Transmissibility and mortality impact of epidemic and pandemic influenza, with emphasis on the unusually deadly 1951 epidemic

Cécile Viboud, Theresa Tam, Douglas Fleming, Andreas Handel, Mark A. Miller, Lone Simonsen

Abstract

There are important gaps in our current understanding of the influenza virus behavior. In particular, it remains unclear why some interpandemic seasons are associated with unusually high mortality impact, sometimes comparable to that of pandemics. Here we compare the epidemiological patterns of the unusually deadly 1951 influenza epidemic (A/H1N1) in England and Wales and Canada with those of surrounding epidemic and pandemic seasons, in terms of overall mortality impact and transmissibility. Based on the statistical and mathematical analysis of vital statistics and morbidity epidemic curves in these two countries, we show that the 1951 epidemic was associated with both higher mortality impact and higher transmissibility than the 1957 and 1968 pandemics. Surprisingly in Liverpool, considered the ‘epicenter’ of the severe 1951 epidemic, the mortality impact and transmissibility even surpassed the 1918 pandemic.

Keywords: Influenza; Epidemic; Pandemic; Mortality; Transmissibility; Reproduction number; Canada; England and Wales

1. Introduction

Influenza is responsible for substantial morbidity and mortality worldwide, despite continued use of vaccine in recent years [1]. Infrequently, a novel influenza virus emerges, against which the human population has little or no pre-existing immunity, thereby causing worldwide pandemics. The 1918 pandemic (A/H1N1) has reportedly caused 20–50 millions deaths worldwide [2], while the 1957 and 1968 pandemic (respectively, A/H2N2 and A/H3N2) were milder with an estimated 1 million deaths worldwide or less [3]. In addition to these rare events, epidemics of influenza recur each year in wintertime due to gradual changes in the virus antigens to evade host immunity [4,5]. The mortality impact of these annual epidemics is highly variable from year to year: in the US some epidemics are associated with no increase in mortality while others are responsible for as many as 74,000 deaths, mainly among the elderly [6]. It is recognized that the influenza impact depends on the predominant subtype of circulating viruses, with A/H3N2 viruses being generally more virulent than B or current A/H1N1 viruses [7]. Yet it is intriguing that within a given subtype, there are large residual fluctuations in the mortality and morbidity impact of epidemics. For instance, in the last two decades in the US, there was a four-fold difference in mortality impact between mild and severe A/H3N2 seasons [6]. The virus

Abbreviations: P&I, pneumonia and influenza; SEIR, Susceptible-Exposed-Infected-Recovered; R, effective reproduction number; CFR, case fatality rate

* Corresponding author at: Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, 16 Center Drive, Bethesda, MD 20892, USA. Tel.: +1 301 496 2146; fax: +1 301 496 8496.

E-mail address: viboudc@mail.nih.gov (C. Viboud).
and population-specific reasons for these variations remain unclear.

Here, we investigate whether the particularly high mortality impact associated with some influenza seasons is related to higher transmissibility of the culprit viruses. To illustrate this hypothesis, we focus on the unusually deadly 1951 influenza epidemic, which occurred in the midst of the early era of A/H1N1 virus circulation (1918–1957) [8–10]. In a companion study, we have shown that England and Wales and Canada were particularly severely affected by the 1951 epidemic [8]. Here we estimate the mortality impact and transmissibility (through the effective reproduction number) of this epidemic from mortality and morbidity epidemic curves recorded in England and Wales and Canada. We compare these estimates with those of other severe influenza seasons, including the three pandemics of the last century.

2. Methods

2.1. Morbidity and mortality data

To estimate the mortality impact associated with the 1951 epidemic and pandemics in England and Canada, we compiled monthly death rates from pneumonia and influenza (P&I) and all-cause from national vital statistics, as described previously [8].

