Effects of school closure on pandemic influenza incidence in Alberta, Canada

David J.D. Earn, DaiHai He, Mark B. Loeb, Kevin Fonseca, Bonita E. Lee, and Jonathan Dushoff


APPENDIX

Normalization of case data

The bottom right of Figure 1 shows a bar plot of the cumulative sum, for each age, of all laboratory-confirmed pH1N1 in 2009; it does not take the age structure of the population into account. We investigated the potential effects of age structure on the shape of the bar plot. Figure 5 compares the total incidence across age (panel a) with per capita incidence (panel b). The distributions of raw incidence and incidence per 100,000 population are very similar. Figure 5 also compares the age distributions of confirmed cases before and after testing restrictions were implemented (on 30 October 2009), emphasizing the sampling bias caused by testing restrictions.

Transmission model

We used a standard SIR framework (23) with two age classes: individuals aged 5–18 (s) and all others (o). We denote the numbers of schoolchildren (other individuals) in each disease compartment by $S_s$ and $I_s$ ($S_o$ and $I_o$), and use $N$ to denote the total population size, including immune individuals. The transmission rate matrix $\beta$ has four entries, corresponding to transmission within and between each of the two age classes,

$$\beta(t) = \begin{pmatrix} \beta_{ss}(t) & \beta_{os}(t) \\ \beta_{os}(t) & \beta_{oo}(t) \end{pmatrix}. \quad (1)$$

We assume $\beta_{os} = \beta_{so}$. The time-dependence of the entries of this matrix are discussed in a separate section below. We denote the disease recovery rate by $\gamma$ (so the mean infectious period is $1/\gamma$). We consider importation of cases from outside the focal region to occur at a constant rate $\iota$ (a proportion $p_s$ of which are assumed occur in school-age children, where $p_s$ is the actual proportion of the population comprised of individuals between the ages of 5 and 18).
The deterministic version of the model equations are

\begin{align}
    \dot{S}_s &= -S_s (\beta_{ss} I_s + \beta_{os} I_o) / N \tag{2a} \\
    \dot{I}_s &= S_s (\beta_{ss} I_s + \beta_{os} I_o) / N - \gamma I_s + \epsilon_p s \\
    \dot{S}_o &= -S_o (\beta_{so} I_s + \beta_{oo} I_o) / N \tag{2c} \\
    \dot{I}_o &= S_o (\beta_{so} I_s + \beta_{oo} I_o) / N - \gamma I_o + \epsilon (1 - p_s), \tag{2d}
\end{align}

where the dot indicates a derivative with respect to time. We simulated the fully stochastic model via the standard fixed-time-step binomial approximation (24; 37) of the exact Gillespie algorithm (44), using the above equations to specify the event rates (45). Laboratory-confirmed case data represent a sample of all cases; we refer to the mean of the sampling distribution as the report ratio, which we denote \( \eta \).

**Time-dependence of the transmission rate**

In our models, each component of the transmission rate matrix (1) was the product of two (potentially) time-dependent factors, one associated with the contact pattern within the group in question (or between the two groups), and the other (identical in all four components) associated with seasonal variation in transmission caused by environmental factors. Thus, for example, the transmission rate among schoolchildren was

\[ \beta_{ss}(t) = \beta_{ss}^*(t) f(t), \tag{3} \]

where \( \beta_{ss}^*(t) \) is a step function that changes on dates when school closes or opens and \( f(t) \) is the seasonal forcing factor (which is the same in all four entries of \( \beta \)).

In the case of sinusoidal seasonality,

\[ f(t) = 1 + \alpha \cos 2\pi (t - t_0 - \phi), \tag{4} \]

where \( \alpha \) and \( \phi \) (both to be fitted) are the amplitude and phase of seasonality, respectively, and \( t_0 \) is a convenient reference time (the beginning of January 2009).

In the case of seasonality induced by temperature changes,

\[ f(t) = 1 - \alpha T(t), \tag{5} \]

where \( T(t) \) is the air temperature at time \( t \) and the seasonal amplitude \( \alpha \) (to be fitted) determines the magnitude of the effect of temperature on influenza transmission.

