still be some 40% slower than for a 'human-like' virus ($T_g$ =2.6 days) with $R_0$ =1.8.

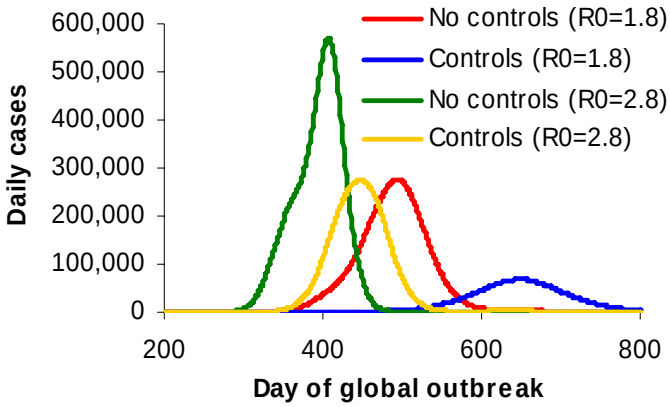


**Figure SI11:** Average (of 10 realisations) epidemic curve (for GB) for an 'avian H5N1-like' pandemic virus (see text), assuming no interventions or policy of 90% case treatment (for 15 days), isolation of 80% of treated cases on day 3 of symptoms, and prophylaxis of household members (with treatment dose, for 15 days). Curves are shown for $R_0$=1.8 and 2.8.

There is a caveat to this discussion: spread is only slower if $T_g$ increases so long as $R_0$ does not also increase. The worst-case scenario would be that a 'H5N1-like' pandemic would have daily infectiousness comparable with peak infectiousness in a 'normal' human influenza infection, but that instead of this level of infectiousness only lasting 1-2 days, it might continue for 7+ days. Such a scenario would result in $R_0$ values of 6-10, with little hope for any of the controls considered in this paper having a significant impact – outside the direct clinical impact of antiviral treatment. Arguably, however, if influenza had the potential to have such a high $R_0$

value, mutation would have already generated such a variant – given the huge epidemiological fitness advantage to be gained from having a 3-5-fold higher total infectiousness.

**Within-place group structure and targeting**

As mentioned in section 2, we incorporated a group structure within the definition of places (e.g. schools, workplaces) in the model to represent school classes or groups of closely working colleagues in a workplace. The key reason for doing so is that the school class in particular is a social unit which might be targeted by control policies. While everyone in a school or workplace can contact any other person in that location, we assume higher contact rates between people in the same group than between people in different groups.

More specifically, our default assumption is that the within-group transmission coefficient is 4-fold higher than the coefficient for between-group transmission. We tested the impact of this assumption by undertaking model runs with no between-group mixing and within between-group mixing occurring at the same level as within-group (i.e. so the group structure is irrelevant), while keeping the overall proportion of transmission occurring in schools and workplaces constant.

The ratio of within- to between- group mixing had no impact on baseline dynamics or the effectiveness of any control measure other than those which specifically target the school class or work-group. For a very intensive policy of next day treatment of 90% of cases, household prophylaxis, reactive school closure and prophylaxis of 80% of individuals in the same class or work group as treated cases, baseline parameters give (for $R_0$=2.0) a cumulative clinical attack rate of 11%, and the policy requires an antiviral stockpile of 93% of population size. If there is no between-group transmission, these figures drop to 10% and 89% respectively, while if between-group and within-group transmission coefficients are identical, the figures rise to 12% and 102%. This scale of differences is similar for lower values of $R_0$.

It is also interesting to compare the relative impact of group- versus place-targeting. If the same policy is adopted, the entire school or workplace is prophylaxed (instead of just individuals in the same class or work group), then the cumulative attack rate is reduced to 5.7%, but at the cost of needing enough drug to treat 300% of the population (for $R_0$=2 – or 200% for $R_0$=1.7). Place-targeting is therefore far less efficient than group targeting, even when group structure has no impact epidemiologically other than to stage prophylaxis in a school or workplace.

**School closure policies**

In this paper we have focused on presenting results for a policy of *reactive* school closure. In this policy, a school is closed for a fixed period following the first case in the school. The school then reopens, but can close again if more cases are diagnosed. In addition, 10% of workplaces are assumed to close in a similar manner (this is intended to crudely represent the impact of absenteeism of carers and spontaneous closure of a minority of non-essential workplaces).

Figure SI12 shows how the impact of a reactive school closure policy depends on the duration of closure following diagnosis of cases in the school population. The best impact is of course achieved by 12 or more weeks of continuous closure, but the dependence is weak, as shorter periods of closure are partly compensated for by schools closing repeatedly (as new cases arise after each school reopens). For the high-transmissibility scenario, all closure periods of 3 weeks or more give cumulative attack rates between 28% and 29%. School closure has a much larger impact on peak attack rates, and paradoxically, 3 weeks of closure at a time minimises peak attack rates for the high-transmissibility scenario. We think this is because this choice of closure

period maximises the desynchrony of timing of closure between schools, while still having sufficient overall impact to substantially reduce transmission rates in schools.

