Pre-symptomatic Influenza Transmission, Surveillance, and School Closings: Implications for Novel Influenza A (H1N1)

G.F. Webb\textsuperscript{1} *, Y-H. Hsieh \textsuperscript{2}, J. Wu \textsuperscript{3} and M.J. Blaser\textsuperscript{4}

\textsuperscript{1} Department of Mathematics, Vanderbilt University, Nashville, Tennessee, 37240 USA
\textsuperscript{2} Department of Public Health and Center for Infectious Disease Epidemiology Research, China Medical University, Taichung, Taiwan
\textsuperscript{3} Laboratory for Industrial and Applied Mathematics, Centre for Disease Modeling, Department of Mathematics and Statistics, York University, Toronto, Canada M3J 1P3
\textsuperscript{4} Department of Medicine and Department of Microbiology, New York University School of Medicine, New York, NY 10016, USA

\textbf{Abstract.} Early studies of the novel swine-origin 2009 influenza A (H1N1) epidemic indicate clinical attack rates in children much higher than in adults. Non-medical interventions such as school closings are constrained by their large socio-economic costs. Here we develop a mathematical model to ascertain the roles of pre-symptomatic influenza transmission as well as symptoms surveillance of children to assess the utility of school closures. Our model analysis indicates that school closings are advisable when pre-symptomatic transmission is significant or when removal of symptomatic children is inefficient. Our objective is to provide a rational basis for school closings decisions dependent on virulence characteristics and local surveillance implementation, applicable to the current epidemic and future epidemics.

\textbf{Key words:} influenza, symptoms surveillance, pre-symptomatic, age of infection model, school closing policy

\textbf{AMS subject classification:} 92B05, 62P10, 37N25

\*Corresponding author. E-mail: glenn.f.webb@vanderbilt.edu
1. Introduction

Prior statistical analyses and models of school closings address several issues, including their effects on attack rates in school children and in the community ([4],[7]), buying time for a strain-specific vaccine ([32]), the impact of the timing and duration of school closure and flow-on effects on other social contacts ([31]), and household responses and costs ([24],[40]). Our model is distinguished from these in its focus on the contribution of pre-symptomatic (or early symptomatic) children to disease transmission. Asymptomatic influenza infection occurs regularly ([2],[16],[29]), accounting for up to one-third of all infections ([33],[34],[37],[38],[39]). Similarly, early (pre-symptomatic) infectivity is well documented ([18]). One well-investigated case of pre-symptomatic influenza transmission ([42]) is partially responsible for World Health Organization recommendations on non-pharmaceutical measures to prevent such transmission in pandemic settings ([45]). In the structured settings of schools, pre-symptomatic infection is a particularly important driver of school closure policies.

Several approaches have been used to model transmission in asymptomatic or subclinically infected persons ([1],[15],[21],[23],[27],[28],[44]). Our approach is a deterministic model, which uses partial differential equations incorporating an age of infection variable beginning at the moment of inoculation and following the disease course through its phases in each individual. In particular, the age of infection variable allows efficient tracking of both (i) the pre-infectious and infectious phases, and (ii) the pre-symptomatic and symptomatic phases. We investigate the critical case that (i) and (ii) are not coincident, but overlap during which time children are infectious, but are not yet symptomatic, or for practical purposes, cannot be identified as symptomatic.

We examine epidemic outcome scenarios by varying two key model parameters: (i) the overlap period of the pre-symptomatic and infectious phases from 0 to 24 h, and (ii) the efficiency of removing symptomatic children from the school environment from 0 to 100% per day. This second parameter is available to school and public health officials for surveillance, identification, isolation, send-home, and stay-at-home policies ([4]); its interpretation is context-dependent, but meaningful in recommending and implementing policies.

2. The SEIR Age of Infection Model

The model (Figure 1) consists of school children divided into susceptible $S(t)$, exposed (latent) $E(t)$ infectious $I(t)$, and removed classes $R(t)$ at time $t$, with the key feature that infected individuals are structured by a continuous variable $a$ corresponding to time since inoculation, and the latent, infectious, pre-symptomatic, and symptomatic stages of the disease are modeled by the disease age of infected individuals. The model equations and analysis are given in the Appendix. The assumptions of the model are that:

(1) The population consists of school children; initially there are 10,000 susceptible children with a small number infected. The model is of standard incidence form ([19]), so the results are valid for any number of initial susceptibles with a given parameterization. The model is thus
Figure 1: Schematic of the SEIR compartment model. Susceptible children become infected, immediately enter the exposed (latent) class in which they are not yet infectious, and then enter the infectious class. Symptomatic children may exit to the removed class, where they are no longer susceptible.

