# An approach to estimate the number of SARS cases imported by international air travel

A. GOUBAR[1], D. BITAR[1]*, W. C. CAO[2], D. FENG[2], L. Q. FANG[2]
AND J. C. DESENCLOS[1]

[1] *Institut de Veille Sanitaire, Infectious Diseases Department, Saint Maurice, France*
[2] *Beijing Institute of Microbiology and Epidemiology, State Key Laboratory of Pathogen and Biosecurity, Feng Tai District, Beijing, P.R. China*

## SUMMARY

The worldwide spread of severe acute respiratory syndrome (SARS) raised questions about the risk of importation of such infection, in particular by air travel. Entry screening was implemented in some countries although poor evidence on its effectiveness is reported. We developed a model to estimate the number of imported SARS cases between regions, using the 2003 SARS epidemic data to apply this model for two scenarios: from Beijing to Frankfurt and from Hong Kong to London. We back-calculated the data to estimate individuals' time of infection and built a model where every individual has a probability of being isolated, of traveling, and of being undetected at arrival. The findings, consistent with what was observed in 2003, suggest that entry screening does not affect the predicted number of imported cases. Inversely, importation depends on the transmission dynamic in the country of origin (including control measures in place) and on the intensity of air travel between regions.

**Key words**: Entry screening, importation, international travel, model.

## INTRODUCTION

Evidence from the severe acute respiratory syndrome (SARS) outbreak illustrated the rapid spread of an emerging infection from one region to another by infected travellers. Thus, the risk of introduction of an infectious pathogen in a community must be assessed, in order to provide governmental authorities with timely information to help in the adoption of appropriate responses. From a risk management point of view, characterization of potentially infected travellers and prediction of their numbers is important.

Travel is a key component for importation to occur. This parameter was used in modelling works that focused on, for example, predicting the worldwide spread of influenza through the air transportation network [1–3]. A more recent study focused on the effect of the heterogeneity of travel frequency in the population on the global spread of infectious agents [4]. These studies, based on 'susceptible–infectious–recovered' or 'susceptible–exposed–infectious–recovered' epidemic models structured per region, used air-travel volume data to estimate travel rate and to model the movement of individuals between regions. The number of travellers was also used by Glass *et al.* [5] to study how control measures in a source region, and entry screening and public health responses in the destination region can impact the

* Author for correspondence: Dr D. Bitar, Institut de Veille Sanitaire, Infectious Diseases Department, 12 Rue du Val d'Osne 94415 (F), Saint Maurice Cedex, France.
(Email: d.bitar@invs.sante.fr)

probability and the size of a SARS outbreak. These authors estimated *the probability that an infected traveller arrives undetected* as a function that depends on the prevalence of infection, the delay to isolation, and the sensitivity of entry screening. This probability could be used to estimate the *number of imported cases* given the number of travellers. These authors applied the population's SARS prevalence to travellers heading to a destination, assuming it to be constant. However, prevalence among travellers might differ from that in the population. In addition, there are obviously some population categories that are more likely to travel than others. It is possible that within these categories, some are at a higher risk of infection than others and thus the prevalence might be different between subgroups of travellers. Health-care workers (HCWs) represent a particular group frequently exposed to in-hospital infections and susceptible to importation of infectious pathogens such as SARS [6, 7]. The extent of their role in the transmission and spread of influenza, for instance, has been recently analysed [8].

In the present paper we estimated the *cumulated number of imported SARS cases* from one source region to a destination since the epidemic began, until a day, $d$, during an outbreak, avoiding any potential bias linked to a possible heterogeneity in the prevalence between population groups. The approach we use required two steps. First, instead of using the prevalence in the source region, we used the distribution of infected individuals by time of infection, back-calculated from the incident cases. In a second step, we performed a simple probabilistic calculation via Monte Carlo simulations, where every infected individual in the source region has a chance to travel and a probability of being undetected by entry screening in the destination region. Our aim was to estimate the number of imported SARS cases, validating this calculation using the 2003 SARS epidemic data and exploring a situation where we assumed that infected HCWs are more likely to travel than other infected individuals. We also investigated whether or not importation of SARS is affected by specific entry screening.

## METHODS

We used data from the 2003 SARS epidemic to estimate the cumulated number of imported infected individuals in two scenarios: Beijing–Frankfurt and Hong Kong–London. Estimates are obtained by time interval $T_d = (0, d)$, during a number of $d$ days since symptom onset of the first case in the source region. Both selected destinations are among the busiest European airports.

The risk of importation of SARS infection depends on the realization of some events. First, an infected individual in a source region must circulate and board a plane in order to travel to a destination. Second, the individual must indeed travel and arrive without being detected by entry screening. We assumed that there is no border screening at departure and considered entry screening for only *symptomatic* travellers. Thus, detection at arrival depends on the efficacy of border screening as well as on the clinical state of the infected traveller. Key parameters and assumptions are shown in Table 1 and are described in a later section.

## Data

The daily number of SARS cases in Beijing was obtained through the Chinese health authorities who provided us with a specific data selection including only cases aged $\geqslant 16$ years and belonging to one of the following categories: HCWs, business workers, students, retired, civil servants and staff [9]. According to expert judgement (D. Feng, personal communication), these professional categories are the most likely to travel. For Hong Kong, the SARS data published by the World Health Organization were used [10, 11]. Figure 1 presents the number of cases in Hong Kong by day of symptom onset and the number of cases in HCWs, who represented 22% of all cases. This proportion was high in the first stage of the outbreak, reaching 67% on 5 March 2003, then decreasing after the global alert on 12 March 2003. In Beijing 19% of the SARS cases were HCWs [9].

## Back-calculation

The distribution of infected individuals by day of infection during an interval $T_d$ can be back-calculated from the incident cases in $T_d$, using the distribution of the incubation period. This estimated distribution does not reflect the actual infection curve because individuals acquiring their infection during $T_d$, but having developed their symptoms after day $d$ are not accounted for. The number of these individuals corresponds to a fraction of secondary cases generated by the incident cases in $T_d$. This number is estimated using the distribution of the serial interval as

Table 1. *Data and assumptions for the parameters used in the Monte Carlo simulation calculation for the prediction of the risk of SARS importation* (*Beijing–Frankfurt and Hong Kong–London*)

| | | Data/values | Source/Assumption |
|---|---|---|---|
| SARS epidemic curve | | | |
| Beijing | | | Table 2*a* (in [9]) |
| Hong Kong | | Figure 1 | WHO [10, 11] |
| Delay between symptom onset and hospitalization, days (mean, s.d.) | | | |
| Beijing | 7 Mar.–9 Apr. | 5·0 (6·1) | Pang *et al.* [17] |
| | 10 Apr.–24 Apr. | 3·8 (3·8) | |
| | 25 Apr. onwards | 2·6 (2·2) | |
| Hong Kong | 26 Feb.–25 Mar. | 4·8 (3·5) | Figure 2B in Donnelly *et al.* [15] |
| | 26 Mar.–1 Apr. | 3·8 (2·5) | |
| | 2 Apr. onwards | 3·7 (3·3) | |
| Number of passengers per day (population in millions), travel rate† (per 10 000) | | | |
| Beijing | Urban population | 545 (5·1), 1·08 | International Civil Aviation Organization [20]; National statistics [18] |
| | Health-care workers | n.a., 2·16 | Assumption |
| Hong Kong | General population | 1482 (6·9), 2·43 | International Civil Aviation Organization [20]; National statistics [19] |
| | Health-care workers | n.a., 4·86 | Assumption |
| Entry screening sensitivity, % (mean, 95 % CI) | | | |
| | | n.a. | Base-case assumption, no border screening |
| | | 13·8 (3·9–31·7) | Samaan *et al.* [23] |
| | | 90 (73·0–99·0) | Assumption |

n.a., Not applicable.