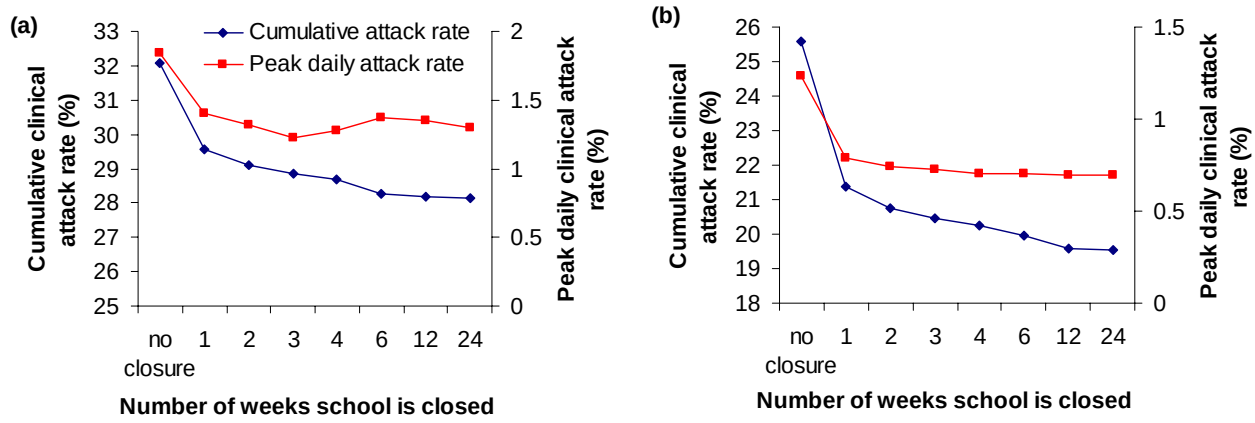


**Figure SI12:** Dependence of the effect of school closure on the number of weeks each school is closed Results assume a school is closed the day after the first case is detected in the staff of pupils of the school. 10% of workplaces are also assumed to close for the same period. A 90% case detection rate is assumed. The cumulative and peak daily attack rates are shown. (a) results for $R_0$=2.0, (b) results for $R_0$=1.7. Treatment of detected cases on the day after symptoms is also assumed. GB model results are shown.

Figure SI13 shows how the impact of reactive school closure depends on the threshold number of cases which need to accumulate in the staff and pupils of a school before closure is triggered. As expected, the impact of closure is highest for closure after the first case, but impact does not diminish significantly so long as closure occurs before 5 cases have accumulated. This is encouraging for the robustness of the reactive school closure policy to poor case reporting rates.

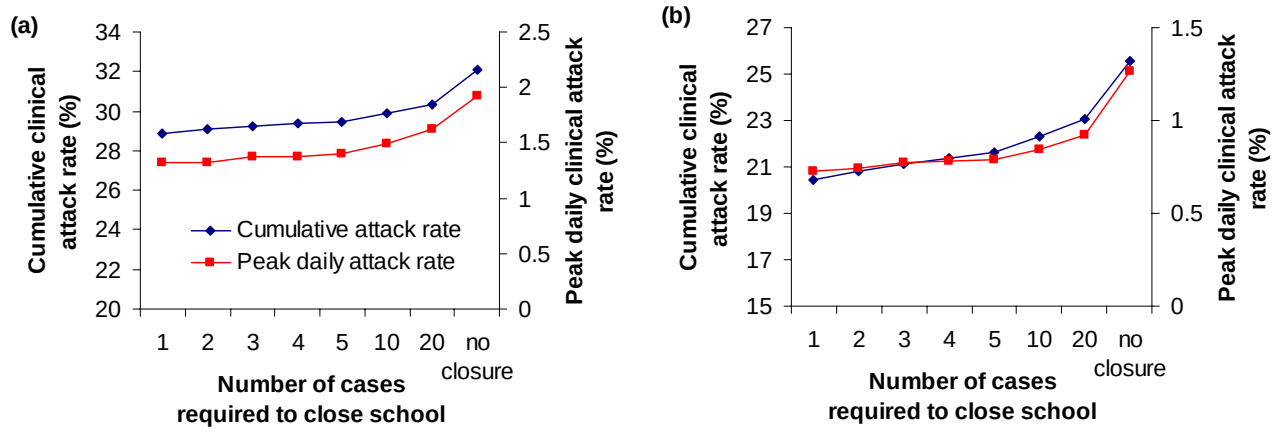


**Figure SI13:** Dependence of the effect of school closure on the threshold number of diagnosed cases which need to accumulate in a school before reactive closure occurs. A 90% case detection rate is assumed. The cumulative and peak daily attack rates are shown. (a) results for $R_0$=2.0 and schools closing for 3 weeks at a time, (b) as (a) but for $R_0$=1.7. Treatment of detected cases on the day after symptoms is also assumed. GB model results are shown.

This is not the only model of school closure of course. As an alternative we also examined area closure, in which when a case arises in an area, all schools within 10km of that case are closed for a fixed period. For $R_0$ =2.0, cumulative and peak daily attack rates in the UK were reduced from 34% and 2.1% respectively to 31.5% and 1.6% respectively by the reactive policy of 3 weeks of closure, while the area policy reduced them to 31.2% and 1.6% respectively for the same duration

of closure. For $R_0$=1.7, cumulative and peak daily attack rates were reduced from 28% and 1.2% respectively to 23.7% and 0.93% respectively by the 3 week reactive policy, while the area policy reduced them to 23.3% and 0.90% respectively. These results are for a policy of school closure alone (i.e. no associated treatment). The marginal difference between the policies reflects the fact that both result in nearly all schools being shut for 95% of the time during the peak 12-14 weeks of the pandemic. Which policy is to be preferred therefore really depends on which is easier to implement in a particular local or national context.

Lastly, we note that compensatory changes in contact rates in other social contexts are likely to accompany school closure. Here we have assumed a 50% increase in contacts in the household, and a 25% increase in community contacts (overall these increases are consistent with the increased time spent at home). The split of the increases between household and community is intended to represent the fact that even if children are advised to stay at home during a school closure period (so that inter-child mixing is limited), childcare maybe performed by people other than members of a child's household (e.g. carers, extended family members etc). Assuming parents do all the additional childcare needed (and no increases in community contact rates) might result in increases in household contact rates of around 70% - but the predicted impact of school closure is relatively insensitive to this change.

**Natural history parameters**

Throughout this work we have used estimates of the incubation period distribution and profile of infectiousness as a time from the end of latency which we believe represent the most accurate characterisation of influenza natural history to date. However, most other published modelling studies of influenza transmission[42,43] have used historical parameters for which the basis in data is unclear[28]. Here we examine how adopting this more commonly used parameterisation (namely a mean latent period of 1.2 days on average, and an infectious period of 4.1 days on average, each distributed as described in the literature[28]) affects the timescale of epidemic dynamics. We preserve the proportions of transmission occurring in households, schools, workplaces and the wider community, and the 50% proportion of infections which become clinical cases.

The observed growth rate using the historical parameters almost exactly matches what would be expected analytically, and is approximately given by $r \approx (R_0 - 1)/T_{eff}$ with $T_{eff}$=4.5 days. Note that this set of parameters gives a $T_{eff}$ substantially less than the actual generation time of $T_g$=5.3 days, unlike the default set of parameters we use in the rest of this paper (for which $T_g$=2.6 days and $T_{eff}$=3.2 days). Nevertheless, for the same value of $R_0$, the growth rate of the pandemic is considerably slower for the historical parameter set than our default parameter set (Figure SI14).

