Estimates of the Prevalence of Pandemic (H1N1) 2009, United States, April–July 2009

Carrie Reed, Frederick J. Angulo, David L. Swerdlow, Marc Lipsitch, Martin I. Meltzer, Daniel Jernigan, and Lyn Finelli

Through July 2009, a total of 43,677 laboratory-confirmed cases of influenza A pandemic (H1N1) 2009 were reported in the United States, which is likely a substantial underestimate of the true number. Correcting for under-ascertainment using a multiplier model, we estimate that 1.8 million–5.7 million cases occurred, including 9,000–21,000 hospitalizations.

Human cases of influenza A pandemic (H1N1) 2009 were first identified in the United States in April 2009 (1,2). By the end of July, >40,000 laboratory-confirmed infections had been reported, representing only a fraction of total cases. Persons with influenza may not be included in reported counts for a variety of reasons, including the following: not all ill persons seek medical care and have a specimen collected, not all specimens are sent to a public health laboratory for confirmatory testing with reverse transcription–PCR (RT-PCR; rapid point-of-care testing cannot differentiate pandemic [H1N1] 2009 from other strains), and not all specimens will give positive results because of the timing of collection or the quality of the specimen. To better estimate the prevalence of pandemic (H1N1) 2009 during April–July 2009 in the United States, we created a simple multiplier model that adjusts for these sources of under-ascertainment.

The Study

Through July 23, 2009, a total of 43,677 laboratory-confirmed infections with pandemic (H1N1) 2009 had been reported in the United States by the 50 states and the District of Columbia, including 5,009 hospitalizations and 302 deaths. To estimate the total number of cases of pandemic (H1N1) 2009, we built a probabilistic multiplier model that adjusts the count of laboratory-confirmed cases for each of the following steps: medical care seeking (A), specimen collection (B), submission of specimens for confirmation (C), laboratory detection of pandemic (H1N1) 2009 (D), and reporting of confirmed cases (E) (Figure). This approach has been used to calculate the underrecognized impact of foodborne illness in the United States (3).

At each step, we identified a range of proportions observed in prior published studies and recent surveys and investigations of pandemic (H1N1) 2009. These include 2 unpublished community surveys on influenza-like illness (ILI) and health-seeking behavior, the 2007 Behavioral Risk Factor Surveillance Survey conducted in 10

Figure. Schematic of the steps involved in adjusting counts of reported cases of pandemic (H1N1) 2009 to estimate total cases.
states and repeated in the same states during May 2009, and field investigations conducted during early outbreaks of pandemic (H1N1) 2009 in Chicago and Delaware (online Technical Appendix, available from www.cdc.gov/EID/content/15/12/2004-Techapp.pdf; [4]). We theorized that, given recommendations for testing, patients hospitalized with pandemic (H1N1) 2009 would more likely have been tested and their cases reported than would outpatients. We therefore stratified our model between hospitalized and nonhospitalized cases (Figure). For hospitalized patients, we used larger estimates of the proportion of specimens collected, tested, and reported, which resulted in smaller multiplier values (Table 1). We also adjusted for the fact that early in the epidemic physicians and health departments were encouraged to collect clinical specimens from all suspect case-patients with ILI and forward them for confirmatory testing with RT-PCR. By May 12, due to the increasing number of cases and the demands on public health laboratories, guidance for confirmatory testing was revised to focus on hospitalized patients. We therefore used a lower estimate for the proportion of specimens collected from patients with mild illness after that date, effectively increasing the multiplier for those patients (Table 1).

Multipliers were calculated as the simple inverses of the proportions at each step. We accounted for variability and uncertainty in model parameters by using a probabilistic (Monte Carlo) approach (built by using SAS version 9.2; SAS Institute, Cary, NC, USA). For each parameter included in the model, we used uniform probability distributions that covered a range of minimum to maximum values, from which the model randomly sampled 10,000 iterations (online Technical Appendix). We generated median, upper, and lower 90% values for the number of total illnesses and hospitalizations.

