Strategies for antiviral stockpiling for future influenza pandemics: a global epidemic-economic perspective

Luis R. Carrasco1, Vernon J. Lee2,3,4,5, Mark I. Chen3,6,7, David B. Matchar8,9,10, James P. Thompson10 and Alex R. Cook1,*

1Department of Statistics and Applied Probability, 2Centre for Health Services Research, and 3Department of Epidemiology and Public Health, National University of Singapore, Singapore 117543, Republic of Singapore
4Biodefence Centre, Ministry of Defence, Singapore
5National Centre for Epidemiology and Population Health, Australian National University, Canberra, ACT, Australia
6Communicable Disease Centre, Tan Tock Seng Hospital, Singapore
7Program in Emerging Infectious Diseases, Duke-NUS Graduate Medical School, Singapore
8Duke Center for Clinical Health Policy Research, Duke Clinical Research Institute, and 9Durham US Department of Veterans Affairs Medical Center, Duke University, Durham, NC, USA
10Program in Health Services and Systems Research, Duke-NUS Graduate Medical School, Singapore

Influenza pandemics present a global threat owing to their potential mortality and substantial economic impacts. Stockpiling antiviral drugs to manage a pandemic is an effective strategy to offset their negative impacts; however, little is known about the adequate long-term size of stockpiles under uncertainty, vaccine availability and the characteristics of different countries, leading to the risk of misallocation of limited public health resources [3].

*Author for correspondence (stacar@nus.edu.sg).


1. INTRODUCTION

Influenza pandemics have occurred over the past few centuries at intervals of between 10 and 40 years [1] causing high morbidity and mortality and enormous economic impacts. Many countries practised stockpiling on antiviral drugs (henceforth, ‘antivirals’) such as oseltamivir (Tamiflu) and zanamivir (Relenza) before the emergence of influenza A (H1N1-2009) as an essential strategy for pandemic response before a new vaccine could be widely distributed [2]; however, we know little about the adequate long-term size of stockpiles under uncertainty, vaccine availability and the characteristics of different countries, leading to the risk of misallocation of limited public health resources [3].

The effect of management strategies like prophylaxis and treatment with antivirals, vaccination and school closures has been studied using complex dynamic models [4–8]. However, such models have infrequently been used to evaluate the cost-effectiveness of management strategies. In addition, the computation time required for such complex models complicates the implementation of model optimization to obtain economically optimal management strategies that account for the stochasticity of the modelled systems. With a considerably lower spatial and temporal resolution,
there are very few insightful studies on the assessment of the cost-effectiveness of antiviral stockpile use [9–12]. These studies are normally non-dynamic or limit the scope to a single pandemic wave; with the economic impacts for longer time horizons estimated by weighting the obtained costs with the probability of a pandemic occurring. While useful, this approach presents some limitations: (i) it might not represent the distribution of potential pandemic combinations, e.g. several pandemics might occur in a short period of time and each pandemic can involve several waves; (ii) stochastic optimization is usually not performed, i.e. estimation of the most adequate stockpile size given the variability in the system; (iii) the development and availability of vaccines are usually not considered; and (iv) it is difficult to account for dynamic logistic problems (e.g. stockpile replenishment). Given the complexity of the epidemic and economic interactions of the system, inter-disciplinary epidemic–economic dynamic models would be necessary to advance the state of knowledge on the economics of antivirals stockpiling [3,13].

We considered the problem of estimating the antiviral stockpile size and the proportion of susceptible and infected individuals that should be managed by antiviral prophylaxis and treatment with the objective of minimizing the expected value of: (i) the net present value of total costs; (ii) number of fatalities; and (iii) costs per quality-adjusted life year (QALY) given the uncertainty of the severity and frequency of future pandemics. A range of countries, spanning different sizes and stages of development, was considered—Brazil, China, Guatemala, India, Indonesia, New Zealand, Singapore, the UK, the USA and Zimbabwe—thereby allowing a global perspective of the potential health and economic impacts of antiviral stockpiling.

2. METHODS

2.1. Overview

We adopted a simulation approach in which the emergence and spread of pandemics over a time horizon occur at random and are modelled with a compartmental epidemic model. The results are then linked to an economic sub-model that quantifies the costs owing to the pandemics. Given that the occurrence of future pandemics is very uncertain, our approach consists of repeating the same simulation experiment a large number of times (using Monte Carlo simulation) to gain insights on the distribution of the potential outcomes. Furthermore, by embedding control by antivirals and vaccines into the epidemic model (appendix A) we can evaluate the distributions of potential results for different levels of antiviral stockpile sizes. Comparing the evolution of the distributions of fatalities and total costs obtained in this way, we gain insight into adequate stockpile sizes for each country.