We considered two possible functional responses of transmission rate to absolute humidity \( AH(t) \) (our calculation of \( AH(t) \) itself is detailed in the next section). The first form is motivated by recent indications (33; 34) that influenza survival is an exponential function of absolute humidity, suggesting

\[ f(t) = 1 + a e^{-bAH(t)}, \tag{6} \]

where \( a \) and \( b \) are parameters to be fitted. The second form attempts to be as simple as possible (and more similar to the other two forms of seasonality considered above), positing a linear relation between absolute humidity and seasonal forcing of transmission rate,

\[ f(t) = 1 - \alpha AH(t). \tag{7} \]

Here, \( \alpha \) is a parameter to be fitted.
Temperature

The top panel of Appendix Figure 1 shows the daily average temperature in Calgary and Edmonton in 2009 (the middle panel shows the daily time series of total confirmed pH1N1 cases for reference). When fitting models to all the case data (for the province as a whole) we used the daily average of the plotted temperatures in Calgary and Edmonton.

Calculation of absolute humidity

We estimated vapor pressure \( v \) from dewpoint temperature \( T_D \) using Teten’s equation (46):

\[
v = a10^{bT_D/(c+T_D)}.
\]

We obtained parameter values from Ref. (47): \( a = 611 \text{ Pa} \), \( b = 7.5 \), \( c = 237.7 \degree \text{C} \). We verified this estimate by calculating relative humidity (RH) as the ratio of vapor pressure to saturated vapor pressure (estimated by applying Teten’s equation to the observed temperature) and comparing to the RH values published by Environment Canada.

Absolute humidity (AH) is the density of water vapor in moist air \((\text{kg/m}^3)\). It can be expressed as

\[
AH = \frac{v}{T \cdot R_w},
\]

where \( v \) is the vapor pressure, \( T \) is the air temperature in degrees Kelvin and \( R_w \) is the gas constant for water vapor, \( R_w = 461.5 \text{ JK/kg} \).

We used equation (9) to calculate hourly AH in Calgary and Edmonton from April to December 2009. From these hourly estimates we computed daily average AH in each city (shown in the bottom panel of Appendix Figure 1). We took the average of the AH and temperature in each of the two major cities to represent Alberta as a whole.

Parameter estimation

Table 1 reports point estimates and 95% confidence intervals for the parameters of our best-fit model. Our model formulation above is not expressed directly in terms of these parameters, so we comment here on the relationships between the model parameters in our formulation above and the (more easily interpretable) parameters that we chose to estimate.

If the recovery rate is \( \gamma \) then the mean infectious period is \( 1/\gamma \).

The subgroup reproductive number \( R_{xy} \) is the expected number of cases in subgroup \( y \) caused by an infectious individual in group \( x \) in a susceptible population. This is given by the product of the transmission rate \( \beta_{xy} \), the mean infectious period \( 1/\gamma \) and the proportion of the population in group \( y \), \( p_y = N_y/N \). We estimated the proportion of school-age individuals (ages 5–18) to be \( p_s = 0.1876 \) for all of the subpopulations we considered (48).

We estimate the (aggregate) basic reproductive number \( R_0 \) by using the time-averaged (49) transmission matrix \( \beta(t) \) [equation (1)] to calculate the matrix \( R_{xy} \), and then calculating the leading eigenvalue of this matrix.

All the parameter estimates we report are based on models with a time step of \( \Delta t = 1 \) day. We repeated fits with \( \Delta t = 1/2 \) day and \( \Delta t = 1/4 \) day and found only slight differences. For example, with our best-fit model, our point estimate and 95% confidence interval for the proportional reduction in transmission induced by closing schools was 0.63 [0.43, 0.84] with \( \Delta t = 1 \) day, 0.66 [0.46, 0.95] with \( \Delta t = 1/2 \) day, and 0.66 [0.45, 0.95] with \( \Delta t = 1/4 \) day.
**Intensive testing during the initial phase**

The number of specimens submitted to ProvLab during the first few weeks of the first wave of the pandemic was about four times as large as the number submitted during the equivalent period at the beginning of the second wave, and is not representative of confirmed cases (see top panel of Figure 1). To account for a period of intensive testing during the initial period of concern about the pandemic, we allowed the report ratio to be a step function,

\[
\eta(t) = \begin{cases} 
\eta_0 & t < t_1 \\
\eta_1 & t \geq t_1 
\end{cases}
\]  

Here, \( t_1 \) is the date on which intensive testing ceased, \( \eta_0 \) is the report ratio during the intensive testing period, and \( \eta_1 \) is the report ratio subsequently.