\* Gamma distribution is fitted to the data in both regions, and the mean and standard deviation (s.d.) are estimated.

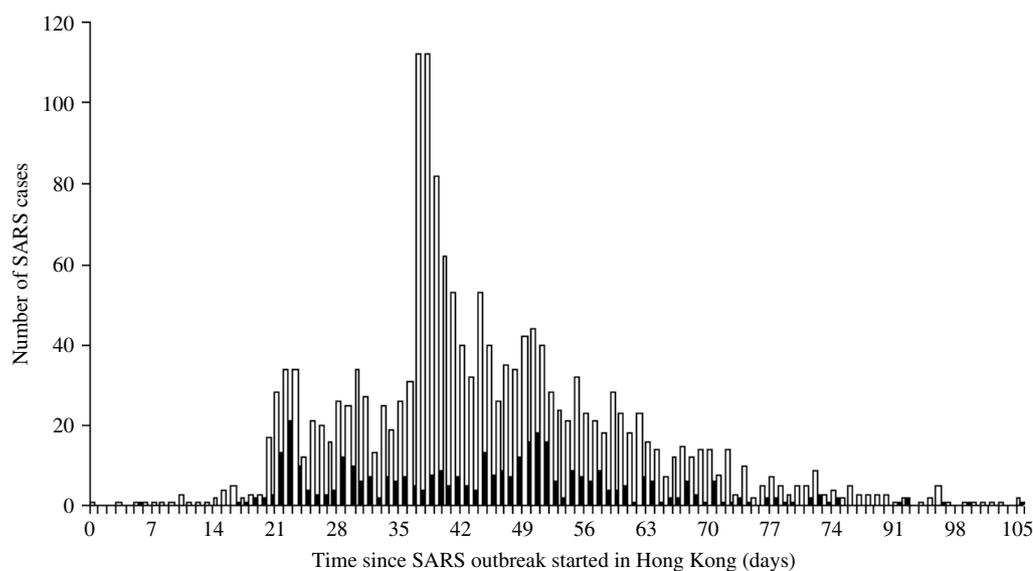† Constant travel rate with no seasonal variation in travel volume.



**Fig. 1.** Distribution of the number of SARS cases by day of symptom onset in Hong Kong. ■, SARS health-care-worker cases; □, All SARS cases.

described by Cauchemez *et al.* [12] (see Appendix A, section 1). Like Cauchemez and colleagues, we assumed that for SARS this distribution is similar in all regions and used a mean (standard deviation) estimate of 8·4 (s.d. = 3·8) days [13]. We applied the back-calculation method [14] on the incident cases in
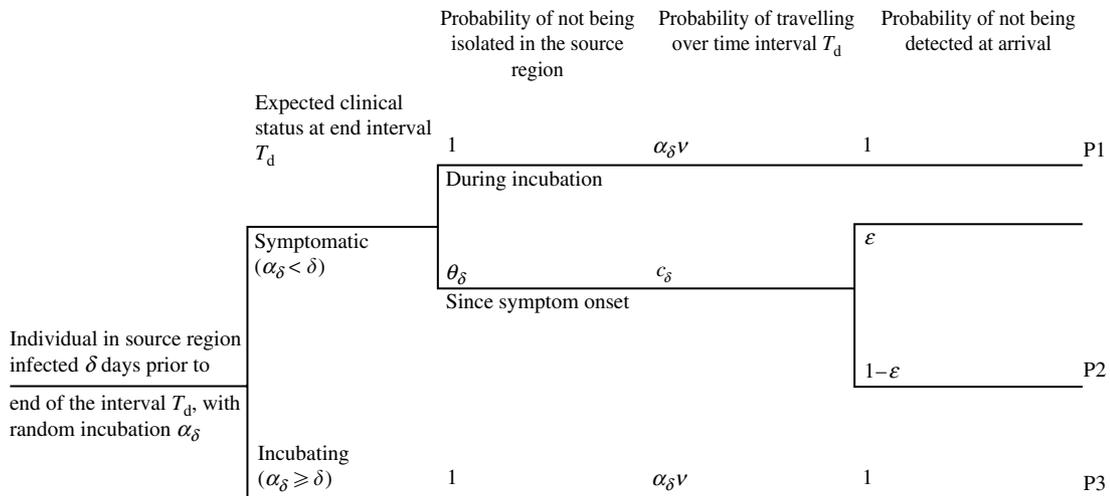
**Fig. 2.** The three pathways for importation of SARS during an interval $T_d$, by an individual infected $\delta$ days since first case in source region. The probability of being isolated after symptom onset is defined by $\theta_\delta$. $v$ is the probability of travelling. The probability of travelling is defined during the incubation period $a_\delta$, or during the symptomatic stage $\delta - a_\delta$. $\varepsilon$ is the probability of detecting a symptomatic traveller at entry to destination.

$T_d$ and on those predicted (Appendix A, section 2). We note by $\bar{h}(\delta)$ ($\delta = 0, 1, \ldots, d$) the estimated number of individuals infected $\delta$ days prior to day $d$ (end of $T_d$).

## Probabilistic Monte Carlo simulation calculation

Over $T_d$, the risk of importation of SARS by every individual $i$ infected $\delta$ days prior to day $d$ ($\delta = 0, \ldots, d$) is simulated using Monte Carlo simulation. At each simulation, we draw a random value $a_{i_\delta}$ from the gamma-distributed incubation period [15]. If $a_{i_\delta}$ is less than $\delta$, then symptom onset for individual $i$ was $\delta - a_{i_\delta}$ days prior to $d$, otherwise the individual was incubating. An infected individual $i$ with $\delta - a_{i_\delta}$ days since symptom onset to day $d$ may travel either during the incubation period (Fig. 2, pathway P1) or after symptom onset (pathway P2). In pathway P2, the individual may be not isolated, may travel some time during the symptomatic phase and arrive undetected at the destination. The infected individual who is expected to be incubating on day $d$ may travel during the incubation period and be undetected by entry screening (pathway P3). The potential occurrence of these events is generated using epidemiological, travel and entry screening parameters (Table 1). The way these parameters flow into the calculation is presented in Appendix A (sections 3–4).

For an infected individual, a single Monte Carlo simulation represents a potential realization of that individual's pathway, quantified as the *probability of arriving by air travel and undetected at entry within time interval $T_d$* (Appendix A, section 4). Summing these probabilities over all possible values of $\delta$, we obtain the total cumulated number of imported infected individuals. The simulations were run with WinBUGs software [16] and results are based on 40 000 runs.