2.2. The model

2.2.1. Epidemic sub-model. We developed a dynamic, hybrid deterministic–stochastic simulation model, with a time horizon of 30 years and a time step of 1 day. We modelled the occurrence of a pandemic in the time horizon as a Poisson stochastic process. Once a pandemic starts, we modelled it using a deterministic susceptible–latent–infected–asymptomatic–recovered (SLIAR) model [14,15] (see appendix A for the mathematical description of the model). The SLIAR model divides the population into compartments that represent the number of individuals at each state in each time step. The effects of the antivirals used for prophylaxis and treatment, the proportions of individuals treated (appendix A) and the population growth rates for each country (table S1.2 in electronic supplementary material, S1) were incorporated in the model.

The basic reproduction number, $R_0$, is the average number of secondary cases from an infected individual in an otherwise susceptible population, and encapsulates the eventual magnitude of an epidemic [14–16]. We considered the $R_0$ of each pandemic to be stochastic and to follow an approximate uniform distribution between 1.4 and 3.9, a range encompassing those of past observed pandemics in 1918–1919, 1957–1958, 1968–1969 and 2009 [17,18]. Keeping the infectious period fixed, we derived the transmission probability from $R_0$ (appendix A). In this way the number, time of occurrence and conditions of each pandemic were stochastic, though once started, the epidemic itself was treated as deterministic.

For countries in temperate regions, each pandemic can consist of several waves owing to climatic seasonality (increased transmission rate in winter) or school terms and holidays. We modelled this seasonality through a yearly sinusoidal oscillation of the transmission rate [19]. A cosine function where the magnitude of seasonal oscillation was parametrized for the H1N1 2009 pandemic was used [20]. No seasonality owing to climate was assumed in tropical countries [21].

We assumed that six to eight months after the beginning of each pandemic an effective vaccine is developed, based upon the times of vaccine development and availability observed for the H1N1 2009 pandemic [22], although in the future this lead time might be shortened.

2.3. Impacts

The proportion of infected individuals leading to fatalities was modelled using a uniform distribution that encompassed the proportions observed in the 1918–1919, 1957–1958, 1968–1969 [12] and 2009 pandemics in the UK. This distribution was used as a baseline distribution that was shifted to reflect different countries characteristics using estimates of the case fatality rate per country that was elicited from vital registration data from the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23]. QALY’s lost owing to death were calculated using the case fatality rates of the 1918–1919 pandemic for all countries [23].

The direct economic impacts owing to the pandemics were: costs of general practitioner consultation, treatment for complications and hospitalization costs (see
For stockpile sizes greater than 15 per cent of the total population, the reduction of fatal casualties tended asymptotically to a constant level of expected number of fatalities that could not be reduced further using only antivirals. Stockpiling beyond this 15 per cent target led to an expected inefficient surplus of antivirals that are not used. The minimum at stockpile sizes of 15 per cent was consistent for both the mean and the 95th percentile, indicating that a relatively small stockpile was also protective for severe projections. If misallocation and misdiagnosis were both 30 per cent, the new stockpile size, for which the asymptotic constant level of fatal casualties will occur, would be for stockpile sizes of 25 per cent, again, fairly constant across countries of different levels of economic development (electronic supplementary material, S2), although the expected level of mortality would be higher, especially in resource-limited countries (electronic supplementary material, S2).

3.2. Economically optimal stockpile size

For small stockpile sizes, the costs owing to job absenteeism and hospitalizations dominated until reaching a minimum at which costs increased because of wastage of the growing unused stockpile, resulting in a J-shaped curve (figure 1d–f, figures S2.1d–f and S2.2e–h in electronic supplementary material, S2). The optimal stockpile size to minimize the expected cost per QALY gained was very similar to the optimal size that minimized the net present value of the total costs and the point at which the number of fatal casualties as a function of stockpile size tended to an asymptotic level (figure 1a–f, and figures S2.1 and S2.2 in electronic supplementary material, S2).

Using the cost-effectiveness criterion, we compare the costs per QALY gained with a threshold of thrice the gross national income per capita of the country (figure 1g–i and figures S2.1g–i and S2.2i–l in electronic supplementary material, S2) [25]. According to this criterion, stockpiling was clearly cost-effective for the higher resourced countries like the USA, the UK, New Zealand and Singapore (figure 2). Under the lower pricing scheme for resource-limited countries [26], stockpiling was cost-effective for Brazil and only marginally so for China; it was however non-cost-effective for Guatemala, India, Indonesia and Zimbabwe (figure 1g–i and figures S2.1g–i and S2.2i–l in electronic supplementary material, S2).