**Comparison of models**

We compared models using the sample-size corrected (second-order) Akaike Information Criterion (50; 38; 51) (AICₖ), which can be written

\[
\text{AIC}_c = -2 \max_\theta \ell(\theta) + \frac{2N_{\text{par}}N_{\text{samp}}}{N_{\text{samp}} - N_{\text{par}} - 1}.
\]

Here, \( \theta \) is the vector of parameters, \( \ell \) refers to the log-likelihood function, \( N_{\text{par}} \) is the number of parameters (dimension of \( \theta \)) and \( N_{\text{samp}} \) is the sample size. We used weekly bi-variate data over 27 weeks (weekly sums of reported cases in the two age classes), so our sample size was \( N_{\text{samp}} = 54 \). AICₖ as defined above is always positive (likelihoods are probabilities, so \( -\infty < \ell(\theta) \leq 0 \), and \( N_{\text{samp}} > N_{\text{par}} + 1 \)); lower values of AICₖ correspond to better models.

**Initial conditions**

Initial conditions (the initial numbers of susceptibles and infecteds in each age class) were treated as free parameters in our models and hence were included in the total number of free parameters when calculating AICₖ. However, we assumed that the initial proportion susceptible was at least 90% in schoolchildren and at least 75% in others. This assumption is plausible given that pH1N1 is believed to have invaded the human population for the first time in 2009; in addition, it helps to address parameter identifiability issues, i.e., anticorrelation between estimates of basic reproductive number \( R_0 \) and the initial number of susceptibles. Although this assumption may affect the scale of \( R_0 \) and the reporting ratio, we do not expect it to affect the temporal changes that are the focus of this paper.

**Sensitivity analyses**

We investigated the robustness of our parameter estimates by examining 20 different model formulations. The parameter estimates listed in Table 1 correspond to the best model according to AICₖ (which was the same model for each city and for the province as a whole), namely the model with

- specified mean infectious period of 4.5 days;
- specified dates on which the transmission rate among schoolchildren dropped (22 June 2009, the cohort-weighted mean date on which schools actually closed in Alberta) and
rose (27 or 31 August 2009, in Calgary and Edmonton, respectively; 29 Aug 2009 for the province as a whole);

- two fitted values of baseline transmission rate in schoolchildren (i.e., the same value of $\beta^*$ as defined in equation (3) before and after the summer and another value of $\beta^*$ during the summer holiday);

- seasonal variation in transmission rate induced by temperature changes (equation (5));

- no latent period (i.e., SIR rather than SEIR);

- out-of-city importation of cases into Edmonton, but not Calgary; out-of-province importation of cases in Alberta as a whole;

- reporting rate high during the intensive testing period and low after that, and identical for both age classes.

In our best model, we fixed the mean infectious period and school closing/opening dates. We also challenged our model by fitting the mean infectious period and school closing/opening dates and the estimates are summarized in Appendix Figure 2. The estimated mean infectious period is close to the fixed value of 4.5 days. The estimated school closing/opening dates match the true values in Alberta as a whole and in Calgary; the mismatch in Edmonton is likely a consequence of the relatively larger effect of demographic stochasticity in this smaller population. With the exception of the model with unspecified mean infectious period, and the model with unspecified school dates, all our model formulations can be obtained from the best model by relaxing the above constraints in various ways (e.g., using a simple sinusoid or absolute humidity as the basis of seasonal forcing, as discussed above). Most of these models fit the data considerably worse than the model for which Table 1 reports parameter estimates, though some were only marginally worse (e.g., including a latent class). Each of the models we investigated is listed together with the associated AICc in Appendix Tables 1, 2 and 3.