To estimate the transmissibility of the 1951 epidemic and other influenza seasons, we compiled three datasets from independent sources on weekly mortality and morbidity from influenza. First, we used mortality rates specifically attributed to influenza for the large cities of England and Wales (the so-called weekly deaths in the Great Towns), representing about 50% of this country’s population. We obtained this data for severe influenza seasons from 1918 to 1951 from Refs. [11,12], including the three waves of the 1918 pandemics and the 1951 epidemic. Data for the 1957 and 1968 pandemics were compiled from Refs. [13,14]. The corresponding epidemic curves are shown in Fig. 1, from the first week of increase in mortality until the week after the peak. Second, we compiled excess mortality rates from respiratory diseases and all-cause for Liverpool, England, where the 1951 epidemic was said to originate and had particularly high mortality impact [8–10]. Since morbidity data may be considered a better proxy for the transmission process than mortality, we also analyzed a third dataset recording the temporal progression of clinical cases of influenza during the 1951 epidemic in Canada and Canadian provinces, from tables published in the weekly Public Health Reports [15].

Regarding the impact and transmissibility of the 1968 A/H3N2 pandemic in England and Wales, we considered the second wave of A/H3N2 virus circulation (1969–1970). In this country, the second wave was associated with a much higher morbidity and mortality impact than the first wave (but not in Canada [16]).

2.2. Estimation of influenza mortality impact in England and Wales and Canada

We applied a seasonal regression model to monthly national vital statistics in order to estimate baseline mortality in the absence of influenza, separately for each outcome (P&I and all-cause) and country. Influenza mortality impact was calculated as the mortality in excess of the baseline during winter months [6,16]. Here we only present overall mortality patterns; detailed age-specific estimates can be found in Ref. [8].

2.3. Estimation of influenza transmissibility (effective reproduction number)

We adapted an algorithm initially developed for estimating the transmissibility of the 1918 pandemic virus from weekly mortality data in US cities [17]. In this approach, transmissi-
bility is estimated through the effective reproduction number \(R\) at the beginning of the epidemic, which represents the number of secondary cases per primary case. Higher values of \(R\) indicate higher transmissibility. Briefly, the algorithm generates epidemic curves of influenza cases for a given value of \(R\), through an SEIR model (Susceptible-Exposed-Infected-Recovered), a dynamic model of influenza transmission. Then, a fraction of cases is transformed to deaths, by applying a given case fatality rate (CFR). The algorithm uses a priori parameters for the latent period of influenza, infectious period, recovery rate, and time from disease onset to death, based on published means and distributions [17].

The SEIR-simulated epidemic curves are then fitted to the observed ones (mortality or morbidity weekly rates) through an optimization procedure, designed to find the values of the two unknown parameters \((R, \text{CFR})\) that jointly minimize the error between simulated and observed data. Here, we fit the 3 weeks associated with the maximum increase in incidence, usually the first three epidemic weeks. We conducted several sensitivity analyses by (i) estimating a single unknown parameter, \(R\), and fixing the CFR at one of two extreme values (2%, as reported for the 1918 pandemic [17], or 0.1% as reported for the 1957 pandemic [13]); (ii) recalculating \(R\) by the more straightforward method of dividing the maximum log increase in cases (or deaths) over two consecutive weeks by a known serial interval between two cases. The serial interval is the time interval between successive cases in a chain of transmission, estimated at 4.1 days for influenza [17].

3. Results

3.1. Mortality impact

In term of overall mortality impact, the 1951 epidemic was unusually severe in England and Wales and Canada, even in comparison with pandemic seasons. In particular, we found that in these countries, the 1951 epidemic had \(\sim 1.5\)-fold higher mortality impact than the 1957 pandemic and \(\sim 2\)-fold higher impact than the 1968 pandemics (with some variations depending on the mortality outcome considered, see Table 1). Surprisingly, the 1951 epidemic had even higher impact in Liverpool, where excess mortality rates from respiratory diseases and all-cause were higher than during the most severe wave of the 1918 pandemic itself (Fig. 2).

3.2. Transmissibility (effective reproduction number)

Overall, our estimates of the transmissibility parameter, \(R\), ranged from 1.9 to 2.5 for the 1951 epidemic, depending on the dataset and outcome considered; this represented an unusually high value for inter-pandemic influenza, as we detail in the section below.