## Parameter description and sources

We considered hospitalization as the main reason for an infected individual to be isolated. We calculated the probability of not being isolated using the distribution of the delay between symptom onset and hospitalization (Appendix A, section 3). Estimates of this distribution are available for three time periods during the SARS outbreak in Beijing and Hong Kong, according to local control measures (Table 1). The mean time from symptom onset to hospitalization fell from 5 to 2·6 days during the outbreak in Beijing [17]. For Hong Kong, it fell from 4·8 to 3·7 days [15].

We considered a flight duration of 10 h for Beijing–Frankfurt and of 12 h for Hong Kong–London. We first assumed that all infected individuals can travel. Their daily probability $v$ was estimated from the daily number of travellers divided by the size of the population in the urban areas for Beijing [18] and of the general population in Hong Kong [19]. Only direct flights from the International Civil Aviation Organization database [20] were considered (Table 1).

In a second alternative, we subjectively assumed that the probability of a HCW travelling was twice that of the general population.

Possible in-flight transmissions of SARS have been documented during the SARS outbreak [6, 21, 22]. Since SARS infectiousness starts after symptom onset, in-flight transmission from incubating travellers is unlikely while symptomatic travellers might generate infections in-flight. To investigate whether or not this impacts our results, we considered the probability of being infectious and transmitting the infection for symptomatic travellers accounting for flight duration (Appendix B).

Information on the sensitivity of symptom screening at entry is scarce. In our analysis, we used two alternatives compared with a base case of no border screening (Table 1). Entry screening with sensitivity 13·8% [95% confidence interval (CI) 3·9–31·7] was used. This value was estimated from a study aimed at detecting symptomatic travellers entering Australia during the SARS outbreak [23]. In the second alternative, we assumed a mean sensitivity of 90% (95% CI 73–99) and applied the same variability (S.D. = 6·9%) as in the above study [23]. We considered this alternative to approximate what a near-perfect screening programme might achieve.

Our main output is the cumulated number of individuals, initially infected in Beijing or Hong Kong, and arriving undetected at their respective destinations during a time-interval $T_d$. Results are presented in terms of a probabilistic distribution, summarized across the 40 000 runs (mean, 95% CI). We start the estimates for 7 March, 1 week before the global alert, corresponding to $d = 14$ for Beijing–Frankfurt and $d = 20$ for Hong Kong–London. Results are presented for several intervals $T_d$.

## RESULTS

### Estimate of the number of individuals by day of infection in the source regions

For illustration we present the cumulative distribution of the new SARS infections in Hong Kong for $T_d$ where $d = 20$, 34, 41 and 48 (Fig. 3). Each graph in Figure 3 includes two distributions: one is back-calculated based on the complete SARS epidemic data and the second is based on the adjusted data until $d$ (Appendix A, section 1). For each $T_d$, the two distributions are comparable except during the last week prior to day $d$, during which a discrepancy is

noticed between the estimates based on the full epidemic data and those based on the adjusted data using Cauchemez et al.'s technique [12]. This discrepancy reflects the uncertainty in the adjusted data that we suspected to be due to the variability in the serial interval distribution. The same conclusion is drawn from the estimated cumulative distributions of the new SARS cases in Beijing.

### SARS importation and effect of border screening at arrival

The cumulated numbers of imported SARS cases [values given are mean (95% CI)] for Hong Kong–London and Beijing–Frankfurt, assuming that all infected individuals were equally likely to travel, are presented in Tables 2a and 2b. When no entry screening was assumed, the cumulated number of SARS cases potentially imported from Hong Kong to London reaches at least one case after 21 March ($T_{34}$). A total of 0·8 imported cases (0·6–0·9) is estimated at $T_{34}$ with a 0·53 probability (0·48–0·59) that this number exceeds one case. At $T_{35}$, 1·4 cases (1·2–1·6) are estimated to be imported with a 0·75 probability (0·70–0·80) of exceeding one case. Results under each of the two other levels of entry screening (13% and 90%) are comparable to those under no screening.

For Beijing–Frankfurt (Table 2b), the cumulated number of imported SARS cases increased during the outbreak but remained <1. It was estimated as 0·7 cases (0·6–0·8) at $T_{63}$ (25 April) and stabilized at 0·8 (0·6–0·9) from $T_{70}$ onwards. Similar to Hong Kong–London, entry screening had no effect on the results.

### SARS importation when infected HCWs are more likely to travel

Under the assumption that HCWs were more likely to travel, we estimated that for Hong Kong–London, at least one case arrives undetected from day 34 (Figure 4, Table 2a). The upper limit of the 95% CI of the cumulated number of imported cases reaches one case at $T_{27}$ (0·4–1·0) assuming no entry screening. For that interval, the probability that at least one case arrives in London is 0·48 (0·36–0·65). A mean of one imported case is estimated at $T_{34}$ (0·7–1·6) and there is a 95% chance that this number is between one and two cases at $T_{35}$ (1·2–2·1). One week after the epidemic peak in Hong Kong at $T_{41}$, a mean of two cases is estimated and the probability of importation
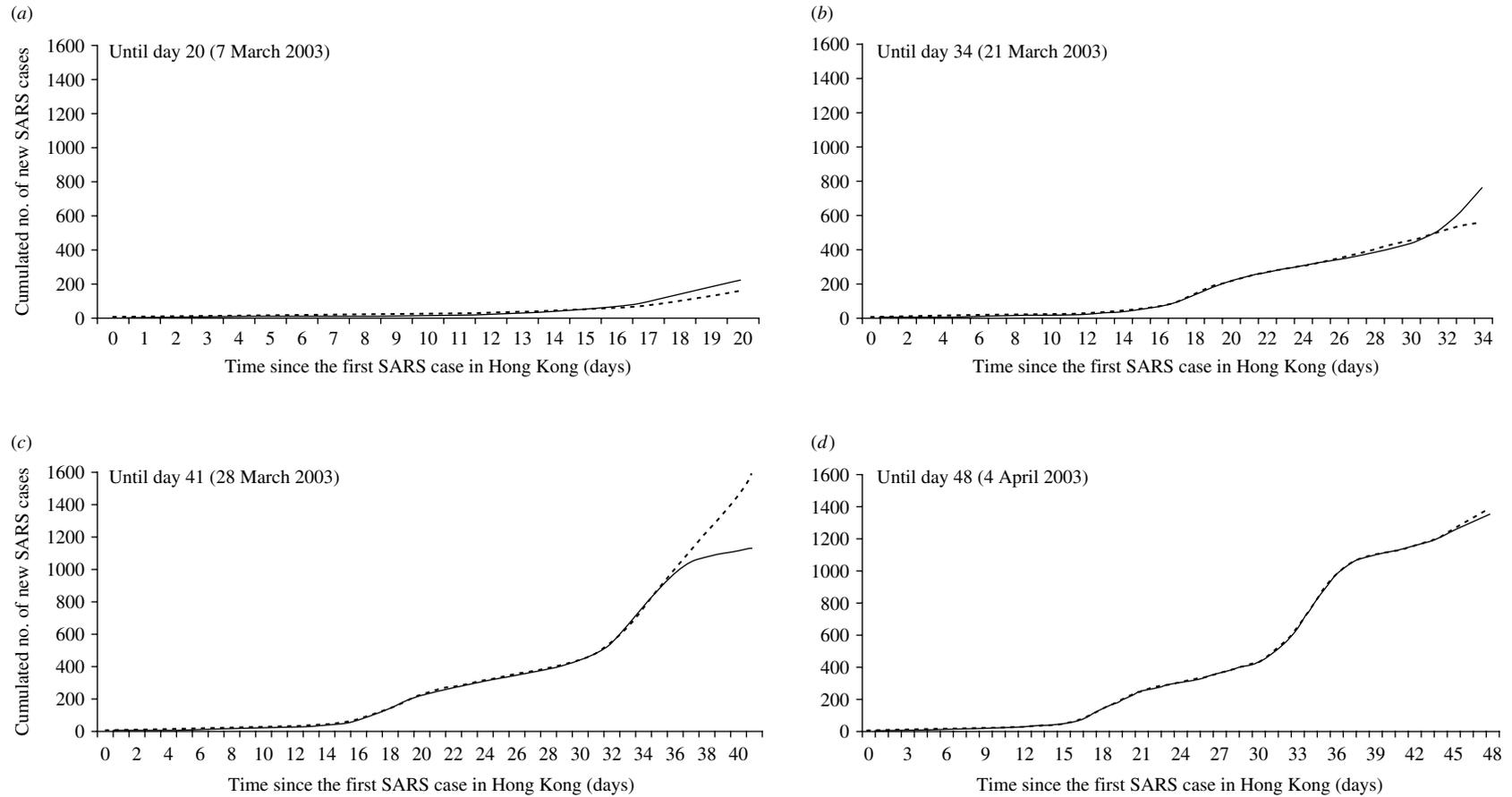
**Fig. 3.** Cumulated number of new infected SARS cases in Hong Kong since the first case's symptom onset until (*a*) day 20, (*b*) day 34, (*c*) day 41 and (*d*) day 48 during the SARS outbreak. The number of new infections is back-calculated from the complete SARS epidemic during the 2003 outbreak (——) and from the SARS data observed until each respective day (- - - -).