(2) The dynamics of the model are localized to the population of school children. Infection of children outside of the school environment is not incorporated specifically, but can be considered included.

(3) A threshold of the infectious phase, denoted by $T_i$, is the time during the disease course at which infected children become infectious (Figure 2). The infectious period lasts for time $F_i$, so that children are infectious from $T_i$ to $T_i + F_i$.

(4) A threshold of the symptomatic phase, denoted by $T_s$, is the time during the disease course at which infected children manifest identifiable systemic or respiratory symptoms (Figure 2). The symptomatic phase lasts from $T_s$ to $T_i + F_i$. We assume that $T_s \geq T_i$.

(5) The removal of symptomatic infected children is $R_{sym}\%$ per day after the threshold of the symptomatic period $T_s$, but 0% per day before $T_s$. This percentage takes into account time lags in identifying presentation of symptoms and isolation of children from the school, both at home by parents and at school by officials.

The thresholds in these assumptions should be viewed as effective average values for typical infected individuals. Their designations are not precise, but rather simplified estimates based on available data such as viral shedding and immune response levels ([18]). Viral shedding before symptoms on-set appears higher in children than in adults ([11],[17],[43]). For isolation of infected children in schools, identification of symptoms is largely equated to fever and coughing, which are problematic as markers at a practical level. Carrat et al. ([5]) showed that viral shedding and total symptoms score overlap over the disease course in adult volunteer studies. However, those data show an initial advance of viral shedding (peak at 2.0 days) compared to total symptoms score.
Figure 2: The age of infection-dependent transmission rate $\beta(a)$. Infectiousness begins at the threshold $T_i$, peaks, and then falls to near 0 at time $T_i + F_i$. The symptomatic period begins at time $T_s$ and lasts until time $T_i + F_i$. The period of pre-symptomatic infectiousness is the gray bar between infection $T_i$ and $T_s$. In the baseline simulation the period of pre-symptomatic infectiousness is 0.25 days.

(peak at 3.0 days). Since the viral shedding measurements range over many orders of magnitude, assigning threshold values is justified. Although the efficiency of viral shedding for transmission from asymptomatic hosts is unknown ([9]), the contribution of pre-symptomatic infectious children to transmission may be critical in school settings. Formulas for the epidemic reproduction number $R_0$ and the daily and cumulative epidemic attack percentages $AR$ under these assumptions are given in the Appendix.

3. The Baseline Model

We specify the parameters of the model at baseline (Table 1) with values supported by studies of past influenza epidemics ([5],[6],[9],[14]) as well as recent reports of the current epidemic, with extrapolation to the context of schoolchildren ([13]). An advantage of this approach is that its parameterization is relatively manageable within a range of acceptable values, which can be adjusted as new information develops.

We use the adult volunteer data in [5] as the basis for our baseline transmission parameter $\beta(a)$ (Figure 2). The choice of $\beta(a)$ is also determined by its compatibility with an epidemic reproduction number $R_0 = 1.29$ and an attack rate = 41%, that are within the range of values reported for the early outbreak in Mexico ([13]) and consistent with past influenza outbreaks with a new strain circulating in the population. The speed at which influenza spreads is not only dependent on a high $R_0$, but also on the serial interval ([26],[30]). The choice of 70%/day removal rate of children assumes a relatively efficient process for their identification and being sent home. Clearly, this percentage will vary considerably, which provides an opportunity for social distancing intervention strategy. The baseline simulation begins with 10 infected students equi-distributed in the age of
Table 1: Baseline parameters of the model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Parameter description</th>
<th>Baseline value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_i$</td>
<td>infectiousness threshold</td>
<td>0.75 days [5]</td>
</tr>
<tr>
<td>$T_s$</td>
<td>symptoms threshold</td>
<td>1.0 day [5]</td>
</tr>
<tr>
<td>$F_i$</td>
<td>length of infectiousness period</td>
<td>4.0 days [5]</td>
</tr>
<tr>
<td>$\beta(a)$</td>
<td>infection age transmission rate</td>
<td>$\beta(a) = \begin{cases} \frac{0.0}{10^{-3}(a - T_i)}e^{-2.0(a - T_i)}, &amp; a \leq T_i \ -0.01R_{sym}, &amp; a &gt; T_i \end{cases}$</td>
</tr>
<tr>
<td>$R_{sym}$</td>
<td>removal of symptomatic children</td>
<td>70% per day</td>
</tr>
<tr>
<td>$\mu(a)$</td>
<td>removal rate of infected children</td>
<td>$\mu(a) = \begin{cases} 0.0, &amp; a \leq T_s \ -log[1.0 - .01R_{sym}], &amp; a &gt; T_s \end{cases}$</td>
</tr>
</tbody>
</table>