**Validation of parameter estimation methodology**

The methodology we have used (e.g., the freely available R package POMP) has been extensively tested in previous work (25; 51; 24). In Appendix Figure 6, we validate our approach by applying it to 256 realizations of our best-fit model and checking that the distribution of estimates for each parameter is consistent with the correct value (marked with a red vertical line).
Appendix Table 1: Comparison of pH1N1 transmission models for Calgary. MLL refers to the maximum log likelihood. $N_{\text{par}}$ is the number of free parameters in the model. AICc is Akaike’s Information Criterion, corrected for small sample sizes. School indicates whether or not transmission is affected by school being in session (either with the dates of summer closure fixed at their known values or free to be fitted). Several types of seasonal forcing were considered, as described in the Methods section (functions of temperature T, absolute humidity AH, or a simple sinusoid). I.P. denotes the mean infectious period, which was either fixed at 4.5 days or free to be fitted in the SIR model (the SEIR model included a fixed mean latent period of 1.5 days and a fixed mean infectious period of 3 days). The fitted Reporting Rate was either identical for both age classes (same), distinct for each age class (different), or distinct for each age class and different in the summer for school-age children (diff summer). In models that allowed for intensive testing, the Reporting Rate was fitted separately before and after a (fitted) date on which intensive testing ceased. Importation, when included, involved a constant (fitted) rate of cases from elsewhere.

<table>
<thead>
<tr>
<th>Model</th>
<th>MLL</th>
<th>$N_{\text{par}}$</th>
<th>AICc</th>
<th>School (date)</th>
<th>Seasonal forcing</th>
<th>I.P.</th>
<th>Reporting rate</th>
<th>Importation</th>
<th>Early intensive testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SIR</td>
<td>-150.21</td>
<td>12</td>
<td>332.03</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>SIR</td>
<td>-154.97</td>
<td>10</td>
<td>335.06</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>SIR</td>
<td>-154.49</td>
<td>11</td>
<td>337.26</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>4</td>
<td>SIR</td>
<td>-154.52</td>
<td>11</td>
<td>337.32</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>free</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>5</td>
<td>SIR</td>
<td>-153.29</td>
<td>12</td>
<td>338.20</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>diff summer</td>
<td>No</td>
</tr>
<tr>
<td>6</td>
<td>SIR</td>
<td>-153.32</td>
<td>12</td>
<td>338.24</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>7</td>
<td>SIR</td>
<td>-155.10</td>
<td>11</td>
<td>338.48</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>diffuse</td>
<td>No</td>
</tr>
<tr>
<td>8</td>
<td>SIR</td>
<td>-149.86</td>
<td>14</td>
<td>338.49</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>SIR</td>
<td>-153.64</td>
<td>12</td>
<td>338.88</td>
<td>Yes (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>SIR</td>
<td>-158.49</td>
<td>10</td>
<td>342.09</td>
<td>Yes (fixed)</td>
<td>1+a*AH</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>11</td>
<td>SIR</td>
<td>-158.24</td>
<td>11</td>
<td>344.76</td>
<td>Yes (fixed)</td>
<td>a<em>exp(-b</em>AH)</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>SIR</td>
<td>-158.42</td>
<td>11</td>
<td>345.12</td>
<td>Yes (fixed)</td>
<td>1+a*AH</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>SIR</td>
<td>-158.75</td>
<td>11</td>
<td>345.79</td>
<td>Yes (fixed)</td>
<td>1+a*sin(t-b)</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>14</td>
<td>SIR</td>
<td>-158.31</td>
<td>12</td>
<td>348.23</td>
<td>Yes (fixed)</td>
<td>1+a*sin(t-b)</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>SIR</td>
<td>-163.68</td>
<td>9</td>
<td>349.46</td>
<td>Yes (fixed)</td>
<td>Constant</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>16</td>
<td>SIR</td>
<td>-157.44</td>
<td>13</td>
<td>349.98</td>
<td>Yes (fixed)</td>
<td>1+a*AH</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>SIR</td>
<td>-158.45</td>
<td>13</td>
<td>352.01</td>
<td>Yes (fixed)</td>
<td>1+a*sin(t-b)</td>
<td>fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>18</td>
<td>SIR</td>
<td>-164.97</td>
<td>11</td>
<td>358.22</td>
<td>No (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>diff summer</td>
<td>No</td>
</tr>
<tr>
<td>19</td>
<td>SIR</td>
<td>-168.47</td>
<td>10</td>
<td>362.06</td>
<td>No (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>different</td>
<td>No</td>
</tr>
<tr>
<td>20</td>
<td>SIR</td>
<td>-167.12</td>
<td>12</td>
<td>365.84</td>
<td>No (fixed)</td>
<td>1+a*T</td>
<td>fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
</tbody>
</table>
## Appendix Table 2: Comparison of pH1N1 transmission models for Edmonton. See caption to Appendix Table 1 for details.