Table 2*a*. *Expected daily number of imported SARS cases (Hong Kong–London) since first case in Hong Kong*

| Sensitivity of entry screening in London (%) … | | Constant and identical travel rate | | | Travel rate in HCWs doubles the rate in non-HCWs | | |
|---|---|---|---|---|---|---|---|
| | | No entry screening | 13·8 (3·9–31·7) | 90 (73–99) | No entry screening | 13·8 (3·9–31·7) | 90 (73–99) |
| Time interval* (days)/date, 2003 | | | | | | | |
| $T_{20}$ | 7 Mar. | 0·11 (0·1–0·13) | 0·11 (0·1–0·13) | 0·1 (0·08–0·12) | 0·15 (0·1–0·23) | 0·14 (0·1–0·23) | 0·13 (0·1–0·21) |
| $T_{27}$ | 14 Mar. | 0·51 (0·45–0·57) | 0·5 (0·44–0·56) | 0·45 (0·38–0·52) | 0·67 (0·44–1·04) | 0·66 (0·43–1·02) | 0·59 (0·38–0·92) |
| $T_{34}$ | 21 Mar. | 0·8 (0·7–0·9) | 0·8 (0·6–0·9) | 0·7 (0·6–0·8) | 1·0 (0·7–1·6) | 1·0 (0·6–1·6) | 0·9 (0·6–1·4) |
| $T_{35}$ | 22 Mar. | 1·4 (1·2–1·6) | 1·4 (1·2–1·6) | 1·3 (1·0–1·5) | 1·6 (1·2–2·1) | 1·6 (1·2–2·0) | 1·4 (1·1–1·9) |
| $T_{41}$ | 28 Mar. | 1·8 (1·6–2·0) | 1·8 (1·5–2·0) | 1·6 (1·4–1·9) | 2·1 (1·6–2·9) | 2·1 (1·6–2·9) | 1·9 (1·4–2·7) |
| $T_{48}$ | 4 Apr. | 2·0 (1·7–2·4) | 2·0 (1·7–2·4) | 1·9 (1·5–2·2) | 2·5 (1·8–3·7) | 2·5 (1·7–3·7) | 2·3 (1·6–3·4) |
| $T_{55}$ | 11 Apr. | 2·3 (1·9–2·8) | 2·3 (1·9–2·7) | 2·2 (1·8–2·6) | 2·9 (2·0–4·2) | 2·8 (1·9–4·2) | 2·7 (1·9–4·0) |
| $T_{62}$ | 18 Apr. | 2·5 (2·0–3·0) | 2·5 (2·0–3·0) | 2·4 (2·0–2·9) | 3·1 (2·1–4·6) | 3·1 (2·1–4·6) | 3·0 (2·1–4·5) |
| $T_{69}$ | 25 Apr. | 2·6 (2·1–3·1) | 2·6 (2·1–3·1) | 2·5 (2·1–3·1) | 3·2 (2·2–4·7) | 3·2 (2·2–4·7) | 3·1 (2·2–4·7) |
| $T_{76}$ | 2 May | 2·6 (2·1–3·2) | 2·6 (2·1–3·2) | 2·6 (2·1–3·2) | 3·3 (2·2–4·9) | 3·3 (2·2–4·9) | 3·3 (2·2–4·9) |
| $T_{83}$ | 9 May onwards | 2·7 (2·2–3·2) | 2·7 (2·2–3·2) | 2·7 (2·2–3·2) | 3·4 (2·3–5·0) | 3·4 (2·3–5·0) | 3·3 (2·2–5·0) |

HCW, Health-care worker.
Values are mean (95% CI).
* Time interval since the first case's symptom onset in Hong Kong was on 15 February 2003.

of at least one case is 0·87 (0·80–0·95). From $T_{55}$ onwards, there is a 95% chance that the cumulated number of imported cases is between two and five.

Under the base-case hypothesis on entry screening, the estimate for Beijing–Frankfurt is 0·5 cases at $T_{55}$ with the upper limit of the 95% CI exceeding one case (0·3–1·1). The probability that the number of imported cases exceeds one case at $T_{55}$ is 0·41 (0·25–0·65). On the following $T_d$, the estimates come within reach of one case with the upper limit of the 95% CI reaching two cases (Table 2*b*). In either scenario, entry screening with 13·8% or 90% sensitivity does not affect the results.

## DISCUSSION

Our results show that the transmission dynamic which defines the size of the epidemic in an affected region determines the time of a potential exportation of the infection through the air transportation network. In both scenarios, the results indicate that the risk of importation increases with amplification of local dissemination in the affected areas.