infection range $[0, T_i + F_i]$ in a population of 10,000 susceptible students. The baseline simulation is given in Figure 3 and shows an epidemic that lasts approximately 60 days, strongly increases around day 25, and develops with characteristics similar to other influenza outbreaks in school-age populations ([7]).

Figure 3: The susceptible population $S(t)$ (Panel A), the cumulative number of cases $C(t)$ (Panel B), the rates of change per day of the exposed population $E(t)$ (Panel C) and infectious population $I(t)$ (Panel D) in the baseline simulation. The epidemic begins with the introduction of 10 infected children into the school population. The duration of the epidemic is approximately 60 days.

4. Simulations of the Model with Variable Pre-symptomatic Infectious Periods and Symptomatic Children Removal Rates

We vary the length of the pre-symptomatic infectious period $T_s - T_i$ from 0.0 to 1.0 days, by holding the threshold of infectiousness fixed at the baseline value $T_i = 0.75$ days and varying the threshold of symptoms $T_s$ from 0.75 days to 1.75 days, and we vary the removal rate of symptomatic children.
from $R_{sym} = 0\%$ to 100\% per day to gain insight into the influence of these elements on the epidemic dynamics. The baseline parameters are otherwise held constant, and we consider the epidemic reproductive number $R_0$ (Figure 4) and the epidemic attack ratio $AR$ (Figure 5) as a function of these two parameters. It is intuitive that a short pre-symptomatic infectious period or an efficient removal of symptomatic children will depress the attack ratio, but it is surprising that both variables must be within a small range to achieve this effect. For example, if the pre-symptomatic infectious period is $> 12$ h, then the removal of symptomatic children must be $> 80\%$ per day to significantly reduce the attack ratio. If the removal of symptomatic children is $< 50\%$ per day, then only if the pre-symptomatic infectious period is $< 6$ h, will the attack ratio be significantly reduced.

![Figure 4: $R_0$ as a function of the pre-symptomatic infectious period $0.0 < T_s - T_i < 1.0$ days, $T_i = 0.75$ days, $0.75 < T_s < 1.75$ days, and the removal rate of symptomatic children $0.0\% < R_{sym} < 100\%$. The height of the red plane is at $R_0$ threshold value 1.0. The green dot is the value of $R_0 = 1.29$ at the baseline parameter values ($T_i = 0.75$ days, $T_s = 1.0$ day, and $R_{sym} = 70\%$ per day).](image)

An illustration of how the simulations in Figure 5 can guide school closing policy is as follows: if the disease carries significant morbidity and mortality, the decision to close schools is advisable when the pairing \{ $T_s - T_i$, $R_{sym}\%$ \} yields an attack ratio $> 60\%$. In such case the attack ratio is not significantly affected by the existing surveillance and removal policies of symptomatic children. However, if the pairing yields an attack ratio $< 30\%$, then surveillance and removal measures can have substantial impact in reducing the attack ratios, and can be further enhanced. Other percent values can be used by officials as acceptable (e.g. 10\% rather than 30\%) or unacceptable (e.g. 40\% rather than 60\%) to predict the attack ratio based on the initialization and parameterization of the
Figure 5: The attack ratio $AR$ as a function of the pre-symptomatic infectious period and the removal rate of symptomatic children as in Figure 4. The height of the green plane is 30% and the red plane 60%. Attack ratios corresponding to pairings of the control variables above the red plane are not effectively controlled by the removal rates of symptomatic children, whereas those below the green plane are. Pairings between the two planes are marginal for school closings, but enhanced surveillance and removal measures may reduce attack ratios to acceptable values. The yellow dot is the value of the removal rate = 41% at the baseline parameter values.