<table>
<thead>
<tr>
<th>Model</th>
<th>MLL</th>
<th>$N_{par}$</th>
<th>AICc</th>
<th>School (date)</th>
<th>Seasonal forcing</th>
<th>I.P. Reporting rate</th>
<th>Importation</th>
<th>Early intensive testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SIR</td>
<td>-145.34</td>
<td>13</td>
<td>325.79</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>SIR</td>
<td>-150.17</td>
<td>11</td>
<td>328.63</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>SIR</td>
<td>-147.09</td>
<td>13</td>
<td>329.27</td>
<td>Yes(fixed)</td>
<td>1+a*sin(t-b) fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>SIR</td>
<td>-149.69</td>
<td>12</td>
<td>331.00</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>different</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>SEIR</td>
<td>-149.98</td>
<td>12</td>
<td>331.57</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>6</td>
<td>SIR</td>
<td>-146.41</td>
<td>14</td>
<td>331.58</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>different</td>
<td>Yes</td>
</tr>
<tr>
<td>7</td>
<td>SIR</td>
<td>-150.13</td>
<td>12</td>
<td>331.87</td>
<td>Yes(fixed)</td>
<td>1+a*T free</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>8</td>
<td>SIR</td>
<td>-148.78</td>
<td>13</td>
<td>332.67</td>
<td>Yes(free)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>9</td>
<td>SIR</td>
<td>-152.36</td>
<td>11</td>
<td>333.01</td>
<td>Yes(fixed)</td>
<td>1+a*sin(t-b) fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>10</td>
<td>SIR</td>
<td>-149.46</td>
<td>13</td>
<td>334.02</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>diff summer</td>
<td>Yes</td>
</tr>
<tr>
<td>11</td>
<td>SIR</td>
<td>-155.28</td>
<td>10</td>
<td>335.68</td>
<td>Yes(fixed)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>12</td>
<td>SIR</td>
<td>-152.32</td>
<td>12</td>
<td>336.25</td>
<td>Yes(fixed)</td>
<td>1+a*sin(t-b) fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>13</td>
<td>SIR</td>
<td>-158.69</td>
<td>11</td>
<td>345.67</td>
<td>Yes(fixed)</td>
<td>1+a*AH fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>14</td>
<td>SIR</td>
<td>-157.70</td>
<td>13</td>
<td>347.01</td>
<td>Yes(fixed)</td>
<td>a<em>exp(-b</em>AH) fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>15</td>
<td>SIR</td>
<td>-157.46</td>
<td>13</td>
<td>350.02</td>
<td>Yes(fixed)</td>
<td>1+a*AH fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>16</td>
<td>SIR</td>
<td>-160.39</td>
<td>12</td>
<td>352.39</td>
<td>No(fixed)</td>
<td>1+a*T fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
<tr>
<td>17</td>
<td>SIR</td>
<td>-163.82</td>
<td>10</td>
<td>352.75</td>
<td>Yes(fixed)</td>
<td>1+a*AH fixed</td>
<td>same</td>
<td>No</td>
</tr>
<tr>
<td>18</td>
<td>SIR</td>
<td>-161.28</td>
<td>12</td>
<td>354.17</td>
<td>No(fixed)</td>
<td>1+a*T fixed</td>
<td>diff summer</td>
<td>Yes</td>
</tr>
<tr>
<td>19</td>
<td>SIR</td>
<td>-159.94</td>
<td>13</td>
<td>354.99</td>
<td>No(fixed)</td>
<td>1+a*T fixed</td>
<td>different</td>
<td>Yes</td>
</tr>
<tr>
<td>20</td>
<td>SIR</td>
<td>-167.75</td>
<td>10</td>
<td>360.62</td>
<td>Yes(fixed)</td>
<td>Constant fixed</td>
<td>same</td>
<td>Yes</td>
</tr>
</tbody>
</table>
### Appendix Table 3: Comparison of pH1N1 transmission models for Alberta as a whole. See caption to Appendix Table 1 for details.