In the base case of no entry screening and when assuming all individuals in the source region have an equal probability of travelling, the method reveals no importation for Beijing–Frankfurt, which was effectively observed in 2003. For Hong Kong–London, the results indicate that importation occurred within 35 days since the time when the first case was reported in Hong Kong. This fits well with the imported case identified in the United Kingdom within this period and reported by the World Health Organization on 18 March [24]. Under the assumption that travel rate in HCWs is twofold that of the travel rate in the population in the source region, the model estimates that at least one case was imported from Beijing–Frankfurt between 11 and 18 April ($T_{56}$). This did

Table 2b. *Expected daily number of imported SARS cases (Beijing–Frankfurt) since first case in Beijing*

| Sensitivity of entry screening in Frankfurt (%) … | Constant and identical travel rate | | | Travel rate in HCWs is twofold the rate in non-HCWs | | |
|---|---|---|---|---|---|---|
| | No entry screening | 13·8 (3·9–31·7) | 90 (73–99) | No entry screening | 13·8 (3·9–31·7) | 90 (73–99) |
| Time interval* (days)/date, 2003 | | | | | | |
| $T_{14}$    7 Mar. | 0·006 (0·004–0·008) | 0·006 (0·004–0·008) | 0·005 (0·004–0·007) | 0·01 (0·004–0·02) | 0·01 (0·004–0·02) | 0·007 (0·003–0·014) |
| $T_{21}$    14 Mar. | 0·020 (0·016–0·023) | 0·019 (0·016–0·023) | 0·017 (0·013–0·021) | 0·03 (0·01–0·05) | 0·03 (0·01–0·05) | 0·02 (0·01–0·05) |
| $T_{28}$    21 Mar. | 0·025 (0·019–0·031) | 0·024 (0·019–0·03) | 0·021 (0·016–0·027) | 0·04 (0·02–0·08) | 0·04 (0·02–0·08) | 0·03 (0·01–0·07) |
| $T_{35}$    28 Mar. | 0·055 (0·046–0·065) | 0·05 (0·04–0·06) | 0·05 (0·04–0·06) | 0·08 (0·04–0·17) | 0·08 (0·04–0·16) | 0·07 (0·03–0·14) |
| $T_{42}$    4 Apr. | 0·14 (0·12–0·16) | 0·14 (0·12–0·16) | 0·12 (0·10–0·15) | 0·2 (0·1–0·4) | 0·2 (0·1–0·4) | 0·2 (0·1–0·4) |
| $T_{49}$    11 Apr. | 0·23 (0·19–0·27) | 0·22 (0·19–0·26) | 0·20 (0·16–0·24) | 0·3 (0·2–0·7) | 0·3 (0·2–0·7) | 0·3 (0·1–0·6) |
| $T_{56}$    18 Apr. | 0·44 (0·37–0·51) | 0·43 (0·37–0·50) | 0·39 (0·32–0·46) | 0·6 (0·3–1·3) | 0·6 (0·3–1·2) | 0·6 (0·3–1·1) |
| $T_{63}$    25 Apr. | 0·69 (0·58–0·82) | 0·68 (0·57–0·80) | 0·62 (0·50–0·74) | 1·0 (0·5–1·9) | 1·0 (0·5–1·9) | 0·9 (0·4–1·7) |
| $T_{70}$    2 May | 0·8 (0·6–0·9) | 0·8 (0·6–0·9) | 0·7 (0·6–0·9) | 1·1 (0·6–2·1) | 1·1 (0·5–2·1) | 1·0 (0·5–1·9) |
| $T_{77}$    9 May | 0·8 (0·6–0·9) | 0·8 (0·6–0·9) | 0·7 (0·6–0·9) | 1·1 (0·5–2·1) | 1·1 (0·5–2·1) | 1·0 (0·5–2) |
| $T_{83}$    16 May onwards | 0·8 (0·6–0·9) | 0·8 (0·6–0·9) | 0·8 (0·6–0·9) | 1·1 (0·5–2·1) | 1·1 (0·5–2·1) | 1·1 (0·5–2·1) |

HCW, Health-care worker.
Values are percentage (95% CI).
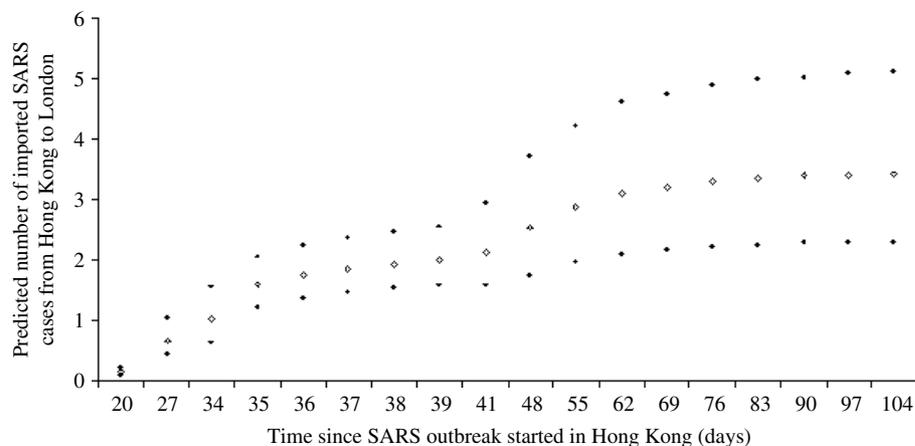* Time interval since the first case's symptom onset in Beijing which was on 21 February 2003.



**Fig. 4.** The mean (95% CI) of the expected cumulated number of SARS cases imported from Hong Kong to London during the outbreak. Estimates are presented by time since first symptom onset of the first case.

not happen in reality, but the probability of that occurring at $T_{56}$ is 0·45 (0·26–0·71). For Hong Kong–London, importation is estimated to occur one day earlier ($T_{34}$ instead of $T_{35}$) when HCWs are assumed more likely to travel. In both scenarios, the results with no entry screening are comparable to

those with 13·8% or 90% screening sensitivity, suggesting that screening which targets SARS symptomatic travellers does not markedly reduce the risk of importation, since symptomatic individuals are likely to be isolated before attempting to travel. The probability of travelling can be regarded as the probability 'of boarding a plane and bypassing exit screening', assuming that exit screening is already considered. Thus, in the first assumption on travel rate (equal daily chance to travel $= v$), an infected HCW has a 0·5 probability of being detected at an exit border and a twofold chance of travelling ($2v$). In the second assumption, HCWs have a travel rate of $2v$ and can be considered as having a 100% chance of bypassing exit screening and a twofold chance of travelling ($2v$). The distribution of the numbers of imported SARS cases remained unaltered when we assumed in-flight transmission (Appendix B). The expected number of generated cases during a 10–12 h flight is almost null, which is consistent with a previous analysis showing that SARS asymptomatic travellers are unlikely to develop symptoms in-flight, and subsequently would not become infectious and generate new infections in-flight [25].

For a real-time application of our method, input for the epidemiological parameters (e.g. the incubation period, serial interval, infectiousness) is required. This is unlikely to be readily available in the very early stage of an outbreak of an emerging infectious disease. In addition, the role of control measures in preventing infected individuals from travelling is not instantly assessed in order to derive assumptions on the delay to isolation. However, inference on this delay from previous outbreaks could be applied. The instant risk estimate also requires the daily number of cases, which should be immediately available, but this could be difficult to achieve during an ongoing outbreak. It requires a worldwide collaboration for a global harmonization of the different surveillance systems as well as an enlarged network for a wide exchange of experiences and data.

We showed that estimates of the distribution of infected individuals by day of infection can be derived from a daily epidemic curve observed in an interval $T_d$ (using the back-calculation method) after adjusting the data by predicting the number of unobserved infected cases (e.g. cases whose symptom onset occur in the following days and their infection occurs during $T_d$). These predicted numbers depend on the epidemiological determinants of spread of the infection

in the source region. Although these factors are not well known in the early stage of an outbreak, the predictions could give satisfactory estimates of the infection curve for $T_d$ with more uncertainty in the first stage of the epidemic than in the later stage, as was seen in Cauchemez et al. [12].