To further divide estimated cases into age groups, we applied the age distribution of confirmed cases and hospitalizations as reported to the US Centers for Disease Control and Prevention through July 23, 2009 (online Technical Appendix), and calculated overall and age-specific incidence of illness and hospitalization, based on the US Census monthly population estimates for May 2009. We did not have age-specific parameter estimates, and thus did not stratify by age group within the model. This approach may not fully capture differences in the probability of ascertainment by age.

Using this approach, between April and July 2009, we estimate that the median multiplier of reported to estimated cases was 79; that is, every reported case of pandemic (H1N1) 2009 may represent 79 total cases, with a 90% probability range of 47–148, for a median estimate of 3.0 million (range 1.8–5.7 million) symptomatic cases of pandemic (H1N1) 2009 in the United States. Likewise, we estimate that every hospitalized case of pandemic (H1N1) 2009 that was reported may represent a median of 2.7 total hospitalized persons (90% range 1.9–4.3). This represents a median estimate of 14,000 (range 9,000–21,000) hospitalizations (Table 2) and thus an estimated ratio of hospitalizations to total symptomatic cases of 0.45% (range 0.16%–1.2%).

We also estimate that incidence of pandemic (H1N1) 2009 over the first 4 months of the pandemic in the United States ranged from a median of 107/100,000 in persons ≥65 years of age, to 2,196/100,000 in persons 5–24 years of age (Table 2). The incidence of hospitalization was estimated to be highest in young children <5 years of age (median 13.0/100,000, 90% range 8.8–20.2).

Table 1. Model parameters and sources of data included in the model estimating prevalence of pandemic (H1N1) 2009, United States, April–July 2009*

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Observed value</th>
<th>Source</th>
<th>Ranges included in the model, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Proportion of persons with influenza who seek medical care, %</td>
<td>42</td>
<td>2007 BRFSS, 9 states†</td>
<td>42–58</td>
</tr>
<tr>
<td></td>
<td>52–55</td>
<td>2009 ILI survey, 10 states†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>49–58</td>
<td>Delaware university survey</td>
<td></td>
</tr>
<tr>
<td></td>
<td>52</td>
<td>Chicago community survey</td>
<td></td>
</tr>
<tr>
<td>B Proportion of persons seeking care with a specimen collected, %</td>
<td>25</td>
<td>2007 BRFSS, 9 states†</td>
<td>19–34</td>
</tr>
<tr>
<td></td>
<td>22–28</td>
<td>2009 ILI survey, 10 states†</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19–34</td>
<td>Delaware university survey</td>
<td></td>
</tr>
<tr>
<td>C Proportion of specimens collected that are sent for confirmatory testing, %</td>
<td>26</td>
<td>Delaware university survey</td>
<td>20–30 (through May 12); 5–15 (after May 12)</td>
</tr>
<tr>
<td></td>
<td>(through May 3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D Test detects influenza</td>
<td>Published studies</td>
<td></td>
<td>90–100</td>
</tr>
<tr>
<td>E Proportion of confirmed cases reported to CDC</td>
<td>Assumption</td>
<td></td>
<td>95–100</td>
</tr>
<tr>
<td>No. reported cases</td>
<td>43,677</td>
<td>Reports to CDC through July 23, 2009</td>
<td>4,759 (through May 12); 33,909 (after May 12)</td>
</tr>
</tbody>
</table>

*BRFSS, Behavioral Risk Factor Surveillance Survey; ILI, Influenza-like illness; CDC, Centers for Disease Control and Prevention. States include California, Colorado, Connecticut, Georgia, Maryland, Minnesota, New Mexico, New York, Oregon, and Tennessee.
†Parameter estimates and sources are described in further detail in the online Technical Appendix; available from www.cdc.gov/EID/content/15/12/2004-Techapp.pdf.
Conclusions

We demonstrate that the reported cases of laboratory confirmed pandemic (H1N1) 2009 are likely a substantial underestimation of the total number of actual illnesses that occurred in the community during the spring of 2009. We estimate that through July 23, 2009, from 1.8 million to 5.7 million symptomatic cases of pandemic (H1N1) 2009 occurred in the United States, resulting in 9,000–21,000 hospitalizations. We did not estimate the number of deaths directly from our model, but among reports of laboratory-confirmed cases though July 23, the ratio of deaths to hospitalizations was 6%. When applying this fraction to the number of hospitalizations calculated from the model—that is, by assuming that deaths and hospitalizations are underreported to the same extent—we obtain a median estimate of 800 deaths (90% range 550–1,300) during this same period. Because this assumption has several limitations (5), more sophisticated models are also being developed to better understand the severity of the US epidemic in the spring of 2009, including intensive care unit admissions and deaths (6).