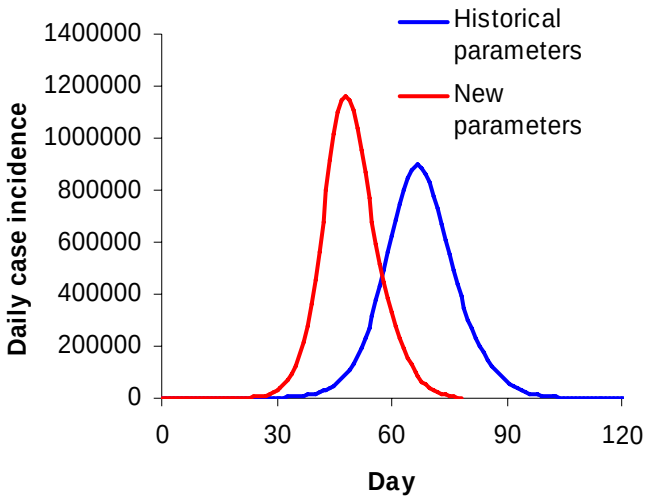


**Figure SI14:** Comparison of GB epidemic curves for the newly derived natural history parameters used in this paper with those obtained using historical parameters, for $R_0$ =2.0 and all other parameters identical. Epidemics are seeded on day 0 with 10 infections, and no other seeding is modelled. The cumulative number of infections is identical for both parameter sets.

The projected impact of most control policies are similar for both old and new natural history parameters for the same value of $R_0$ when delays in detecting and treating clinical cases and implementing prophylaxis are short (<24h) – though in general, policies relying in case detection have slightly larger impacts due to the smaller proportion of onward transmission which has occurred by the time a case is treated. For >24h delays in treatment the difference in predicted policy efficacy for the two parameter sets increases, with the historical parameters producing substantially more optimistic conclusions (e.g. a 2/3 rather than 1/2 reduction in cumulative clinical attack rates from a policy of next day case treatment, household prophylaxis and reactive school closure).

However, both models need to be calibrated against the growth rates seen in past pandemics. Since the historical parameters have longer $T_g$ (and thus $T_{eff}$), then estimates of $R_0$ from past pandemics increase. Hence the 'equivalent' (in terms of maximum real-time growth rate) to the $R_0$ =1.7 and $R_0$ =2.0 scenarios used in our work would be scenarios with $R_0$ =2.0 and $R_0$ =2.4 if historical natural history parameters were to be used. This increase in $R_0$ means that the predicted impact of controls is generally lessened (it requires more effort to control an epidemic with larger $R_0$).

The choice of natural history parameters is therefore extremely significant - while conclusions about the effectiveness of control measures are similar for the same $R_0$ value of 2.0, in one case this represents almost an absolute 'worst-case' 1918-19-like scenario, and in the other it represents a more moderate 1957-58 like scenario.

We would therefore encourage more analysis of primary data sources on influenza natural history parameters by epidemiological researchers working on influenza, rather than the continued uncritical reuse of parameters[28] for which the source data are unavailable.

**Proportion of school/workplace transmission**

As stated repeatedly in the main text, it is necessary to make assumptions about the proportion of transmission which occurs in schools and workplaces, as data do not exist to allow these parameters to be estimated. For GB and the US, our baseline assumption is that 37% of transmission occurs in these contexts, with the within-school transmission coefficient being twice that of the within-workplace coefficient. However, this choice is arbitrary.

We therefore explored the impact of doubling the within-place transmission coefficient while keeping the proportion of transmission occurring in the household fixed (at 29% for GB). This increased the proportion of transmission in schools and workplaces to 57%, and decreased the proportion of random community transmission from 33% to 13% (values comparable to those assumed by other authors [42,43]).

The impact of this change is marked. Overall infection attack rates for $R_0$=2.0 drop from 68% to 54% (due to mixing being much more focused on school-age children), with clinical attack rates assumed to be half of these (34% and 27% respectively). Peak daily attack rates are also reduced – from 1.9% to 1.45%. The effect of socially-targeted control policies is enhanced. While same-day treatment of 90% of clinical cases on its own reduces cumulative and peak daily clinical attack rates to 23% and 1.0% respectively, next-day treatment coupled with reactive school closure to the policy brings these values down to 18% and 0.6% - a far greater impact than school closure has with the baseline transmission parameters. Adding household prophylaxis brings attack rates down further, to 11% and 0.3%, respectively – a significantly greater reduction than for the same policy with default parameter values. Equally significantly, the treatment-only policy requires a 21% antiviral stockpile, and the household prophylaxis policy only increases this to 28%. Going still further and adding next day socially targeted prophylaxis (of 80% of close work colleagues or school class mates) brings cumulative and peak daily clinical attack rates down to 3% and 0.1% while keeping antiviral stockpile requirements (since the policy dramatically reduces transmission in the population) below 30% of population size.

We therefore think our default parameter choice is relatively conservative in assuming 1/3 of transmission will occur in the community, and therefore be largely untraceable (from the modelling perspective) and not able to be targeted except via very resource intensive geographically targeted policies.

**Proportion of infections which become clinical cases**

We have assumed 50% of infected individuals will be ill enough to potentially seek healthcare and thus be defined as a clinical case. These individuals are also assumed to be more infectious than non-cases, and to be more likely to be absent from school or work (see sections 2 and 3 above). Additionally, we assume 90% of cases would actually be diagnosed/detected by the healthcare system during a pandemic, meaning only 45% of infections are detected overall.

Other work [42,43] assumes 67% of infections become cases, but that 80% or fewer of these would be diagnosed – meaning 54% or less of infections are detected. Assuming a 67% case rate for our model increases baseline cumulative clinical attack rates from 34% for $R_0$=2.0 to 45%, and from 28% to 36% for $R_0$=1.7. However, the impact of case-targeted policies also increases.