Since antiviral stockpiling is non-cost-effective for Guatemala at current prices, it seems reasonable to assume that for countries with lower per capita Gross National Income than Guatemala antiviral stockpiling will be non-cost-effective. In this case, around 46 per cent of the world population lives in countries where antiviral stockpiling is non-cost-effective. Including China in this group (since stockpiling is only marginally cost-effective there) causes this figure to rise to around 66 per cent. Were generic antivirals not protected by a patent and available at a price below $2 per course, purchasing costs would drop to the level at which stockpiling would become cost-effective for China, Indonesia and India (figure 2), though it would still
be cost-ineffective in Zimbabwe, and by extension, least resource-limited countries.

3.3. Proportion of antivirals for prophylaxis and treatment and vaccines availability

Both the minimum expected number of fatalities and the minimum level of total costs occurred consistently when the antiviral stockpile was used exclusively for treatment and not for untargeted prophylaxis (see figures S2.3 and S2.4 in electronic supplementary material, S2). Antivirals used for prophylaxis had a very limited mitigating effect on the pandemic but had a high opportunity cost, i.e. stocks that are useful for treatment are better used for this purpose than prophylaxis, and once sufficient stocks for treatment are reached, increasing the stock for prophylaxis use led to costs outweighing benefits.

A mean delay in vaccine availability for subsequent waves of the pandemics from 240 to 350 days did not have a noticeable effect on the total number of fatalities (compare figures S2.5–2.7 with S2.12–2.14 in electronic supplementary material, S2). This result represents a lower bound of vaccine effectiveness as we did not consider, in the model, waning of immunity in recovered individuals and the antigenic drift in the pandemic virus strain.

4. DISCUSSION

The results demonstrated that stockpiles—covering 15 per cent of the population for perfect allocation and 25 per cent for 30 per cent misallocation—attain considerable reductions of both mortality and total costs in all countries we considered, suggesting that the pattern holds for countries with varied sizes, influenza seasonality and stages of economic development.

The results also showed that, for all the countries considered, it was preferable not to use antivirals for prophylaxis of susceptible individuals but to reserve them for treatment of infected individuals, since, if antivirals are scarce, their use for prophylaxis implies a
large opportunity cost because they could be more effectively deployed as treatment for infected individuals. To some extent, this result can be attributed to the epidemic model used, as the assumption of homogeneous mixing prohibits targeted prophylaxis, which may be feasible via contact-tracing at an early stage of the outbreak [5–7] and has been demonstrated to be effective in semi-closed populations such as army camps [27]. However, on the whole, our results have clear implications: especially in the case of advanced epidemics where contacts with exposed individuals cannot be fully traced back, the use of antivirals for prophylaxis for the general public would represent an inefficient allocation of valuable resources and paradoxically lead to an increase of the number of fatalities if an efficient stockpile size had been put in place.

The administration of antivirals is not always achievable within 1 day of symptom onset and not all influenza-like-illness cases are due to the pandemic strain leading to some potential misallocation of antivirals. The proportion of misallocation of antivirals will depend on the country’s health and antiviral distribution systems. Apportionment for greater stockpile sizes (increase from 15 to 25%) will partially compensate for misallocations and misdiagnosis; however this does not prevent all of the resulting increase in mortality.

This study presents some limitations and could be extended in several ways. Firstly, we decided not to incorporate the possibility of emergence of antiviral-resistant strains because, although they have been observed to emerge in patients treated with oseltamivir [28], the fitness costs of resistance involved a reduction in infectiousness of the resistant strains such that they may not be capable of outcompeting wild strains [19]. Secondly, the analysis was based on a generalization from a sample of the three observed historical pandemics in the twentieth century and also one from the twenty-first century. The approach we took was to treat the severe 1918–1919 pandemic as an upper bound for pandemic severity, but it is very possible that pandemics having characteristics beyond the observed range might occur, black-swan-style i.e. failure of past data to provide sufficient information for adequate future predictions [29,30], so that the sample distribution of their characteristics underestimates the uncertainty in the true distribution. Thirdly, certain population age groups or medical conditions present higher risk to pandemic influenza. Owing to the large scale of the study and because case fatality rates per age group vary for each pandemic, we preferred to model all the susceptible individuals as homogeneous. In reality, as the pandemic unfolds, high-risk groups can be identified (e.g. elderly in the 1968–1969 pandemic) and antivirals can be preferentially allocated accordingly, leading to a higher number of avoided fatalities per antiviral course used. By contrast, because we are treating all the individuals as homogeneous, we are implicitly assuming that information on high-risk groups is not available. Therefore, our results of optimal antiviral stockpile sizes are conservative estimates since smaller stockpile sizes would attain the same number of fatalities reductions if the high-risk age groups could be identified.