Our analysis involves several assumptions. Data for parameter estimates were collected in limited periods and areas and thus may not be fully representative of the entire United States. To account for some of this uncertainty, a range of values was included for each proportion. Additional data from surveys of health-seeking behavior, physician testing practices, and policies for confirmatory testing at public health laboratories could help refine the parameter estimates. In addition, parameters were obtained from studies of persons with ILL, defined as fever with cough or sore throat. Persons with milder illness may be less likely to seek care or be tested, and thus may not be fully captured in these estimates. Likewise, in some heavily affected areas, the size of the outbreak quickly exceeded the capacity to ascertain and test case-patients. Thus, our results may reflect a conservative estimate of total cases.

As pandemic (H1N1) 2009 continues to spread through the United States and the world, laboratory-confirmed cases will continue to greatly underestimate the number of actual cases that occur. Surveillance for influenza does not traditionally rely on complete case ascertainment, which would be impractical, but on focused case ascertainment with well-characterized surveillance systems and special studies. Unfortunately, relying on laboratory-confirmed cases limits the ability to understand the full impact and severity of the epidemic, especially when severe cases are more likely to be recognized (5).

This model provides a relatively quick and simple approach to estimate the human health impact of the epidemic in advance of more rigorous analysis of surveillance and health care data that will be available over the next few years. Health systems and infrastructure may be unprepared in the short-term if plans are based on a number of confirmed cases that substantially underestimates the impact of the epidemic. We estimate that the total number of pandemic (H1N1) 2009 cases in the United States during April–July 2009 may have been up to 140× greater than the reported number of laboratory confirmed cases. A spreadsheet version of the model has been developed and is available online (www.cdc.gov/h1n1flu/tools). Using this tool, health officials and policy makers could adjust...
the model parameters to represent their local experience, which may provide useful estimates of the prevalence of pandemic (H1N1) 2009 in their areas and help plan for a subsequent wave of the epidemic in the fall and winter months of 2009–2010.

M.L. acknowledges support from the US National Institutes of Health Models of Infectious Disease Agent Study program through cooperative agreements 5U01GM076497 and 1U54GM088588. M.L. has received consulting fees from the Avian/Pandemic Flu Registry (Outcome Sciences), funded in part by Roche.

Dr Reed is an epidemiologist in the Influenza Division at the Centers for Disease Control and Prevention. She recently completed a 2-year fellowship as an Epidemic Intelligence Service Officer in the Influenza Division at CDC.

References


Technical Appendix

Sources for Parameter Estimates

Parameters were estimated from a number of different surveillance systems and special investigations. These sources are briefly described below.

2007 Behavioral Risk Factor Surveillance Survey (BRFSS)

The BRFSS, a random-digit-dialed telephone survey was established by the Centers for Disease Control and Prevention (CDC) and state health departments in 1984 to obtain a representative sample of adults ≥18 years of age in each state and the District of Columbia. The BRFSS collects information annually on a core set of health behaviors, and also may include a variety of additional public health modules. In 2007, a module on influenza-like illness (ILI) was included in 9 states, designed to assess the incidence of ILI, health-seeking behavior, physician diagnosis of influenza, and treatment of influenza with antiviral medications. The 9 participating states included California, Colorado, Connecticut, Georgia, Maryland, Minnesota, New Mexico, Oregon and Tennessee. A total of 51,249 persons ≥18 years of age participated in these states during the 12 month survey. Analysis of BRFSS data involves weighting for the probability of selection of a telephone number, the number of adults in a household, and the number of telephones in a household, as well as to reflect the age and gender distribution of the underlying population.