We first compared the \(R\) estimates for the 1951 epidemic with earlier influenza seasons in the early A/H1N1 era using mortality data from the largest cities of England and found that transmissibility was unusually high for the 1951 epidemic (Fig. 3A). As expected, \(R\) was generally higher for pandemic than epidemic seasons. However, the 1951 epidemic appeared as an outlier, with higher \(R\) than for the 1957 and 1968 pandemics \((R = 2.0\) versus 1.6 and 1.8, respectively); in fact \(R\) for the 1951 epidemic was on the order of that found for the most transmissible wave of the 1918 pandemic (2.1). Setting the CFR at 2% or 0.1% instead of considering it as an unknown parameter had very little effect on \(R\) estimates (the difference in \(R\) estimates across this large range of CFR was only 4% on average, max 9%).

In Liverpool, the 1951 epidemic ‘epicenter’, transmissibility of this epidemic was even higher than for the rest of England \((R = 2.2\) based on influenza-specific death rates, Fig. 3B). More surprising in this city, transmissibility of the 1951 epidemic appeared higher than for all three waves of the 1918 pandemic, and the 1957 pandemic as well. In general, estimates of \(R\) based on influenza-specific deaths were higher than those based on all-cause excess deaths (e.g., for the 1951 epidemic, \(R = 2.2\) versus 1.9, respectively). This is probably because it is difficult to retrieve the exact contribution of influenza to all-cause mortality, when the mortality

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Comparison of the mortality patterns of the unusually deadly 1951 epidemic with the 1957 and 1968 pandemics in England and Wales and Canada</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>England and Wales (43.8 millions(^a)); relative risk of death in pandemic season vs. 1951 epidemic season(^b) (seasonal excess deaths per 100,000)</td>
</tr>
<tr>
<td>P&amp;I All-cause</td>
<td></td>
</tr>
<tr>
<td>1951 epidemic (A/H1N1)(^e)</td>
<td>1.00 (50.1)</td>
</tr>
<tr>
<td>1957 pandemic (A/H2N2)(^d)</td>
<td>0.71 (35.8)</td>
</tr>
<tr>
<td>1968 pandemic (A/H3N2)(^e)</td>
<td>0.78 (39.3)</td>
</tr>
</tbody>
</table>

\(\text{P&I: Pneumonia and influenza. For comparison purposes, the second wave of the 1968 pandemic is considered for England and Wales (1969–1970 season), because it had much higher impact than the first [16].}\)

\(^a\) Population in 1951.

\(^b\) Relative risk calculated as (pandemic mortality impact)/(mortality impact of the 1951 epidemic).

\(^c\) Periods of increased mortality: 1951 epidemic: England: January 1951–March 1951; Canada: January 1951–April 1951.


Fig. 2. Comparison of the 1951 influenza epidemic with the 1918 and 1957 pandemics in Liverpool, England. Time series of weekly excess death rates from (A) respiratory deaths (pneumonia, influenza and bronchitis) and (B) all-cause. Epidemics were aligned at the week of peak mortality (peak week = week ended February 22, 1919; January 13, 1951; October 12, 1957). The 1918 pandemic occurred in three waves in Liverpool (summer 1918, autumn 1918, winter 1919); the “third wave” was associated with the highest mortality impact and is presented here.

Fig. 3. Estimates of the effective reproduction number ($R$), based on influenza weekly death rates and using the algorithm detailed in Ref. [17]. (A) Estimates based on mortality in the large cities of England and Wales. (B) Estimates based on mortality in Liverpool. Red bars: pandemic seasons; black bars: inter-pandemic seasons; blue bar: 1951 epidemic. $R$ is estimated separately for the three waves of the 1918 pandemic (summer wave = 1918–19 (1), autumn wave = 1918–19 (2), winter wave = 1918–19 (3)).

increase due to influenza is superimposed on a large background of unrelated deaths (‘noise’). Hence considering a less specific outcome such as all-cause excess mortality results in underestimation of $R$. Nevertheless for both mortality outcomes in Liverpool, we found the same intriguing pattern of higher transmissibility in the 1951 epidemic than in the 1918 and 1957 pandemics.

Canadian morbidity data gave estimates of $R$ for the 1951 epidemic consistent with those derived from mortality data, both at the national and regional scale (estimate for the whole of Canada, $R = 2.1$; range for Canadian provinces $R = 1.9–2.5$). Unfortunately, historical morbidity records were not available in Canada for early A/H1N1 epidemics to compare with these estimates. As a crude comparison however, the preceding two influenza seasons in 1948–1949 and 1949–1950 produced, respectively, mild and moderate morbidity [15] and were associated with substantially lower $R$ values ($R = 1.3$ and 1.5, respectively).

