Supplementary Material

Initial Human Transmission Dynamics of the

Pandemic (H1N1) 2009 Virus in North America

Babak Pourbohloul, Armando Ahued, Bahman Davoudi, Rafael Meza, Lauren A. Meyers, Danuta M. Skowronski, Ignacio Villaseñor, Fernando Galván, Patricia Cravioto, David J. D. Earn, Jonathan Dushoff, David Fisman, W John Edmunds, Nathaniel Hupert, Samuel V. Scarpino, Jesús Trujillo, Miguel Lutzow, Jorge Morales, Ada Contreras, Carolina Chávez, David M. Patrick, Robert C. Brunham

The network approach used in our analysis in the main text provides the initial reproduction number using the number of daily reported cases (symptoms onset) J_t, the *removal distribution function* ψ_τ, transmissibility *T*, and the average number of contacts of an infected individual with susceptible individuals z, as input parameters. In the following, t is the time passed since the first infection in the population and τ is the infection period for an individual.

There are three characteristic periods defining the disease states in each individual, namely latent τ^l, asymptomatic τ^a, and symptomatic τ^s. The transmissibility, defined as the probability of disease transmission along one contact during the whole period of infection $\tau = \tau^l + \tau^a + \tau^s$, is negligible for the latent period and it is given by

$$T_\tau = 1 - \prod_{\tau' < \tau} (1 - \beta_{\tau'}) \qquad \text{S(1)}$$

β_τ is the transmission probability per day. In general β_τ can vary from day to day (β_τ is the discrete form of the continuous function $\beta(t)$ described in Ref (xx) in the main text). Typically β_τ is initially zero (i.e., during the latent period), increasing to a certain level and then declining (during the infectious period) before finally vanishing and remaining zero (e.g., in the case of permanent recovery or death) [see Figure S1]. In practice, the functional form of β_τ should be estimated from the actual transmission profile corresponding to a specific disease. Given that little was known about the actual profile of β_τ corresponding to the S-OIV (H1N1) strain at the time of publication of this paper, we assumed $\beta_\tau = \bar{\beta}$ in the $\tau^a + \tau^s$ period and zero for the latent period. This assumption was used for the bulk of the calculations presented in Figure 4 in the main text (388,080×6 different combinations). Additionally, we applied the infection profile corresponding to the A/H1N1 strain provided by Carrat F *et al.* (American Journal of Epidemiology, 2008 167(7):775-85), and Ferguson NM *et al.* (Nature 442, 448-452 (27 July 2006)) to a 10% subset of these combinations and the results for *R* were within the range reported in Table 1.

The removal distribution function ψ_τ, is the probability that each infected individual is removed from the infectious population either due to experiencing severe symptoms, recovery, etc. This distribution can be easily constructed from the time period τ^{sr} for infected individuals, τ^{sr} being the period between symptom onset and being removed from the population (see Figure S1). Using hospitalization/self-isolation dates and the self-declared symptoms onset dates that were provided to us by the Mexico City health authorities, we derived the removal distribution function, which is depicted in Figure 2.

The expected transmissibility can be defined as

$$T = \sum_\tau T_\tau \psi_\tau \qquad \text{S(2)}$$

The reproduction number can then be defined as $R = zT$ (Newman MEJ, Phys. Rev. E 66, 016128 (2002); Davoudi B *et al*. arXiv:0905.0728v1 [q-bio.QM]).

For an ongoing infection a fraction of infected individuals have recovered at day t . The total number of removed (R_t) or infectious (I_t) individuals up to time t depends on the number of individuals who were infected for some period of time during the whole range $0 < t' < t$ as well as on the removal distribution function, namely

$$R_t = \sum_{\tau=0}^{t} \psi_\tau \sum_{\tau'=\tau}^{t} J_{t-\tau'}$$
$$I_t = \sum_{\tau=0}^{t} \psi_\tau \sum_{\tau'=0}^{\tau} J_{t-\tau'} \qquad \text{S(3)}$$

We define the expected transmissibility of removed individual up to time t as T_t^r . This is a generalization of the expected transmissibility considering the correction that not all infected individuals are removed at time t and it can be obtained from equation (1.1) replacing ψ_τ with ${\psi_{\tau,t}}^r$

$$T_t^r = \sum_\tau T_\tau {\psi^r}_{\tau,t} \qquad , \qquad \text{S(4)}$$

where ${\psi_{\tau,t}}^r$ is given by

$${\psi_{\tau,t}}^r = \psi_\tau \frac{\sum_{\tau'=\tau}^{t} J_{t-\tau'}}{\sum_{\tau=0}^{t} \psi_\tau \sum_{\tau'=\tau}^{t} J_{t-\tau'}} \qquad \text{S(5)}$$