2009 ILI Community Survey

In May 2009, after the identification of pandemic (H1N1) 2009 in the United States, a random-digit dialed telephone survey sampled similarly to the BRFSS was conducted using only the ILI module from the 2007 BRFSS and some limited demographic information. Respondents were adults ≥18 years of age living in the same 9 states where the ILI module was included.
during the 2007 BRFSS plus New York State. Participants were asked the same set of questions included in the ILI module during the 2007 BRFSS, including ILI in the past month, care-seeking behavior, receipt of antiviral treatment, and influenza vaccination. Participants were also asked the same questions about all members of their household. A total of 1,788 adults responded during a 3-week period. As with analysis of the BRFSS, data were weighted for the probability of selection and the age and gender distribution of the population.

Chicago Community Survey

Illness with pandemic (H1N1) 2009 virus was first documented in Chicago in late April 2009. By early May, a community in northeast Chicago accounted for approximately one third of the city’s confirmed cases. To further investigate this early outbreak of pandemic (H1N1) 2009 in Chicago and better characterize community transmission, a community survey was conducted in this neighborhood. Using a multistage cluster design, investigators obtained a representative sample of households in six census areas and then administered a standardized questionnaire to enrolled households. A total of 240 households and 643 persons were enrolled, including adults and children.

Delaware University Survey

A large outbreak of pandemic (H1N1) 2009 occurred on a university campus in Delaware in late April 2009. Data were available on clinic visits and influenza testing among students from the campus health center. To further describe the extent of the outbreak and effect of illness on campus, an online survey was conducted to assess health-seeking behaviors, influenza vaccination status, risk factors for illness, prevention practices, and measures to reduce transmission. A total of 6,049 students (32% response rate) and 1,401 faculty/staff (24% response rate) responded to the online survey over a 1-week period.

Values for Parameter Estimates

A) We estimate that between 42% and 58% of persons with influenza-like illness sought medical care for their illness. The low of 42% was calculated through a weighted analysis of the 2007 BRFSS, during an annual influenza season before the emergence of pandemic (H1N1) 2009. A similar survey conducted in the same states in May 2009 during early pandemic (H1N1) 2009 outbreaks found that a slightly higher proportion (52%) of respondents with ILI reported...
seeking medical care for their illness. Participants in this survey were also asked about other members of their household, and 55% of household members with ILI in May 2009 were reported to have sought medical care. Similarly, a university-wide survey in Delaware indicated that 58% of students with ILI and 49% of faculty and staff with ILI reported that they sought medical care for their illness. Finally, 52% of persons with ILI included in a community household survey in a Chicago neighborhood reported seeking medical attention.

These values are all calculated from self-reported medical visits following ILI, and thus may overestimate care-seeking behavior. Some recent studies on the accuracy of self-reported medical visits find that self-reported care seeking in population surveys tends to overstate the actual number of medical visits seen in the same community (1,2). If the true value for this parameter is lower, it would result in a higher multiplier and thus a higher estimate of cases than we present.

B) Of patients who reported seeking outpatient care, we estimate that between 22% and 34% had a swab taken by the physician for an influenza test. Weighted analysis of the BRFSS data from 2007 indicated that 25% of respondents reported having been swabbed for an influenza test. A similar analysis of the data from 2009 found that 22% of respondents and 28% of their household members reportedly were swabbed for an influenza test. During the Delaware investigation, 34% of students and 30% of faculty who sought medical care reported that they were swabbed for an influenza test. All of these values come from self-reports of influenza testing. Data from the campus health center were also available from the Delaware investigation, and these showed that 19% of patients presenting with febrile respiratory illness had a specimen collected for a rapid influenza test. Anecdotally, physicians continued collecting clinical specimens from patients with ILI for rapid antigen tests, even though not all were sent for further subtyping and confirmation with reverse transcription–PCR (RT-PCR).