5. Discussion

School closings policy is a critical issue for epidemics such as the 2009 novel influenza A (H1N1) epidemic. This outbreak has highly impacted younger age groups in Mexico, the US, and elsewhere ([13],[36]), and has potential for even greater impact on youths in future waves ([10],[22]). A key epidemiological factor for younger age cohorts is the importance of asymptomatic and pre-symptomatic individuals, since these cases may evade parental and institutional identification and isolation but still contribute significantly to transmission, and since surveillance and removal policies in schools are necessarily linked to manifestation of symptoms.

We provide a mathematical model that can guide public health and school officials in ascertaining when school closures are warranted, by quantifying the tension between asymptomatic transmission and efficiency of surveillance methods. For school environments the model provides a rationale for improved surveillance methods. For example, non-invasive temperature monitoring with real-time thermal sensors, as in current use in some airports ([41]), offers a potential advance for improved surveillance and rapid isolation of infected students. Another important advance for
surveillance methods is to implement antibody-based testing, to see who is immune and no longer susceptible ([35]). In the long term, the benefits of implementing such new surveillance technologies in schools should be evaluated. Near term, informed decisions for school closings will become increasingly important, if the recent spring A (H1N1) epidemic in North America and elsewhere in the world is merely a herald wave for a fall-winter wave of the new H1N1 variant, which may occur with greater virulence, as has been widely anticipated ([8],[20]).

Our modeling approach is based on a model that tracks infected individuals through their disease course with a continuum variable corresponding to age of infection. Age of infection models were introduced by Kermack and McKendrick ([25]) in 1927 and have been used extensively to model epidemics (a review is given by Brauer ([3]). Our age of infection model focuses on features important to school closings, namely the age of infection tracked symptomatic and infectious phases of infectives and the efficiency of surveillance and isolation controls. These features are readily recognizable to public health officials and relevant to their recommendations. Our model quantifies the relationships of these features and predicts the consequences of control measures. In Fraser et al. [12], epidemic outbreak control interventions were studied in linear models based on time since infection, with emphasis on the overlap of the pre-symptomatic and infectious periods. The basic model in [12] can be recast as a well-posed age of infection model similar to ours, but without incorporation of the susceptible population and without any removal of the infected population.

We argue that a useful indicator derived from an epidemic model of school closings is not only the epidemic reproduction number $R_0$ (which is defined abstractly as the average number of secondary infections per infected individual over an infection lifetime in a given susceptible population context), but also the epidemic attack ratio (or attack rate) $AR$, which is the fraction or percentage of the susceptible population ultimately infected. In our model $AR$ depends on the initial value of the infected population, whereas $R_0$ does not. The initialization of the infected population of the epidemic is critical for determination of its long-term behavior. In Figure 6 we illustrate the dependence of $AR$ on the initial infected population in our model, which illustrates that the size of the initial infected population strongly influences the ultimate attack ratio. In addition, school re-opening policies can be evaluated by a re-set of the initial conditions, which again underscores their importance in modeling approaches. Our model provides a means to identify $AR$ based on parameterizations that are recognizable to public health officials, namely, the length, severity, and overlap of the symptomatic and infectious periods (see the Appendix). Beyond $R_0$ and $AR$, our modeling approach quantifies through simulations the full course of an epidemic in a way also recognizable to public health officials, namely the consequences of policies that identify and isolate infected children from the school.

Our modeling methodology is deterministic, with a small number of parameters, and can be completely analyzed for parameter sensitivity and parameter dependent outcomes. We provide online open access to our simulations codes, which can be readily adapted to specific school settings, at

http://www.med.nyu.edu/medicine/labs/blaserlab/h1n1-math-model.html

Our model is further applicable to other epidemic-at-risk populations, including those at large
Figure 6: The dependence of the attack ratio $AR$ on the initial value through the formula (6.11) while $AR_i$ is given by formula (6.10) (see the Appendix). The distributed parameters $\beta(a)$ and $\mu(a)$ are at baseline values and the initial condition $i(a, 0) = i_0(a)$ is varied to give $AR_i$ via formula (6.11). We vary $i_0(a)$ from 0 to 250 for the first two days and 0 otherwise. With value 150 we have $AR_i$ approximately 0.2, which gives an attack ratio 0.62. The initial level of the infected population strongly influences the attack ratio, especially at lower levels corresponding to outbreak. School re-opening policies can be guided by re-setting the initialization to a new level after intervals of school closure.

 workplaces, military installations, event (large crowd) gatherings, mass transit sites, and other settings suitable for social distancing policies.

Acknowledgements

This work was supported by NIH GM 63270 (M.J.B.), NSC97-2314-B-039-013-MY3 (Y-H.H.), CRC, GEOIDE, NSERC and MITACS (J.W.), and NIH GM 63270, NSF/NIH GM 83607 (G.F.W.).