In recent studies on the effect of international travel on spread of infections, travel rates between regions were assumed constant and homogenous between peoples [2, 3, 25–30]. In a more recent work, estimates of the number of exported infected cases with SARS or influenza infection were derived using simulated outbreaks distinguishing between individuals with high- and low-travel frequency [4]. These authors studied the potential impact of the relative difference in travel frequency between these categories. Technically different from this model, our method allows for assigning of different travel patterns to different infected groups, provided that demographic information of the diseased individuals is available. In the absence of information on travel rate in HCWs (a group at risk, and a potential spreader of germs), our assumption is a subjective way of exploring the potential effect that different travel rates assigned to different population groups could have on the expected number of imported SARS case. However, we advise that the resulting risk should be interpreted with caution.

One of the key points in our analysis is that border screening has no impact on the expected number of imported SARS cases. This screening targets only symptomatic travellers who are possibly less likely to travel given the delay between their symptom onset and their hospitalization. It misses the incubating travellers who can successfully enter their destination. Thus, as a worldwide concern in delaying the spread of infection, efforts should go into promoting and strengthening interventions to reduce local spread in infected regions.

the communication between the project leaders to access some of the data. We particularly thank Martin Eichner and anonymous readers for their helpful advice and comments.

## DECLARATION OF INTEREST

None.

## REFERENCES

1. **Caley P, Becker NG, Philp DJ.** The waiting time for inter-country spread of pandemic Influenza. *Public Library of Science* 2007; **2**: 1 1–8.

2. **Cooper BS, et al.** Delaying the international spread of pandemic influenza. *Public Library of Science* 2006; **3**: 1–11.

3. **Ferguson NM, et al.** Strategies for mitigating an influenza pandemic. *Nature* 2006; **442**: 448–452.

4. **Hollingsworth TD, Ferguson NM, Anderson RM.** Frequent travelers and rate of spread of epidemics. *Emerging Infectious Diseases* 2007; **13**: 1288–1294.

5. **Glass K, Becker N, Clements M.** Predicting case numbers during infectious disease outbreaks when some cases are undiagnosed. *Statistics in Medicine* 2007; **26**: 171–183.

6. **Desenclos JC, et al.** Introduction of SARS in France, March–April, 2003. *Emerging Infectious Diseases* 2004; **10**: 195–200.

7. **Wilder-Smith A, Paton NI.** Imported cases of severe acute respiratory syndrome to Singapore had impact on national epidemic. *British Medical Journal* 2003; **326**: 1393–1394.

8. **Barnes B, Glass K, Becker NG.** The role of health care workers and antiviral drugs in the control of pandemic influenza. *Mathematical Biosciences* 2007; **209**: 403–416.

9. **Feng D, et al.** The SARS epidemic in Mainland China: bringing together all epidemiological data. *Tropical Medicine and International Health* (in press).

10. **WHO.** Epidemic curves – Severe Acute Respiratory Syndrome (SARS). World Health Organization, 2003 (http://www.who.int/csr/sars/epicurve/epiindex/en/). Accessed 18 May 2006.

11. **WHO.** SARS: how a global epidemic was stopped. Geneva: World Health Organization's Regional Office for the Western Pacific, 2006.

12. **Cauchemez S, et al.** Real-time estimates in early detection of SARS. *Emerging Infectious Diseases* 2006; **12**: 110–113.

13. **Lipsitch M, et al.** Transmission dynamics and control of Severe Acute Respiratory Syndrome. *Science* 2003; **300**: 1966–1970.

14. **Becker NG, Watson LF, Carlin JB.** A method of non-parametric back-calculation and its application to AIDS data. *Statistics in Medicine* 1991; **10**: 1527–1542.

15. **Donnelly CA, et al.** Epidemiological determinants of spread of causal agent of severe acute respiratory syndrome in Hong Kong. *Lancet* 2003; **361**: 1761–1766.

16. **Spiegelhalter DJ, et al.** WinBUGS User Manual: Version 1.4, p. 1. WinBUGs, 2001.

17. **Pang X, et al.** Evaluation of control measures implemented in the severe acute respiratory syndrome outbreak in Beijing, 2003. *Journal of the American Medical Association* 2003; **290**: 3215–3221.

18. **National Bureau of Statistics of China.** *2005 China Statistical Yearbook*. China Statistics Press, 2005.

19. **Hong Kong Statistics.** Hong Kong Government, 2007 (http://www.censtatd.gov.hk/hong_kong_statistics/statistics_by_subject/index.jsp). Accessed 18 May 2006.

20. **ICAO.** Number of passengers from the Traffic by Flight Stage database on international scheduled services. International Civil Aviation Organization, 2006 (http://www.icaodata.com/). Accessed 18 May 2006.

21. **Olsen SJ, et al.** Transmission of the Severe Acute Respiratory Syndrome on Aircraft. *New England Journal of Medicine* 2003; **349**: 2416–2422.

22. **Wilder-Smith A, Leong H, Villacian J.** In-flight transmission of severe acute respiratory syndrome (SARS): a case report. *Journal of Travel Medicine* 2003; **10**: 299–300.

23. **Samaan G, et al.** Border screening for SARS in Australia: what has been learnt? *Medical Journal of Australia* 2004; **180**: 220–223.

24. **WHO.** Cumulative number of reported suspect and probable SARS cases: From 1 February to 18 March 2003. World Health Organization, 2003 (http://www.who.int/csr/sars/tablemarch18/en/). Accessed 18 May 2006.

25. **Pitman RJ, et al.** Entry screening for severe acute respiratory syndrome (SARS) or influenza: policy evaluation. *British Medical Journal* 2005; **331**: 1242–1243.

26. **Brownstein JS, Wolfe CJ, Mandl KD.** Empirical evidence for the effect of airline travel on inter-regional influenza spread in the United States. *Public Library of Science* 2006; **3**: 1826–1835.

27. **Colizza V, et al.** The role of the airline transportation network in the prediction and predictability of global epidemics. *Proceedings of the National Academy of Sciences USA* 2006; **103**; 2015–2020.

28. **Colizza V, et al.** Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *Public Library of Science* 2007; **4**: 95–110.

29. **Guimera R, et al.** The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles. *Proceedings of the National Academy of Sciences USA* 2005; **102**; 7794–7799.

30. **Hufnagel L, Brockmann D, Geisel T.** Forecast and control of epidemics in a globalized world. *Proceedings of the National Academy of Sciences USA* 2004; **101**: 15124–15129.