As a sensitivity analysis, we recalculated all estimates of $R$ by considering the maximum rate of increase in deaths (cases) over two consecutive weeks (Fig. 1) and a predetermined serial interval of 4.1 days. This sensitivity analysis gave estimates consistent with those derived from an SEIR model, for all seasons (Pearson correlation coefficient between $R$ values produced by the two methods = 0.96, $P < 0.01$; mean difference in $R$ values = 5%). Importantly, the general pattern of $R$ values was preserved, in that transmissibility of the 1951 epidemic in England and Wales and Canada was found similar to that of pandemic seasons rather than inter-pandemic ones.

4. Discussion

We have shown that the 1951 influenza epidemic had unusually high mortality impact in England and Wales and Canada, especially given that it occurred well into the A/H1N1 inter-pandemic period, when the A/H1N1 subtype was already well established. In particular the 1951 epidemic had higher mortality impact than the 1957 and 1968 pandemics, and higher mortality impact than all three pandemics in Liverpool. The analysis of various mortality and morbidity epidemic curves revealed another unusual feature of the 1951
epidemic: its transmissibility was similar to that of pandemic seasons rather than inter-pandemic ones.

Our estimates of the effective reproduction number for the three influenza pandemics of last century in England and Wales are remarkably consistent with previous studies [17–19], even though these studies used different methodological approaches or data from different countries (Table 2). The largest differences are perhaps found for the 1968 pandemic. We have previously reported on the surprising ‘smoldering’ pattern of delayed mortality impact for the 1968 pandemic in Europe and Asia [16]. It is possible that this pattern was associated with true geographical and temporal differences in influenza transmissibility.

By contrast to pandemic influenza, estimates of transmissibility for inter-pandemic influenza are rare, let alone for the era before the 1957 pandemic when A/H1N1 viruses circulated. Yet our estimates for severe seasons in the old A/H1N1 inter-pandemic period ranged between 1.2 and 1.8 (excluding the 1951 epidemic outlier), which seems reasonably in line with an estimated $R = 1.5$ for the 1984–1985 A/H3N2 inter-pandemic season in France [20].

There are some potential caveats to our study. First, we have used epidemic data aggregated at the scale of a country (England and Wales, Canada), which may lead to underestimation of $R$ because of spatial heterogeneities. However our conclusion of unusually high transmissibility for the 1951 epidemic relies on the comparison of the relative values of $R$ across several epidemic and pandemic seasons. Moreover, the robustness of our results is guaranteed by parallel analyses of mortality and morbidity data at the regional and city level (Canadian provinces, Liverpool). Second, to estimate the effective reproduction number for influenza, it is important to use mortality time series truly reflecting the influenza virus activity. This is traditionally done by estimating influenza-related mortality as the P&I or all-cause mortality in excess of a baseline [6,16]. This approach fails for mild influenza seasons, when the contribution of influenza to mortality becomes so small that it cannot be discriminated from background mortality (noise). This is why we have focused on the most severe influenza seasons in this study. Further, we have used deaths coded specifically as influenza, which is believed to be a very specific indicator of influenza activity [21] (in particular very few influenza-specific deaths occur outside of the influenza period). A third caveat relates to the comparison of the effective reproduction number across several influenza seasons to study temporal changes in transmissibility. The reproduction number estimates come with uncertainty, since they are derived from noisy data. Here, we have carefully checked that our conclusions hold for independent datasets and independent estimation algorithms. In addition, previous works have used a similar approach to compare influenza transmissibility across seasons or geographical units, and draw conclusions about the natural history of the disease [17,22].

The laboratory data collected by the WHO surveillance network during the 1951 epidemic provides clues as to what the unusually severe 1951 influenza strain was. Influenza viruses circulating in the winter of 1951 in England and Wales and Canada were confirmed as A/H1N1, and hemagglutinin inhibition tests conducted at the time did not evidence unusual changes in the hemagglutinin antigen [23,24]. In line with these laboratory results, we have previously shown that the age pattern of mortality in 1951 was not reminiscent of pandemics, since there was no age shift of deaths towards younger age groups [8]. This epidemiological and virological evidence point to a virus that did not have entirely novel antigens (i.e., no shift).