A global international comparison of antiviral stockpiling cost-effectiveness demands the challenging consideration of the epidemic and socioeconomic heterogeneities between resourced and less-resourced countries. Owing to limitations in health systems resources, less-resourced countries are more vulnerable to pandemics and present higher mortality rates [23]. Despite the high number of potential fatalities that could be avoided, stockpiling antivirals in resource-limited countries implies that valuable resources are not allocated to ongoing severe health problems like measles, malaria or acquired immunodeficiency syndrome. The differences in the priorities between countries need to be reflected by using different discount rates which, in turn, can be surrounded by intense debate [31]. In addition, whereas resource-limited countries risk higher mortality rates, resourced countries have a higher stake on economic losses derived from future pandemics. For instance, resourced countries have a more highly specialized labour, which is more difficult to replace during a pandemic leading to very high productivity losses.

Given the high connectedness of today’s world by international air travel, the interventions most likely to be effective at delaying the spread of a pandemic are those based on local control with antivirals [27,32]. Hence, it makes economic and ethical sense, both from a global perspective and from the perspective of resourced countries, for resourced-limited countries to be able to have antiviral stockpiles. We estimated, however, that two-thirds of the world’s population live in countries where antiviral stockpiling is not cost-effective. The use of generic antivirals could make stockpiling cost-effective for China, Indonesia and India, enabling large stockpiles to be developed at low cost. Antivirals would still not be cost-effective for countries like Zimbabwe, highlighting the need for international cooperation. This is a particularly relevant finding given that both Tamiflu and Relenza go off-patent in the next five years (2016 and 2013, respectively), especially since lower antiviral prices are necessary to make their life-saving potential cost-effective in the most resource-limited countries. We conclude that generic antivirals, international cooperation and a global network of national antiviral stockpiles will provide for a very high avoidance of economic impacts and fatalities owing to future pandemics worldwide.

We are thankful for the valuable comments of two anonymous reviewers and the editor. L. R. Carrasco and A. R. Cook are thankful for research funding from the National University of Singapore.

APPENDIX A

A.1. Epidemic model

We considered the compartments susceptible (S), latent (E), infectious asymptomatic (A), infectious symptomatic (I), recovered (R) and dead (D). The parameters of the model are: the transmission probability between infectious symptomatic ($\beta$) and infectious asymptomatic
The latent period (\(\theta_{\text{sym}}\)) and susceptible; the population size (\(N\)); the proportion of infectious symptomatic individuals treated with antivirals (\(\theta_{\text{pro}}\)), the proportion of susceptible individuals administered antivirals for prophylaxis (\(\theta_{\text{pro}}\)), the latent period (\(\mu\)), the proportion of latent individuals becoming asymptomatic (\(\theta_{\text{asym}}\)), the infectious period of infectious symptomatic (\(\epsilon\)) and asymptomatic (\(\epsilon_{\text{asym}}\)) individuals; the rate of fatal casualties owing to the flu (\(\alpha\)); \(d_1\), \(d_2\), \(d_3\), and \(d_4\) are the effectiveness of the antivirals in reducing the infectious period, the death rate, the transmission rate from infectious non-treated to susceptible treated and from infectious treated to susceptible with no prophylaxis, respectively. The model is expressed by a set of differential equations:

\[
\frac{dS}{dt} = -\Omega
\]

\[
\Omega = \frac{\rho}{N}(1 - \theta_{\text{pro}})I_t(1 - \theta_{\text{pro}})S_t
\]

\[
+ \frac{\rho(1 - d_1)}{N} \theta_{\text{tre}} I_t(1 - \theta_{\text{pro}})S_t
\]

\[
+ \frac{\rho(1 - d_1)}{N} (1 - \theta_{\text{sym}})I_t \theta_{\text{pro}} S_t
\]

\[
+ \frac{\rho(1 - d_3)}{N} (1 - \theta_{\text{sym}})I_t \theta_{\text{pro}} S_t
\]

\[
+ \frac{\rho(1 - d_4)}{N} \theta_{\text{asym}} I_t \theta_{\text{asym}} S_t + \frac{\rho_A}{N}
\]

\[
\times A_t (1 - \theta_{\text{pro}})S_t
\]

\[
+ \frac{\rho_A (1 - \theta_{\text{pro}})}{N} A_t \theta_{\text{pro}} S_t.
\]  