## APPENDIX A

### 1. Prediction of the unobserved secondary infections based on number of cases with onset currently observed

Given a set of the number of observed cases $O(T)$ with symptom onset available up to a day $T$ ($n_t$, $t \leqslant T$), we predict $X_t^+(T)$, the number of cases initially infected before $T$ whose symptom onset occurs after $T$. We used the method of Cauchemez et al. [1] where the authors developed a technique to timely estimate the reproductive number ($R_t$) of SARS. The $R_t$ is represented by the ratio of $n_t$ and the number of cases they infected, $X_t$: $X_t = X_t^+(T) + X_t^-(T)$. $X_t^-(T)$ is the number of cases initially infected before $T$ that showed symptoms before $T$. As did Cauchemez and colleagues, we estimated the distribution of $X_t(T)/O(T)$ within a Bayesian framework. A decomposition of $X_t(T)$ in the number of early and secondary generated cases follows the equation:

$$
\left.
\begin{aligned}
P[X_t = r/O(T)] &= \sum_{k=0}^{r} P[X_t^+(T) \\
&= r - k/X_t^-(T) = k] \times P[X_t^-(T) = k/O(T)],
\end{aligned}
\right\} \quad (1)
$$

where $P[\ ]$ indicates the probability density. The distributions of $X_t^+(T)/X_t^-(T)$ and $X_t^-(T)/O(T)$ are both estimated. The later is approximated by a sum of independent binomial distributions in the same way as in Wallinga & Teunis [2]:

$$
X_t^-(T)/O(T) \sim \sum_{k \leqslant T} \text{binomial}\,(\psi_{t,k}, n_k), \quad (2)
$$

where $\psi_{t,k}$ is the probability that a case detected at day $k$ has been infected by a case detected at day $t$ ($t, k \leqslant T$) (see supplementary material in Cauchemez et al. [1] for further details). A value of $X_t^+(T)/O(T)$ is derived for each $t$ ($t < T$) by subtracting a value of $X_t^-(T)/O(T)$ from $X_t(T)/O(T)$. $X_t^+(T)/O(T)$ is hence redistributed by day of symptom onset over the 10 days following $T$ ($x_t(T+s)$, $s = 1$–10) using the probability distribution of the serial interval.

### 2. Back-calculation of the observed and the predicted number of incident cases

For a given day $d$, we note by $\Theta(t)_{t,\,0 < t \leqslant d+10}$ the number of cases by day of symptom onset $t$ until day 10 after $d$:

$$
\Theta(t) = \sum_{\tau=0}^{t-1} h(t-\tau, \tau) \quad \text{for} \quad 0 < t \leqslant d+10, \quad (3)
$$

where $h(t-\tau, \tau)$ is the number of individuals infected on day $t-\tau$ (prior to $t$), with an incubation period of $\tau$ days. $h(t-\tau, \tau)$ are assumed independently Poisson-distributed with mean $\bar{h}(t)f_\tau$ where $\bar{h}(t)$ is the expected mean of the number of individuals infected on day $t$ and $f_\tau$ is the probability that incubation time is $\tau$. We used the Expectation Maximization algorithm that gives Smoothed (EMS) predictions [3] to estimate $\bar{h}(t)_{0 < t \leqslant d}$. We refer the reader to Becker et al. [3] for explicit details on the EMS algorithm. The EMS algorithm was coded with R software [4]. In the following, we note $\bar{h}(t)$ by $\bar{h}(\delta)$; the number of individuals infected $\delta$ days prior to $d$ ($\delta = 0, \ldots, d$).

### 3. Distribution of time since symptom onset to hospitalization

The probability of not being isolated equals 1 for an incubating individual. For an individual $i$ with $\delta - a_{i_\delta}$ since symptom onset, this probability is defined using the cumulative density function ($F(\ )$) of the distribution of the delay between symptom onset and hospitalization:

$$
\left.
\begin{aligned}
\theta_{i_\delta} &= \begin{cases} (1 - F(\delta - a_{i_\delta})) & a_{i_\delta} < \delta \\ 1 & a_{i_\delta} \geqslant \delta, \end{cases} \\
F(\delta - a_{i_\delta}) &= \int_0^{\delta - a_{i_\delta}} g(u)\mathrm{d}u,
\end{aligned}
\right\} \quad (4)
$$

where $g$ is the probability density function of this distribution. This is a gamma distribution specific to three time-periods during the SARS outbreak either in Beijing or Hong Kong (see Table 1 of main paper).

### 4. Probability of departing and arriving undetected at destination

At each simulation, for every infected individual $i$ with incubation period $a_{i_\delta}$ ($\delta = 0, \ldots, d$), we calculated the probability of the individual travelling and arriving undetected at the destination as some time between infection and day $d$ ($p_{i_\delta}$). Assuming $\varepsilon$ to be the sensitivity of screening at border entry, so that $1 - \varepsilon$ represents the probability for a symptomatic individual to bypass border control at arrival, we assumed that there is no chance of detecting an individual travelling during the incubation period.

An infected individual $i$ with $\delta - a_{i_\delta}$ days since symptom onset has a probability $\nu a_{i_\delta}$ of travelling during the incubation period $a_{i_\delta}$ (see Figure 2, pathway P1 in the main paper), or a probability $c_{i_\delta}$ of

travelling after developing symptoms during time $\delta - a_{i_\delta}$ (pathway P2):

$$c_{i_\delta} = \underbrace{(1-\nu)(1-\nu)(1-\nu) \cdot \cdots \cdot (1-\nu)}_{a_{i_\delta}}$$
$$\times \underbrace{(\nu + \nu + \nu + \nu + \nu + \cdots \cdots + \nu)}_{\delta - a_{i_\delta}} \tag{5}$$
$$= (1-\nu)^{a_{i_\delta}} (\delta - a_{i_\delta})\nu.$$

For an individual $i$ likely to be incubating on day $d$, the probability of departing some time during the incubation period (pathway P3) is the cumulative probability of the individual travelling during incubation time $\nu a_{i_\delta}$. The specific probability $p_{i_\delta}$ of an individual $i$ among $\bar{h}(\delta)$ infected cases ($\delta = 0, \ldots, d$) is given by:

$$p_{i_\delta} = \begin{cases} \nu a_{i_\delta} & a_{i_\delta} \geqslant \delta \quad (6) \\ \nu a_{i_\delta} + \theta_{i_\delta} c_{i_\delta}(1-\varepsilon) & a_{i_\delta} < \delta. \quad (7) \end{cases}$$

We draw the time until scheduled travel from the uniform distributed over the interval $(0, \delta)$:

$$\text{time}_{i_\delta} \sim \text{uniform}\,(0, \delta).$$

The sum of the probabilities $\{(p_{i_\delta})_i\}_\delta$ approximates the expected number of initially infected individuals in the source region, arriving undetected at the region of destination during the time interval $T_d = (0, d)$; since the beginning of the outbreak until day $d$:

$$E_{T_d} = \sum_{\delta=0}^{d} \sum_{i=1}^{\bar{h}[\delta]} p_{i_\delta}, \tag{8}$$

At each simulation, $E_{T_d}$ is calculated and a summary statistic is derived over 40 000 runs. We assumed that SARS importation from one region to another is a non-homogenous Poisson process characterized by a time-dependent rate $\gamma(t)$. The expected number of arrivals during an interval $T_d$ is defined by the integral $\int_0^d \gamma(t)\partial t$ which we approximated by $E_{T_d}$. Therefore, the number of arrivals of undetected infected travellers in the time interval $T_d$ is regarded as a random variable, which is Poisson-distributed with mean $E_{T_d}$.