<table>
<thead>
<tr>
<th>Model</th>
<th>MLL</th>
<th>$N_{\text{par}}$</th>
<th>$\Delta\text{IC}_c$</th>
<th>School (date)</th>
<th>Seasonal forcing</th>
<th>I.P. Reporting rate</th>
<th>Importation</th>
<th>Early intensive testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>SIR</td>
<td>-193.89</td>
<td>13</td>
<td>422.88</td>
<td>Yes (fixed)</td>
<td>1+$a^T$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>2</td>
<td>SIR</td>
<td>-193.25</td>
<td>14</td>
<td>425.27</td>
<td>Yes (free)</td>
<td>1+$a^T$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>3</td>
<td>SIR</td>
<td>-193.55</td>
<td>14</td>
<td>425.88</td>
<td>Yes (fixed)</td>
<td>1+$a^T$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>4</td>
<td>SIR</td>
<td>-199.31</td>
<td>11</td>
<td>426.91</td>
<td>Yes (fixed)</td>
<td>1+$a^T$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>5</td>
<td>SIR</td>
<td>-201.35</td>
<td>14</td>
<td>441.47</td>
<td>Yes (fixed)</td>
<td>$a^\exp(-b^A^H)$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>6</td>
<td>SIR</td>
<td>-204.00</td>
<td>13</td>
<td>443.11</td>
<td>Yes (fixed)</td>
<td>1+$a^\sin(t-b)$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>7</td>
<td>SIR</td>
<td>-204.28</td>
<td>14</td>
<td>447.33</td>
<td>Yes (fixed)</td>
<td>$a^\exp(-b^A^H)$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>8</td>
<td>SIR</td>
<td>-210.14</td>
<td>13</td>
<td>455.39</td>
<td>Yes (fixed)</td>
<td>1+$a^A^H$</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>9</td>
<td>SIR</td>
<td>-224.47</td>
<td>12</td>
<td>480.55</td>
<td>Yes (fixed)</td>
<td>constant</td>
<td>fixed</td>
<td>same</td>
</tr>
<tr>
<td>10</td>
<td>SIR</td>
<td>-225.97</td>
<td>12</td>
<td>483.54</td>
<td>No (fixed)</td>
<td>1+$a^T$</td>
<td>fixed</td>
<td>same</td>
</tr>
</tbody>
</table>
Appendix Figure 1: Daily average temperature (top panel), daily confirmed pH1N1 cases (middle panel) and daily absolute humidity (bottom panel) in Calgary and Edmonton in 2009. The grey shaded region in the middle panel highlights the time period when virological testing restrictions were in place. Models were fitted only to data that preceded the initiation of testing restrictions.
Appendix Figure 2: Comparison of fitted vs fixed value of the mean infectious period (column 1) and fitted vs known values of the dates on which the transmission rate dropped (column 2) and rose (column 3) as a result of schools closing for the summer in Calgary (row 1), Edmonton (row 2) and the province of Alberta as a whole (row 3). Panels (a,d,g) in column 1 show the likelihood profile for the mean infectious period, estimated using our best-fit model (Table 1), but allowing the mean infectious period to be fitted. The 95% confidence lies between the two dotted vertical blue lines, while the a priori fixed value of 4.5 days is indicated with a solid vertical red line. Similarly, columns 2 and 3 show likelihood profiles for fitted vs actual school closing and opening dates.
Appendix Figure 3: Comparison of pH1N1 data for Calgary (blue) with simulations (box plots are based on 1000 realizations of our best-fit model) as specified in Appendix Table 1. The left panels show data and simulation results for school-age children and the right panels show the corresponding data and simulations results for the rest of the population. The top panels compare the data with simulations of our best-fit model (Appendix Table 1). The bottom panels show what the model predicts if schools had been left open throughout the summer.
Appendix Figure 4: Comparison of pH1N1 data for the city of Edmonton (blue) with simulations (box plots). See caption to Appendix Figure 3 for details.
Appendix Figure 5: Age distribution of laboratory-confirmed cases of pH1N1 in Alberta in 2009. Top panel: Cases. Bottom panel: Case rates (per 100,000 population, based on data from Ref. (48)).
Appendix Figure 6: Validation of parameter estimation methodology. With parameters set to their MLE values, we generated 256 stochastic realizations of our best-fit model. Using POMP, we then refitted our model to each of these 256 simulations. For each parameter, a histogram shows the distribution of MLEs and a red vertical line shows the true value (which always lies close to the centre of the distribution of MLEs). The final (bottom right) panel shows the distribution of the maximum log likelihood (MLL) for each simulation compared with the MLL for the observed data (red vertical line). While the MLL is not a parameter and should not be directly compared for different data sets, the MLL associated with a given model fitted to observed data should be similar to the MLL obtained by fitting to simulations of the same model (as we find here).