For $R_0$=2.0, same day treatment of all detected cases can reduce clinical attack rates to 37%, but requires an antiviral stockpile of 34% of population size. If same-day treatment is not feasible, then next-day treatment coupled with reactive school closure gives the same impact on cumulative attack rates (and a larger reduction in the peak daily attack rate) for the same stockpile

size. Adding household prophylaxis gives an overall halving of clinical attack rates to 22% and requires an antiviral stockpile sufficient to treat 50% of the population. Socially targeted prophylaxis scenarios also have slightly increased efficacy and slightly decreased stockpile requirement (down to 82%). Similar differences are seen for other values of $R_0$.

Except in estimating the potential demand for antivirals in a treatment-only strategy, we therefore believe that our assumption of 50% of infections being clinical cases and thus able to be detected is moderately conservative in evaluating the likely impact and requirements of control policies which rely on case identification.

Conversely, assuming 67% of infections become cases *and* that a higher proportion of transmission occurs in locations which can be readily identified gives much more optimistic conclusions about the feasibility of certain control measures[43]. In particular, assuming 57% of transmission occurs in schools/workplaces (as in the last subsection) and a 67% case rate dramatically increases the predicted impact of prophylaxis targeted at both households and school classes/workgroups – with predicted cumulative clinical attack rates falling to 5% and stockpile requirements to 47% of population size (for $R_0$=2.0).

**Infectiousness of infected individuals with mild illness**

Our default assumption is that those ill enough to be defined as clinical cases are twice as infectious as those with milder disease. This assumption was motivated by parameter estimates derived from household study data[6]. Combined with our assumptions about the behaviour of symptomatic individuals (withdrawal to the home), this means the overall contribution of clinical cases to transmission is only ~40% more than that of non-cases.

Given the relative weakness of data on this important parameter, we evaluated the effect of varying the relative infectiousness of non-cases – to 25% of that of cases. Transmission coefficients in the household, school/workplace and community were adjusted to give the same overall proportion of transmission in these settings, and reduces the size of antiviral stockpile needed to implement prophylaxis polices (Figure SI15). As expected, this change increases the effectiveness of policies based on case identification (e.g. treatment, targeted prophylaxis), though the effect is only marked for the most intensive policy (socially targeted prophylaxis). We also examined the impact of additionally varying the proportion of infections which become clinical cases to 67% - which further increased policy effectiveness (Figure SI15).