6. Appendix

The compartments of the model are given in Figure 1. The model equations are

$$\frac{d}{dt}S(t) = - \left( \int_{T_i}^{T_i+F_i} \beta(a)i(a,t)da \right) S(t), \quad t \geq 0$$  \hspace{1cm} (6.1)
where \( t \) is time, \( a \) is age of infection, \( S(t) \) is the number of susceptibles at time \( t \), \( i(a, t) \) is the density of infectives of infection age \( a \) at time \( t \), and \( R(t) \) is the number of removed infectives at time \( t \). The initial conditions are specified in (6.5). The baseline parameters \( \beta(a), \mu(a), T_i, F_i, T_s, R_{sym} \) are given in Table 1. The period of pre-symptomatic infectiousness is \( T_s - T_i \). At time \( t \), the total numbers of exposed (latent) infectives \( E(t) \), infectious infectives \( I(t) \), pre-symptomatic infectives \( PS(t) \), symptomatic infectives \( SI(t) \), and pre-symptomatic infectious infectives \( PSI(t) \) are

\[
E(t) = \int_0^{T_i} i(a, t)da, \quad I(t) = \int_{T_i}^{T_i+F_i} i(a, t)da, \quad PS(t) = \int_0^{T_s} i(a, t)da,
\]

\[
SI(t) = \int_{T_i}^{T_i+F_i} i(a, t)da, \quad PSI(t) = \int_{T_i}^{T_s} i(a, t)da.
\]

It can be verified directly that the following formula for \( i(a, t) \) gives a generalized solution of (6.2) in terms of \( i(0, t) \):

\[
i(a, t) = \begin{cases} 
  i(0, t-a)e^{-\int_0^a \mu(b)db}, & a < t, \\
  i_0(a-t)e^{-\int_a^{T_i} \mu(b)db}, & a \geq t.
\end{cases}
\]  

(6.6)

Further, from (6.1) and (6.3)

\[
i(0, t) = -\frac{d}{dt} S(t) = \left( \int_{T_i}^{T_s} \beta(a)i(a, t)da \right) e^{-\int_0^t \int_{T_i}^{T_i+F_i} \beta(a)i(a, s)dsda} S_0.
\]  

(6.7)

The proof of the following theorem uses the method of steps to obtain a unique solution to (6.6) and (6.7). The proof of the existence of the solution and its asymptotic behavior is similar to the proof given in ([46]).

**Theorem 1.** Let \( T_i, F_i > 0, \beta, \mu \in L_+^\infty(0, T_i + F_i), S_0 > 0, \) and \( i_0 \in L_+^1(0, T_i + F_i) \). There exists a unique solution to (6.6) and (6.7). If \( \mu(a) \geq \bar{\mu} \) for some \( \bar{\mu} > 0, a \in [0, T_i + F_i] \), then \( \lim_{t \to \infty} S(t) = S_\infty > 0, \lim_{t \to \infty} R(t) = 0, \) and \( \lim_{t \to \infty} I(t) = 0. \)
The formula for the epidemic reproduction number is

\[ R_0 = S(0) \int_{T_i}^{T_i+F_i} \beta(a) e^{-\int_a^\infty \mu(b) db} da. \]  

(6.8)

The formula for the cumulative number of cases \( C(t) \) at time \( t \) is \( \int_0^t i(0, t) dt \), the fraction of the initial susceptible population that has been infected by time \( t \) is \( \int_0^t i(0, t) dt / S(0) \), and the epidemic attack ratio for the entire course of the epidemic is \( AR = \int_0^\infty i(0, t) dt / S(0) \). The total cumulative number of cases during the course of the epidemic is

\[ \int_0^\infty i(0, t) dt = (1 - e^{-\int_0^\infty (AR_i + R_0 AR) dt}) S(0). \]  

(6.9)

where \( AR_i \) is obtained from the infective initial data by the formula

\[ AR_i = \int_{T_i}^{T_i+F_i} \beta(a) \int_0^a i(\hat{a}, 0) e^{-\int_\hat{a}^\infty \mu(b) db} d\hat{a} da. \]  

(6.10)

Thus, from (6.8), (6.9), and (6.10), \( AR \) satisfies uniquely the equation

\[ AR = 1 - e^{-(AR_i + R_0 AR)}. \]  

(6.11)

References