The total number of contacts between the removed individuals and the rest of population is given by zR_t, a fraction T_t^r of which will lead to infections up to time t. Those infected individuals who acquired the infection from the removed individual up to time t are either infectious or removed, which we denote by I_t^r and R_t^r respectively. This means that the following equation holds

$$R_t z T_t^r = R_t^r + I_t^r \qquad \text{S(6)}$$

$R_t^r + I_t^r$ can be calculated from the following equation

$$R_t^r + I_t^r = \sum_{t'=0}^{t} J_{t'} \left[1 - \frac{\sum_{\tau=0}^{t} J_{t,t-t'+\tau}{}^{i} \Delta T_\tau}{\sum_{\tau=0}^{t} J_{t',\tau}{}^{i} \Delta T_\tau} \right] \qquad \text{S(7)}$$

where $\Delta T_\tau = T_{\tau+1} - T_\tau$, $J_{t,\tau}{}^{i} = J_{t-\tau}\psi_\tau{}^{c}$ and $\psi_\tau{}^{c} = \sum_{\tau'=\tau} \psi_\tau$.

The distribution of $\psi_\tau{}^{c}$ - the probability that an individual remains infectious for *at least* a period of τ – for the Mexico City data is shown in Figure S3.

The algorithmic procedure to calculate the initial reproductive number is as follows:

1) Prepare the symptoms onset notification data J_τ.
2) Calculate the removal distribution ψ_τ from τ^{sr} for a set of infected individuals from notification/hospitalization data.
3) Using the results in (2), calculate $\psi_\tau{}^{c}$ and other related quantities.
4) Start with an arbitrary value (reasonable guess) for β_τ and calculate T_τ from equation S1.
5) Calculate $R_t^r + I_t^r$ by using equation S7 from values obtained in (3) and (4).
6) Calculate $\psi_{\tau,t}{}^{r}$ using equation S5.
7) Calculate T_t^r using equation S4.
8) Insert all relevant values in S6:

- If the two sides of equation S6 are equal: Accept the value of β_τ in (4); proceed to (9);

- If the two sides of equation S6 are not equal: Go to (4), choose a better estimate for β_τ using some root-finding algorithm; repeat the subsequent steps.

9) Use β_τ accepted in (8) to calculate T using equations S1 and S2.

10) Use $R = zT$ to calculate the initial reproduction number.

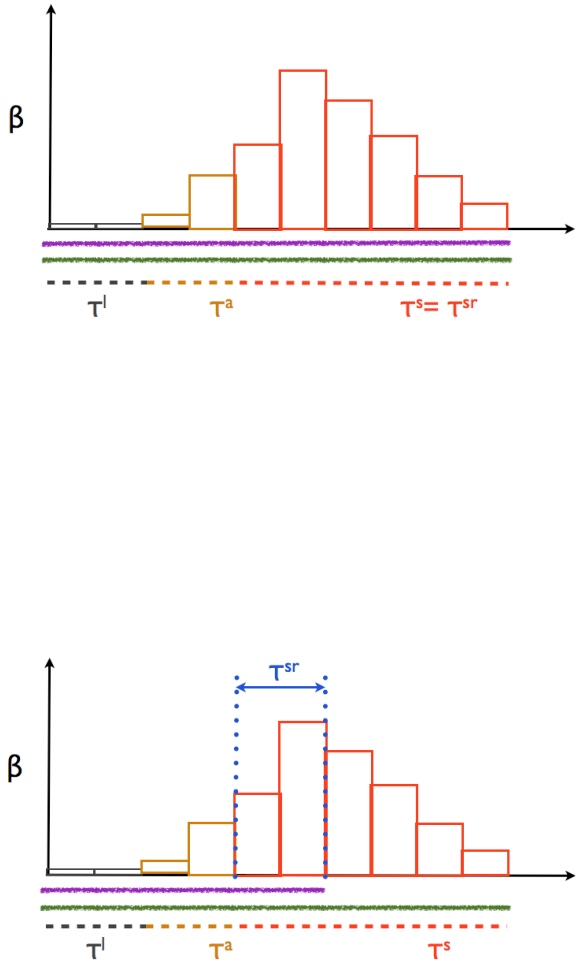

Figure S1. A hypothetical infectivity profile β for an infected individual: Top panel) τ^l, τ^a and τ^s are the latent, asymptomatic and symptomatic periods, respectively. The green line shows the total generation time corresponding to this hypothetical disease for a typical person. The length of the line may vary from person to person. In the absence of intervention, the person has a chance to complete the natural course of infection (purple line). Bottom panel) If intervention is in place (self-isolation, hospitalization, etc.) or if the patient succumbs to death due to the severity of symptoms, the person is removed before he/she has a chance to complete the generation time. In this case he/she has the potential to infect others for a shorter period of time (purple line) although the generation time, as the biological characteristic of disease, remains the same (green line). The interval τ^{sr} – the time between symptoms onset and hospitalization/self-isolation – can be calculated from the databases of suspect and/or confirmed cases provided to us.