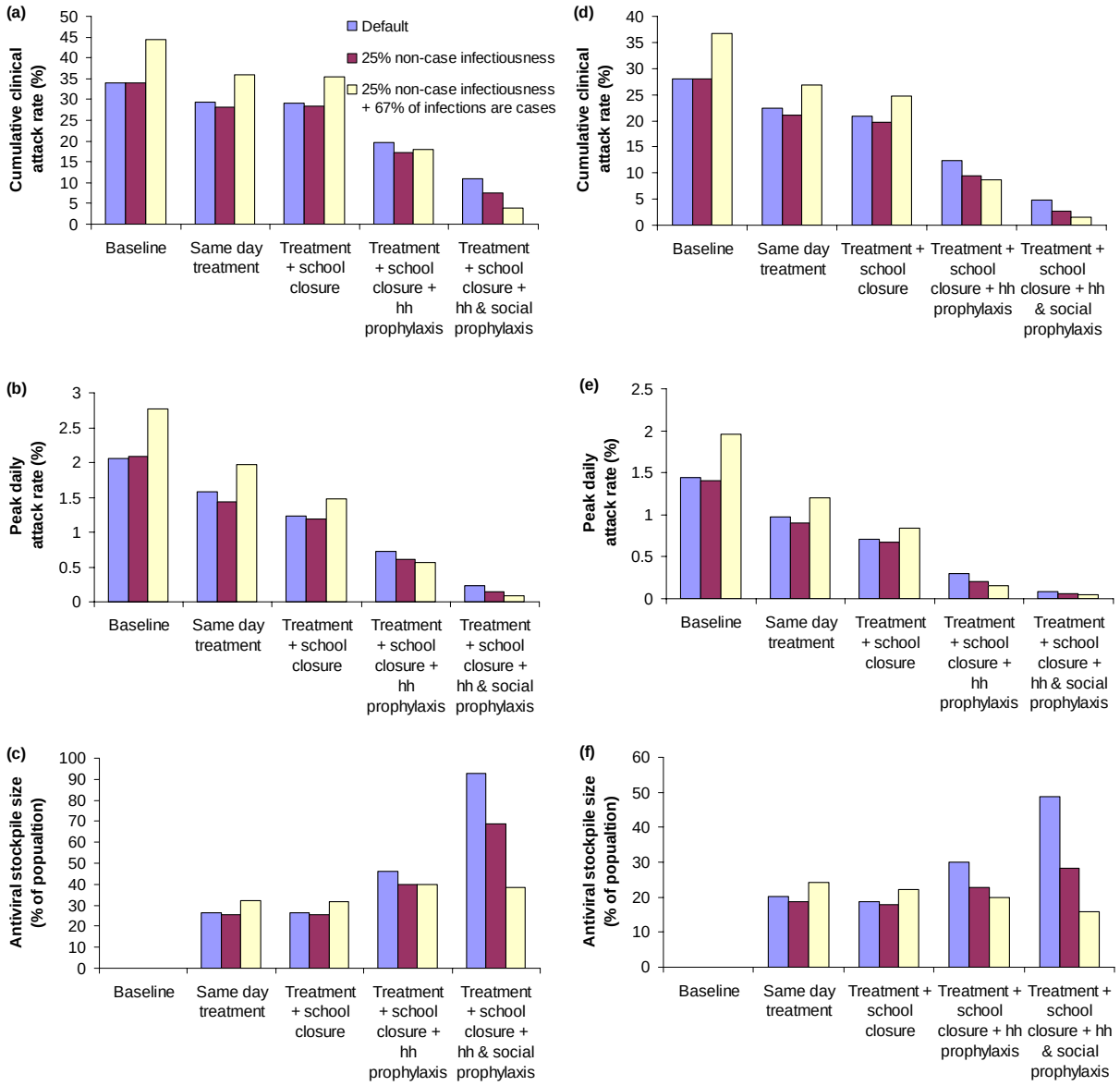


**Figure SI15:** Effect of varying assumptions about the infectiousness of cases versus non-cases. Results for an epidemic in the absence of interventions and for 4 intervention policies are shown. Same day treatment is assumed to be provided to 90% of cases on the same day symptoms start, while treatment coupled with prophylaxis is assumed to be given the day after symptoms start in the index case. Where applicable household prophylaxis is provided for all members of 90% of households of diagnosed cases the day after symptom onset in the case. School closure is assumed to be reactive, and to result in a school closing for 3 weeks after the first case is reported. Social prophylaxis is the prophylaxis of 90% of the school class mates or close work colleagues of 90% of cases, occurring the day after symptom onset in the case. Results are shown for three parameter sets: (i) the default of 50% of infections becoming cases, and non-cases being half as infectious as cases; (ii) 50% of infections becoming cases, and non-cases being 25% as infectious as cases; (iii) 67% of infections becoming cases, and non-cases being 25% as infectious as cases. (a) cumulative clinical attack rate for $R_0$=2, (b) peak daily clinical attack rate for $R_0$=2, (c) antiviral stockpile requirements for $R_0$=2, (d)-(f) as (a)-(c), but for moderate transmissibility ($R_0$=1.7) scenario, GB model results are shown.

## Behavior of symptomatic individuals

Our default assumption is that all children who are ill enough to be clinical cases will stay home from school, but that only 50% of symptomatic adults might stay home from work. Additionally, we assume that symptomatic individuals have a 50% reduced community contact rate. We have

found no data with which to verify these assumptions (particularly in a pandemic setting), but they appear conservative (in the sense of not over-estimating the extent of withdrawal to the home).

To check whether these assumptions significantly affected model projections, we undertook simulations assuming 75% of symptomatic adults are absent from work, and that community contact rates are reduced by 75% for all symptomatic individuals. Keeping $R_0$ fixed, the effect of these changes was small (less than a 1% increase in cumulative clinical attack rates), but slightly reduced the impact of control measures relying on case detection. This apparently paradoxical results arises because if one assumes symptomatic individuals reduce their contact rates more, then transmission from asymptomatic or mildly symptomatic individuals (who are assumed not to be detected) must increase to keep $R_0$ constant.

**Timing of start of mitigation policies**

Figure SI16 shows the effect of varying the number of cases that have to occur in the country before mitigation policies are started, excluding policies involving border controls or movement restrictions. Virtually no difference is seen between starting after 10 cases or 10,000 for either GB or the US (results are shown for the GB model).

The effectiveness of border controls and internal travel restrictions is clearly dependent on when such policies are implemented. In the case of border restrictions, no significant differences were seen in the effectiveness of border controls to reduce importation of infections at delaying a national epidemic as a function of whether they were started on day 0, 15 or 30 of the global epidemic. This was true for 90%, 99% and 99.9% reductions in importations of infections and for both the moderate and high transmissibility scenarios examined. However, the efficacy of border controls declines markedly for delays of 45 days or more in policy initiation, since at that point there is a high probability that infected individuals will have already entered the country.

Within country travel restrictions (which only have any significant effect in the US) show a similar dependency, though in that case the restriction is that the national epidemic needs to be small at the time such measures are imposed – so that optimal effectiveness is seen when such measures are imposed within 2 weeks of the first case in the US

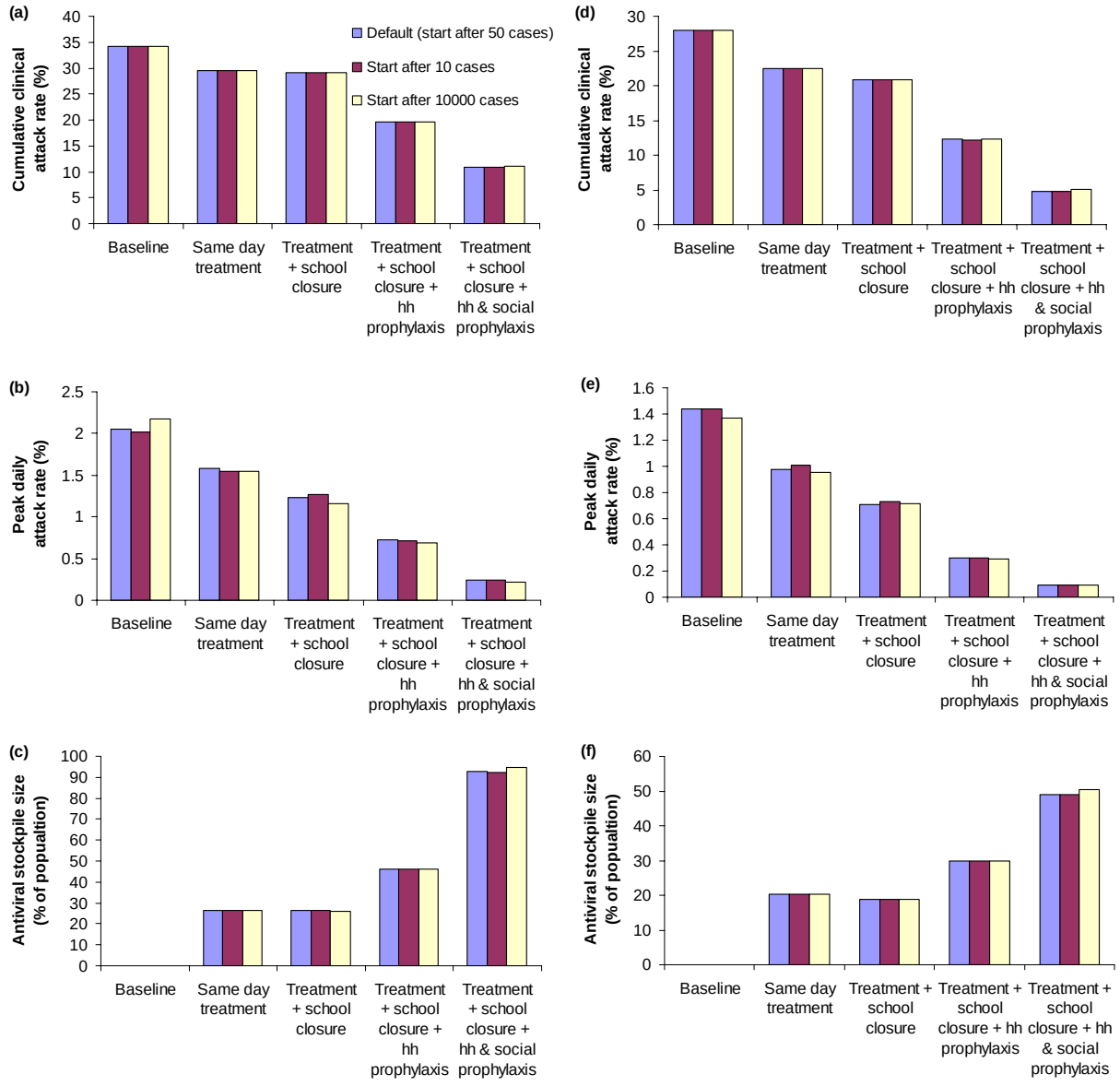


**Figure SI16:** Effect of varying assumptions about when mitigation measures start. As Figure SI15, except that results for starting policies after 50 cases (default), 10 cases and 10000 cases have been diagnosed in the country are shown.