## APPENDIX B

### Transmission during flight and effect on the expected number of imported cases

For a transmission to occur during flight, an infected traveller must be infectious and transmit the infection to other passengers during the flight. We assumed that infectiousness and transmission mechanisms during flight are same as while mixing in the population. The probability of being infectious during the flight depends on the time since symptom onset to the scheduled travel. We estimated this probability using the results from the analysis of viral shedding, based on isolation of SARS coronavirus or RT–PCR detection in the respiratory secretions of 415 cases in Hong Kong [5]. This study indicated that viral excretion (infectiousness) was detected in the first or second day after symptom onset in 20% of cases, and in more than 33% of cases one week following symptom onset. Viral excretion was detected in only 17% of cases 2 weeks after symptom onset and in none of cases 3 weeks following symptom onset. $\lambda_{i_\delta}$ is the probability of being infectious for an $i$th individual expected to be symptomatic on day $d$ and travelling while being symptomatic (pathway P3), given time since symptom onset (time$_{i_\delta} - a_{i_\delta} + \Delta_F$), where time$_{i_\delta}$ is the time since infection to scheduled travel and $\Delta_F$ is the flight duration. An infectiousness event is thus generated at each Monte Carlo simulation, using Bernoulli distribution, $r_{i_\delta} \sim \text{Bernoulli}\,(\lambda_{i_\delta})$, so that a random value $r_{i_\delta} = 1$ indicates that the individual is infectious.

We first considered that all infected individuals have their own expected number of secondary cases (over each individual's infection history) given the population basic reproductive number $R_0$ (Table B1). For an $i$th infectious traveller ($r_{i_\delta} = 1$), the expected number of secondary cases $\pi_{i_\delta}$ is gamma-distributed with mean $R_0$ and shape $\alpha$ [6], and the number of secondary cases generated is considered as a random variable ($Z$) which is described by an offspring distribution that has a negative binomial distribution with mean $\pi_{i_\delta}$ and dispersion parameter $1/\alpha$. The fraction of $\pi_{i_\delta}$ caused during a flight duration $\Delta_F$ is represented by $\eta_{i_\delta}$ given a median infectiousness period of $D$ days. A value $\hat{\pi}_{i_\delta}$ is drawn from the gamma distribution at each simulation, and an estimate $\hat{\eta}_{i_\delta}$ is obtained from equation (10). The probability of causing no secondary infections during a flight ($\kappa_{i_\delta}$) is calculated from the probability distribution with mean $\hat{\eta}_{i_\delta}$ and dispersion $1/\alpha$ [eqn (11)] and $\varphi_{i_\delta}$; the probability to generate new cases is derived, and is used drawing a value from a Bernoulli distribution: $r_{\varphi_{i_\delta}} \sim \text{Bernoulli}\,(\varphi_{i_\delta})$. Whenever $r_{\varphi_{i_\delta}} = 1$, a random number of generated cases during flight is drawn from the Poisson distribution with mean $\hat{\eta}_{i_\delta}$ [eqn (11)]:

$$\left.\begin{array}{l} \hat{\pi}_{i_\delta} \propto \Gamma(R_o, \alpha) \\ \hat{\eta}_{i_\delta} = \Delta_F \hat{\pi}_{i_\delta}/D, \end{array}\right\} \tag{9}$$

Table B1. *Data used to estimate transmission parameters when accounting for in-flight transmission*

|  | Data/values | | Source |
|---|---|---|---|
| Infectiousness rate | Detection of viral excretion in respiratory secretions | | Figure 3 in Chan *et al.* [5]. Curve based on respiratory specimens data |
| Basic reproductive number $R_0$ (shape) | | | |
| Beijing | Before intervention | 1·88 (0·12) | Lloyd-Smith *et al.* [6] |
|  | After intervention | 0·28 (0·006) | |
| Hong Kong | Before intervention | 3·6 (0·18) | Lloyd-Smith *et al.* [6] |
|  | After intervention | 0·7 (0·08) | |

Table B2. *Estimates of the number of imported SARS cases (Hong Kong–London) assuming identical travel rate, and no border screening at arrival*

| Time interval*/date 2003 | | Not accounting for in-flight transmission | Accounting for in-flight transmission |
|---|---|---|---|
| $T_{20}$ | 7 Mar. | 0·11 (0·1–0·13) | 0·11 (0·1–0·13) |
| $T_{27}$ | 14 Mar. | 0·51 (0·45–0·57) | 0·51 (0·45–0·57) |
| $T_{34}$ | 21 Mar. | 0·8 (0·7–0·9) | 0·8 (0·7–0·9) |
| $T_{35}$ | 22 Mar. | 1·4 (1·2–1·6) | 1·4 (1·2–1·6) |
| $T_{41}$ | 28 Mar. | 1·8 (1·6–2·0) | 1·8 (1·6–2·0) |
| $T_{48}$ | 4 Apr. | 2·0 (1·7–2·4) | 2·1 (1·7–2·4) |
| $T_{55}$ | 11 Apr. | 2·3 (1·9–2·8) | 2·3 (1·9–2·8) |
| $T_{62}$ | 18 Apr. | 2·5 (2·0–3·0) | 2·5 (2·0–3·0) |
| $T_{69}$ | 25 Apr. | 2·6 (2·1–3·1) | 2·6 (2·1–3·1) |
| $T_{76}$ | 2 May | 2·6 (2·1–3·2) | 2·6 (2·1–3·2) |
| $T_{83}$ | 9 May onwards | 2·7 (2·2–3·2) | 2·7 (2·2–3·2) |

\* Time interval (days) since the first case's symptom onset in Hong Kong on 15 February 2003.
Values are mean (95 % CI).

$$\left.\begin{array}{l} \kappa_{i_\delta} = (1 + \alpha\hat{\eta}_{i_\delta})^{\frac{1}{\alpha}} \\ \varphi_{i_\delta} = (1 - \kappa_{i_\delta}), \end{array}\right\} \tag{10}$$

$$\rho_{i_\delta} \propto \text{Poisson}\,(\hat{\eta}_{i_\delta}), \tag{11}$$

The expected total number of generated cases during a flight is given by:

$$G_{T_d} = \left( \sum_{\text{Infectious},\, i_\delta} \rho_{i_\delta} \right). \tag{12}$$

The cumulated number of SARS cases arriving undetected from an infected region at a destination since the beginning of the outbreak until a day $d$ includes those imported cases who are initially infected in a source region ($E_{T_d}$) and those generated by infectious travellers during flight $G_{T_d}$ [eqn (12)]:

$$S_{T_d} = E_{T_d} + G_{T_d}, \tag{13}$$

In the case where we account for in-flight-generated infections (Table B2), the number of arrivals of undetected infected travellers in the time-interval $T_d$ is Poisson-distributed with mean $S_{T_d}$.

## REFERENCES

1. **Cauchemez S, et al.** Real-time estimates in early detection of SARS. *Emerging Infectious Diseases.* 2006; **12**: 110–113.
2. **Wallinga J, Teunis P.** Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *American Journal of Epidemiology* 2004; **160**: 509–516.
3. **Becker NG, Watson LF, Carlin JB.** A method of non-parametric back-calculation and its application to AIDS data. *Statistics in Medicine* 1991; **10**: 1527–1542.
4. **R (computer program).** R: A language and environment for statistical computing, version 2.5.1. Vienna, Austria, 2007.
5. **Chan PKS, et al.** Laboratory diagnosis of SARS. *Emerging Infectious Diseases* 2004; **10**: 825–831.
6. **Lloyd-Smith JO, et al.** Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; **438**: 355–359.