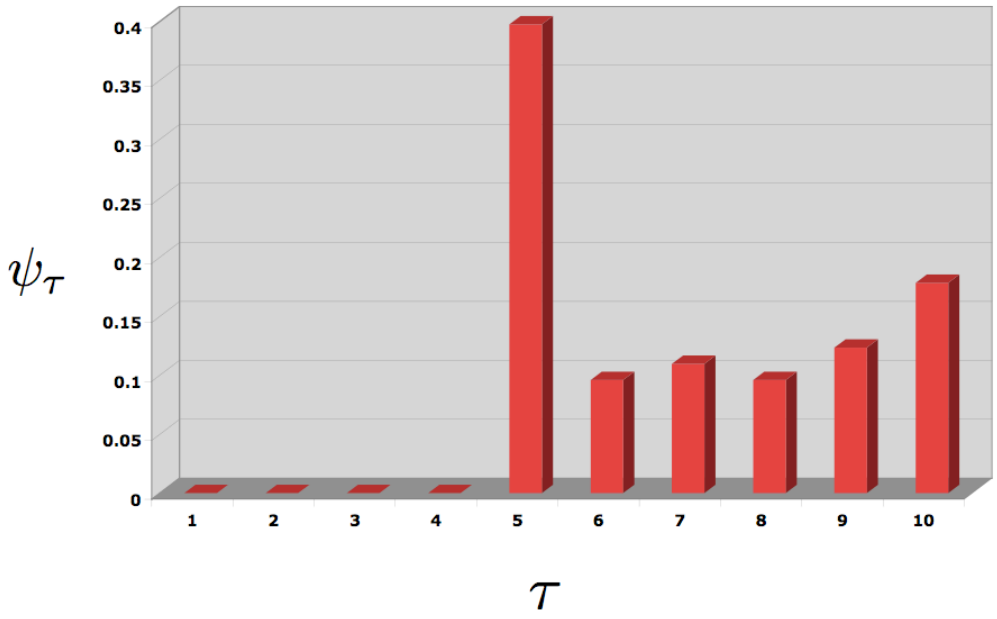

Figure S2. The *removal distribution function* ψ_τ: The interval τ is the sum of the latent period τ^l, asymptomatic period τ^a, and τ^{sr}. This latter period is the difference between the hospitalization/self-isolation date and the self-declared symptoms onset date recorded in the Mexico City health registry for suspect cases before April 19. The distribution is shown for the baseline parameter set in which the latent and asymptomatic periods are set to 3 and 1 days, respectively. The large column on Day 5 shows that about 40% of patients reported the date of symptoms onset on the same day that they visited hospitals or other healthcare facilities. During sensitivity analysis, the values of latent and asymptomatic periods vary, which results in shifting the above distribution to the right or left. A similar approach was used for all time series introduced in the main text.

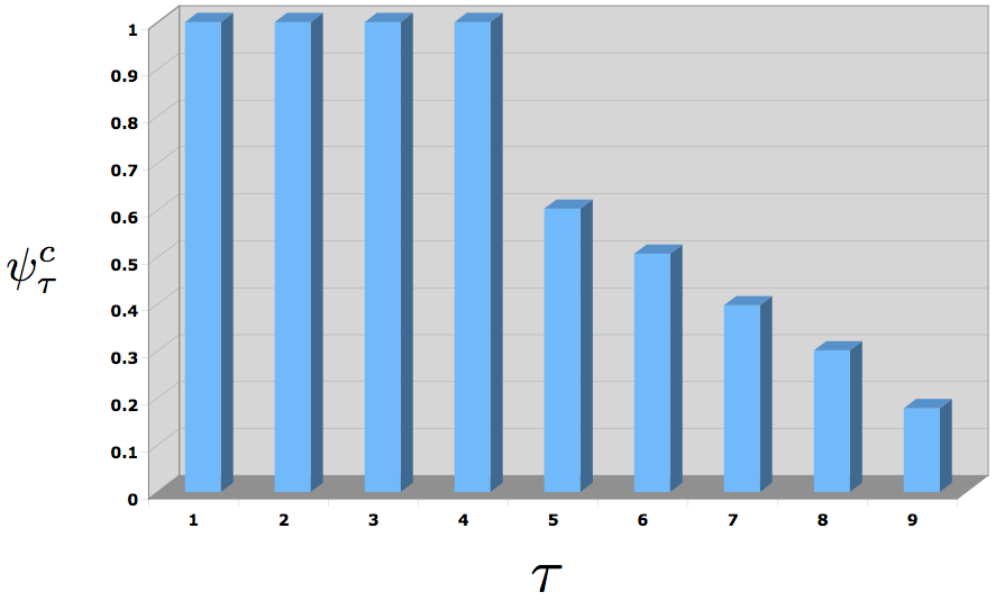

Figure S3. The distribution of ψ_τ^c – the probability that an individual remains infectious for *at least* a period of τ – for the Mexico City data.