## Effect of kernel shape on US pandemic dynamics

We investigated the effect of variations in kernel shape on the dynamics of a pandemic in the US. Within the range of kernel shapes shown in Figure SI4, almost no difference could be observed (Figure SI17). There were marginal differences in peak daily case incidence, which while not statistically significant appeared to suggest that the marginally fatter long-distance tail of the default US kernel favours slightly greater epidemic synchrony between different locations and therefore a slightly greater peak daily incidence. Control policies are equally largely unaffected by kernel choice in the range detailed by Figure SI4: the largest effect is on internal movement restrictions, but even here kernel choice affects timing of the peak of a pandemic by more than a week or so.

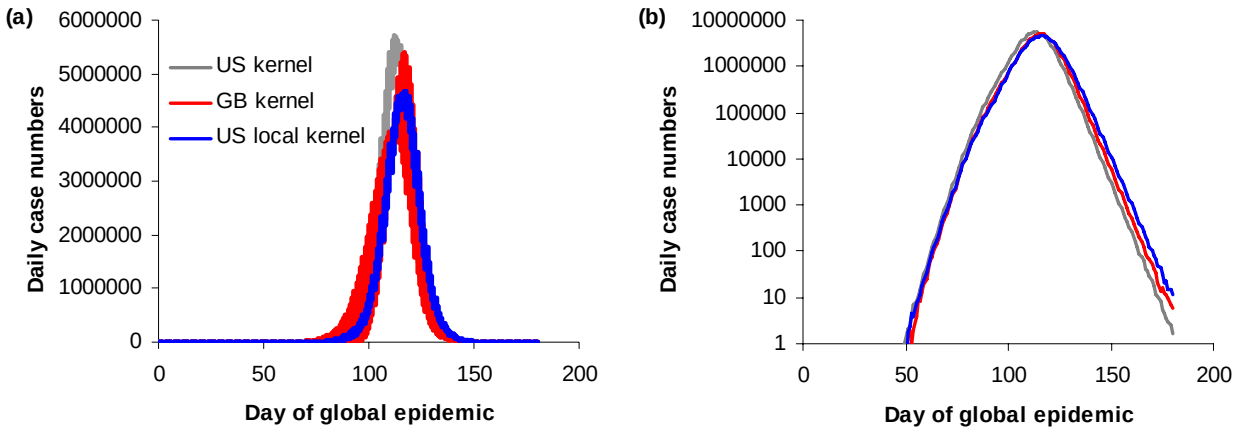


**Figure SI17:** Effect of kernel shape on baseline US pandemic dynamics. Three kernels are compared: the default US kernel derived from STP64 census data (see Figure SI4b), the default GB kernel (see Figure SI4a), and a 'localised' version of the US kernel (see Figure SI4c) (a) Epidemic curve for $R_0$=2.0 for US when the epidemic. Averages of 10 realizations are shown. The vertical thickness of the curve represents twice the standard deviation seen between runs in the incidence at that time point. (b) As (a) but with showing average incidence only (for clarity without error bars or an envelope) on a log scale. Real time growth rate of the epidemics is seen to be independent of kernel choice.

## 'Containment' policies applied in GB or the US

Figure SI18 shows the potential impact of applying the type of intensive transmission reduction policy proposed in previous work to contain a pandemic at source to limit spread in the UK. In principle such a policy is very effective: however, unlike containment at source, infection is constantly being reseeded in the country, resulting in very large stockpile requirements for the policy and immense logistical requirements in delivering the policy (for the high transmissibility scenario, over a 6 month period, every individual in the country would have been prophylaxed 5 times. Also, virus is never truly eliminated from the population – a national incidence of around 200 cases per day is maintained, even in the absence of reseeding.