During 2 recent influenza seasons, capture-recapture analyses were performed in areas that had two independent surveillance systems operating concurrently for influenza-associated hospitalizations in children <5 years of age. These analyses found that laboratory-based surveillance from routine clinical testing identified 38%–39% of all pediatric influenza-related hospitalizations, and prospective surveillance and testing of all children with acute febrile illness still only identified 69%–74% of children with influenza (3,4). Although no similar data are
available on surveillance for influenza-related hospitalizations in older children and adults, we would hypothesize that detection may even be lower for older age groups. Given the variety of approaches to surveillance between states and hospitals, we include a fairly broad range of 40%–75%.

C) In the United States, health departments were encouraged to seek laboratory confirmation of suspect cases with RT-PCR at federal or state public health laboratories early in the epidemic; however, as the epidemic progressed and transmission became widespread, complete case ascertainment became prohibitive, and by May 12, 2009, physicians were recommended to primarily test patients in special risk groups or those with severe illness. Data from the campus health clinic during the Delaware outbreak (before May 12) indicate that ≈26% of specimens collected by physicians may have been sent to the state for confirmatory testing; we included a range of 20%–30% in the model for this period. The proportion of samples sent for confirmatory testing likely decreased after activity became more widespread and testing recommendations were changed to focus on severe cases. Through May 12, hospitalized cases made up 4% of all the cases reported, but accounted for 12% of the reported cases after May 12. Thus, confirmatory testing of non-severe cases may have decreased three-fold following the change in recommendations. We included a $2\times$–$4\times$ lower range in our model for cases reported after May 12.

Likewise in hospitalized patients, although many physicians continued to collect clinical specimens for rapid point-of-care testing, not all may have been forwarded to public health laboratories for confirmation of pandemic (H1N1) 2009. We do not have any direct estimates of the proportion of specimens sent for confirmation during this period. We do know that in some locations with large outbreaks, specimens from hospitalized patients were only sent for confirmation if the patient had a positive rapid influenza test result. Given the sensitivity of rapid diagnostic tests for pandemic (H1N1) 2009 in early studies (5,6), this finding may represent as few as 50% of true influenza cases. Due to the uncertainty in this parameter estimate, we include a broad range of 50%–90%.

D) Although RT-PCR is known to have a high sensitivity, the ability of the test to detect influenza in a clinical specimen may be influenced by the quality of the specimen, specimen handling, timing of collection, or age of the patient (7,8). We included a range of 90%–100%
detection, which includes a range of sensitivities for RT-PCR seen in published studies. One previous study based on Canadian influenza surveillance, however, estimated that the sensitivity of testing given the quality of specimens collected and submitted in practice may be lower (9). If the true proportion is lower, this would lead to a higher multiplier and thus a higher estimate of total cases than we present.

E) Finally, due to miscommunication or delays in reporting, some patients with a positive test result may not be reported by the state health department to CDC and thus included in official case counts. Although we do not have any direct estimates of this fraction, we included a range of 95%–100% reporting.

Data on ascertainment of fatal cases are even more limited. Consequently, we chose not to estimate fatal cases directly in the model, but include a rough extrapolation based on the ratio of reported deaths to reported hospitalizations. More sophisticated models are being developed to estimate the severity of pandemic (H1N1) 2009 during the spring, including admission to an intensive care unit and death (10).

References


<table>
<thead>
<tr>
<th>Age group, y</th>
<th>Cases, † %</th>
<th>Hospitalizations, † %</th>
<th>Deaths, † %</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–4</td>
<td>13.0</td>
<td>20.1</td>
<td>2.5</td>
</tr>
<tr>
<td>5–24</td>
<td>59.6</td>
<td>36.3</td>
<td>17.4</td>
</tr>
<tr>
<td>25–49</td>
<td>20.1</td>
<td>25.0</td>
<td>44.9</td>
</tr>
<tr>
<td>50–64</td>
<td>5.9</td>
<td>13.9</td>
<td>25.7</td>
</tr>
<tr>
<td>&gt;65</td>
<td>1.4</td>
<td>4.7</td>
<td>9.4</td>
</tr>
</tbody>
</table>

*Age recorded for 85% of cases, 95% of hospitalizations, and 91% of deaths reported through July 23, 2009.
Figure. Distribution of the estimated number of total cases of pandemic (H1N1) 2009 (in millions) calculated from Monte Carlo simulation (10,000 iterations showing median estimate and 90% range).