Expression (A 1) indicates how the number of susceptibles decreases over time owing to contacts between susceptible individuals not undergoing prophylaxis with untreated infectious symptomatic individuals (first element of \(\Omega\)) and infectious symptomatic individuals treated with antivirals (second element of \(\Omega\)); the contact between susceptible individuals undergoing prophylaxis with untreated symptomatic infectious individuals (third element of \(\Omega\)) and treated symptomatic infectious individuals (fourth element of \(\Omega\)); and the contact between susceptible individuals not undergoing prophylaxis (fifth element of \(\Omega\)) and undergoing prophylaxis (sixth element of \(\Omega\)) with infectious asymptomatic individuals.

\[
\frac{dE}{dt} = \Omega - (1 - \theta_{\text{sym}})\mu^{-1}E_t - \theta_{\text{sym}}\mu^{-1}E_t.
\]  

Expression (A 2) reflects how the number of latent individuals (non infectious) increases as a result of the contacts described above (\(\Omega\)) and decreases because of latent individuals reaching the end of the incubation period and becoming either symptomatic infectious (third term in expression (A 2)) or asymptomatic infectious (fourth term in expression (A 2)).

\[
\frac{dA_t}{dt} = \theta_{\text{asym}}\mu^{-1}E_t - \epsilon_{\text{asym}}^{-1}A_t.
\]  

The dynamics of the number of infectious asymptomatic (A 3) depends on the latent individuals becoming new infectious asymptomatic (second element in equation (A 3)) and the reduction of asymptomatic individuals as they reach the end of their infectious period (third element in equation (A 3)).

\[
\frac{dI_t}{dt} = (1 - \theta_{\text{sym}})\mu^{-1}E_t - \epsilon^{-1}(1 - \theta_{\text{tre}})I_t
\]

\[
- \epsilon^{-1}(1 - d_1)^{-1}\theta_{\text{tre}} I_t - \alpha(1 - \theta_{\text{tre}})I_t
\]

\[
- \alpha(1 - d_2)\theta_{\text{tre}} I_t.
\]

The dynamics of the number of infectious symptomatic individuals described by equation (A 4) depends on the number of latent individuals that become infectious symptomatic (second term in equation (A 4)) and the proportions of infectious that get recovered. Recovery rates will vary depending on the proportion of individuals untreated with antivirals (third element in equation (A 4)) and treated with antivirals (fourth element in equation (A 4)). A proportion of infectious symptomatic individuals die with two different rates: the first corresponding to infectious individuals untreated (fifth element in equation (A 4)) and treated with antivirals (sixth element in equation (A 4)).

\[
\frac{dR_t}{dt} = \epsilon^{-1}(1 - \theta_{\text{tre}})I_t + \epsilon^{-1}(1 - d_1)^{-1}\theta_{\text{tre}} I_t + \epsilon^{-1}_{\text{asym}} A_t.
\]  

Expression (A 5) describes the increase of individuals recovered from infectious symptomatic untreated (second element in expression (A 5)) and treated with antivirals (third element in expression (A 5)) and infectious asymptomatic individuals (fourth element in expression (A 5)).

\[
\frac{dD_t}{dt} = \alpha(1 - \theta_{\text{tre}})I_t + \alpha(1 - d_2)\theta_{\text{tre}} I_t.
\]  

The number of fatal casualties in (A 6) is derived from the fifth and sixth elements in equation (A 4).
was obtained as a function of $R_0$. The expression of $R_0$ was obtained from the positive real eigenvalue of the next generation matrix [15] for the case of the continuous SLIAR model without control measures:

$$R_0 = \frac{\kappa}{\mu} + \varepsilon (1 - \theta_{\text{asymp}}) + \varepsilon_A (\rho_A \theta_{\text{asymp}}), \quad (A.7)$$

where $\kappa = 0$ is the relative infectiousness of latent with respect to asymptomatic infectious individuals and $\rho_A = 2 \mu$ [19].

After a period of time $T$ a vaccine of effectiveness $\gamma$ against the pandemic strain is assumed to become available. A proportion ($\gamma_{\text{vacc}}$) of the susceptible individuals is inoculated. The inoculated individuals are transferred to the recovered state after the vaccination according to:

$$S(T^+) = (1 - \gamma_{\text{vacc}}) S(T^-),$$

where $T^-$ and $T^+$ indicate the times before and after the application of the vaccine.

REFERENCES