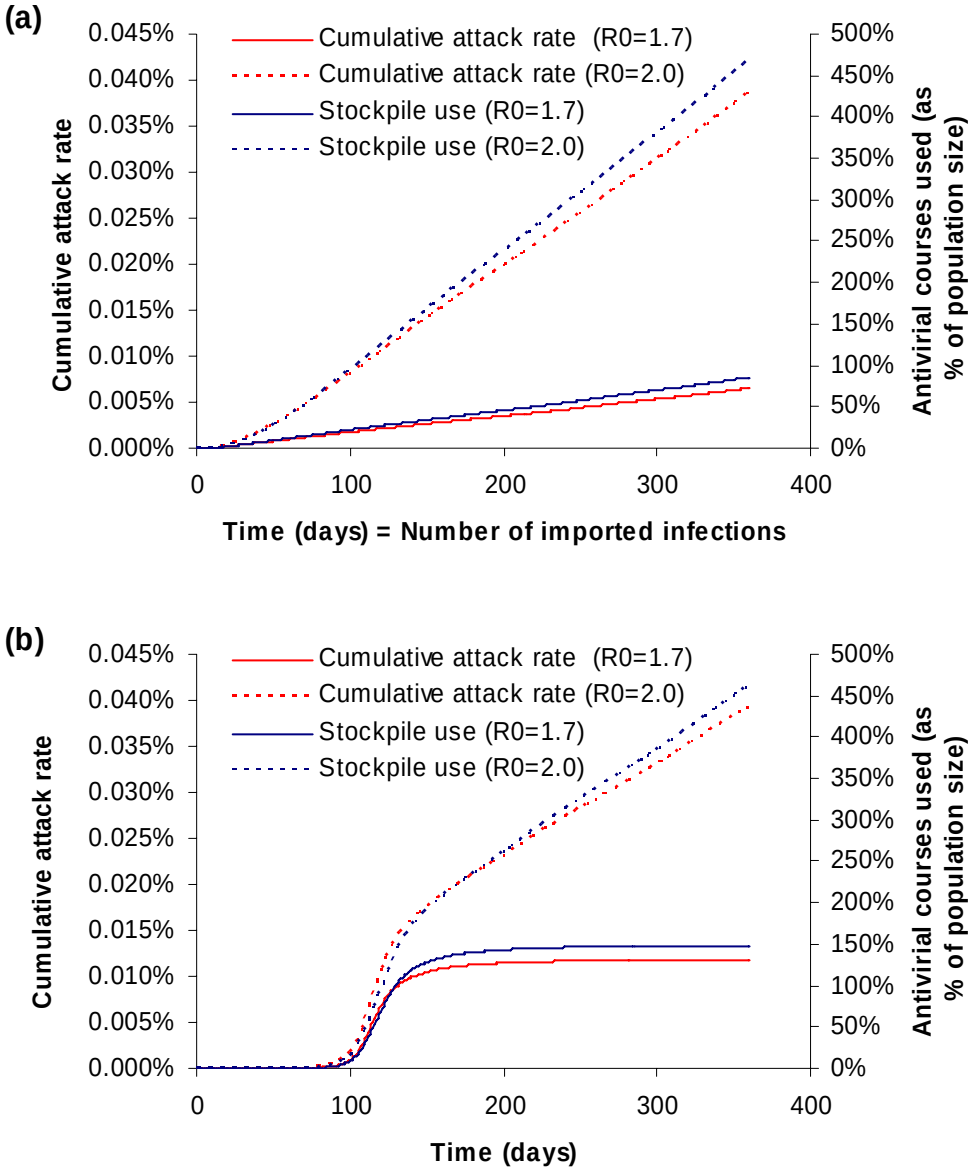


**Figure SI18:** Application in GB of a 'containment at source' strategy of case treatment, household prophylaxis, school closure, prophylaxis of school classes and close work colleagues, blanket population prophylaxis of the nearest 20,000 people to each case, and a 75% reduction in travel in and out of affected zones of 5km around each case[6]. (a) Effectiveness of policy (as quantified by the cumulative clinical attack rates and antiviral stockpile required) assuming 1 infected individual enters the country per day. Results for $R_0$=2.0 and $R_0$=1.7 are shown. (b) as (a), but for realistic seeding (see Figure SI7 for profile of infections per day entering GB).

## Internal travel restrictions

Figure SI19 shows the potential impact of blanket reductions on travel over 50km (for the duration of a pandemic) on pandemic spread in the US, as a function of the extent of reduction achieved. All policies include 99.9% effective border controls, since within country reductions on travel have very limited impact in the absence of border restrictions, due to infection being widely seeded solely from external imports. The effect of internal restrictions is largely to spread the epidemic over a much longer time period. The peak of the epidemic is delayed little, but the magnitude of the peak is substantially reduced for reductions in travel over 50km of at least 90%. Internal travel restrictions have a very limited impact on overall attack rates (no more than a 2%

reduction). The effectiveness of internal travel restrictions drops dramatically for less than 90% reductions – a 75% reduction gives results very similar to having no travel reductions at all (not shown).

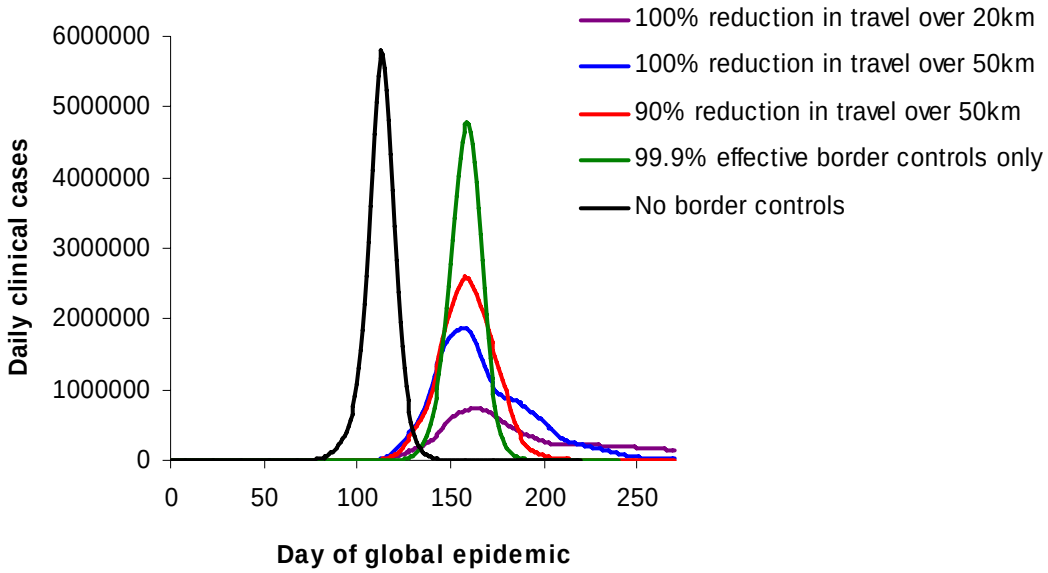


**Figure SI19:** Single example realisations of high transmissibility scenario US pandemic, showing effect of 99.9% border restrictions (relative to default scenario), and then addition of reductions in within-country travel involving journey distances of either 20km or 50km. Realisations shown are representative of overall qualitative impact of internal and external travel restrictions.

## Bibliography


1. Oakridge National Laboratory. Landscan global population data http://www.ornl.gov/sci/gist/landscan (2003).
2. US Bureau of the Census. U.S. Interim Population Projections http://www.census.gov/ipc/www/usinterimproj/ (2005).
3. Office for National Statistics (UK). Mid-2004 Population Estimates: UK; estimated resident population by age and sex http://www.statistics.gov.uk/statbase/Expodata/Spreadsheets/D9081.xls (2005).
4. Office for National Statistics (UK). Percentage of households by size http://www.statistics.gov.uk/STATBASE/Expodata/Spreadsheets/D6715.xls (2003).
5. US Bureau of the Census. Households by Size http://www.census.gov/population/socdemo/hh-fam/tabHH-4.pdf (2005).
6. Ferguson, N. M. et al. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* **437**, 209-14 (2005).
7. Department for Education and Skills. *Statistics of Education: Schools in England* (HMSO, 2004).
8. UK Department for Transport. Factors leading to increased school journey length http://www.dft.gov.uk/stellent/groups/dft_susttravel/documents/page/dft_susttravel_017877-01.hcsp (2003).
9. National Centre for Educational Statistics. Common Core of Data http://nces.ed.gov/ccd/bat/index.asp (2004).
10. US Environmental Protection Agency. Travel and Environmental Implications of School Siting http://www.epa.gov/smartgrowth/pdf/school_travel.pdf (2003).
11. US Bureau of Transportation Statistics. Transportation Statistics Annual Report: Table 27 http://www.bts.gov/publications/transportation_statistics_annual_report/2003/html/appendix_b/ (2003).
12. Axtell, R. L. Zipf Distribution of U.S. Firm Sizes. *Science* **293**, 1818-1820 (2001).
13. US Bureau of Labor Statistics. Quarterly Census of Employment and Wages: establishment size and employment http://www.bls.gov/cew/size2003.pdf (2003).
14. US Bureau of the Census. Census 2000: Census tract of work by census tract of residence (STP64) http://www.census.gov/mp/www/spectab/stp64-webpage.html (2004).
15. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462-5 (2006).
16. Viboud, C. et al. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. *Science* **Advanced online publication**, (www.sciencexpress.org) 10.1126/science.1125237 (2006).
17. US Bureau of Transportation Statistics. Airline Origin and Destination Survey (DB1B) http://transtats.bts.gov/Databases.asp?Mode_ID=1&Mode_Desc=Aviation&Subject_ID2=0 (2005).
18. US Bureau of Transportation Statistics. National Transportation Statistics:Table 1-39 http://www.bts.gov/publications/national_transportation_statistics/2005/html/table_01_39.html (2005).
19. Xia, Y., Bjornstad, O. N. & Grenfell, B. T. Measles metapopulation dynamics: a gravity model for epidemiological coupling and dynamics. *Am Nat* **164**, 267-81 (2004).
20. Ortuzar, J. & Willumsen, L. G. *Modelling Transport: Third Edition* (Wiley, Chichester, 2001).
21. Travel Industry Association of America. Tourism Facts. (2005).
22. Rvachev, L. A. & Longini, I. M. A Mathematical-Model for the Global Spread of Influenza. *Mathematical Biosciences* **75**, 3-23 (1985).
23. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans; Dynamics and Control* (Oxford University Press, Oxford, 1991).
24. Hufnagel, L., Brockmann, D. & Geisel, T. Forecast and control of epidemics in a globalized world. *Proc Natl Acad Sci U S A* **101**, 15124-9 (2004).
25. Office for National Statistics (UK). Travel Trends 2004 www.statistics.gov.uk/downloads/theme_transport/traveltrends2004.pdf (London, 2005).
26. US Office of Travel and Tourism Industries. International Arrivals to U.S. - Historical Visitation 1997-2004 http://tinet.ita.doc.gov/view/f-2004-04-001/index.html (2005).
27. Yang, Y. & Longini, I. M. Technical Report 04-09: Design and evaluation of prophylactic interventions using infectious disease incidence data from close contact groups. http://www.sph.emory.edu/bios/tech/Tech_Report_04-09.pdf (Department of Biostatistics, Emory University, 2004).

28. Elveback, L. R. et al. An influenza simulation model for immunization studies. *Am J Epidemiol* **103**, 152-165 (1976).
29. Moser, M. R. et al. An outbreak of influenza aboard a commercial airliner. *American Journal of Epidemiology* **110**, 1-6 (1979).
30. Hayden, F. G. et al. Use of the oral neuraminidase inhibitor Oseltamivir in experimental human influenza. *JAMA* **282**, 1240-1246 (1999).
31. Cauchemez, S., Carrat, F., Viboud, C., Valleron, A. J. & Boelle, P. Y. A Bayesian MCMC approach to study transmission of influenza: application to household longitudinal data. *Statistics in Medicine* **23**, 3469-3487 (2004).
32. Longini, I. M., Jr., Koopman, J. S., Haber, M. & Cotsonis, G. A. Statistical inference for infectious diseases. Risk-specific household and community transmission parameters. *Am J Epidemiol* **128**, 845-59 (1988).
33. Taubenberger, J. K. et al. Characterization of the 1918 influenza virus polymerase genes. *Nature* **437**, 889-93 (2005).
34. Diekmann, O. & Heesterbeck, J. A. P. *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation* (ed. Levin, S.) (John Wiley & sons, New York, 2000).
35. HMSO. Forty-Eighth Annual Report of the Local Government Board, 1918-19. Supplement Containing the Report of the Medical Officer for 1918-19. 24-27 (London, 1919).
36. Mills, C. E., Robins, J. M. & Lipsitch, M. Transmissibility of 1918 pandemic influenza. *Nature* **432**, 904-6 (2004).
37. UK Ministry of Health. The Influenza epidemic in England and Wales 1957-58 (London, 1960).
38. Fraser, C. & Ferguson, N. M. Explaining variation in transmissibility in 20th Century influenza pandemics. *In preparation* (2005).
39. Barry, J. M. *The Great Influenza: The Epic Story of the Deadliest Plague in History* (Penguin, 2005).
40. Chotpitayasunondh, T. et al. Human disease from influenza A (H5N1), Thailand, 2004. *Emerging Infectious Diseases* **11**, 201-209 (2005).
41. The Writing Committee of the WHO Consultation on Human Influenza A/H5. Current concepts - Avian influenza A (H5N1) infection in humans. *New England Journal of Medicine* **353**, 1374-1385 (2005).
42. Longini, I. M., Jr., Halloran, M. E., Nizam, A. & Yang, Y. Containing pandemic influenza with antiviral agents. *Am J Epidemiol* **159**, 623-33 (2004).
43. Longini, I. M., Jr. et al. Containing pandemic influenza at the source. *Science* **309**, 1083-7 (2005).