The present study indicates that the unusually large mortality burden of the 1951 influenza epidemic in England and Wales and Canada was associated with unusually high transmissibility of the circulating virus. This finding corroborates the simple observation that this epidemic was particularly intense and short [9,10] (see also Fig. 2). Unfortunately, we were not able to compare estimates of the case fatality rate between the 1951 epidemic and other seasons since this parameter value varied substantially across datasets. Yet, investigation of age-specific morbidity and mortality data for Liverpool, the epidemic ‘epicenter’, suggests a higher case fatality for the 1951 epidemic than the 1957 pandemic among children and working adults [9,10], and in turn higher pathogenicity for the 1951 virus. Unfortunately, this observation is limited to these age groups and city. The elderly are the population subgroup at highest risk of death during influenza epidemics in general [21] – and during the 1951 epidemic in particular [8]. In the absence of morbidity data for the elderly, the pathogenicity of the 1951 virus remains unclear.

We see two biological mechanisms potentially responsible for increased transmissibility in some influenza viruses. First, the theory of the dynamics of infectious diseases predicts that transmissibility increases with the proportion of susceptible hosts in the population [25]. For influenza, the pool of susceptible hosts becomes large following antigenic drift of the influenza virus [5] and even larger following pandemic shifts [1,17]. Epidemiological and virological evidence for the 1951 epidemic rule out a pandemic shift in the circulating virus.

Table 2
Comparison of the effective reproduction numbers estimated by different studies for the three influenza pandemics of last century

<table>
<thead>
<tr>
<th>Year</th>
<th>Study</th>
<th>1918 pandemic</th>
<th>1957 pandemic</th>
<th>1968 pandemic</th>
</tr>
</thead>
<tbody>
<tr>
<td>This study</td>
<td>2.1</td>
<td>1.5</td>
<td>1.8</td>
<td></td>
</tr>
<tr>
<td>Gani et al.</td>
<td>2.0</td>
<td>1.7</td>
<td>2.2</td>
<td></td>
</tr>
<tr>
<td>Mills et al.</td>
<td>2.0</td>
<td>–</td>
<td>–</td>
<td>1.9</td>
</tr>
<tr>
<td>Longini et al.</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

When the pandemic came in several waves, the wave of highest impact was considered.

a Data for England and Wales.
b Data for England and Wales; estimated from the final epidemic size (including asymptomatic cases).
c Median estimate for 45 US cities, interquartile range (1.7–2.3); algorithm similar to ours.
d SEIR model fitted to on international data.
but antigenic drift was possible — although we think the drift would have to be unusually pronounced given the high estimated value of $R$. The alternative mechanism is that transmissibility can also increase with viral fitness, in particular with advantages on viral replication. A recent study indicates that the virus responsible for the 1918 pandemic had a replication advantage via particularly efficient polymerases, explaining in part its unprecedented pathogenicity [26]. Whether similar features existed in the 1951 virus remains elusive: unfortunately, no sequence is available for this virus in the public domain. If historical specimens could be recovered, it would be extremely interesting to revisit this question at the molecular level and examine changes not only in surface antigens (to study drift) but also internal genes (to study viral fitness). Both are equally important, as there appears to be a balance between drift and viral fitness. In particular, a study of recent A/H3N2 virus genomes revealed that some influenza viruses gained some advantage on population immunity through antigenic drift early on, but did not become widespread until they later acquired a backbone of fit internal genes [27].

The comparative similarity of total mortality in pandemics and major epidemics provide a justification for pandemic planning to be based on the available data for the 1957 and 1968–1969 pandemics. The reasons for the enhanced severity of the 1951 epidemic remain unclear and illustrate important gaps in our current understanding of the influenza virus. It is intriguing that the specific example of the 1951 epidemic is not unique and is echoed in the modern era of A/H3N2 virus circulation, albeit less dramatically. In the recent inter-pandemic period, the most severe seasons are not always associated with antigenic novelty (drift), but with antigenically similar strains returning for several consecutive seasons: striking examples are the 1989–1990 epidemic in the UK/A/England/89(H3N2)strain [28] or the 1999–2000 epidemic in the US/A/Sydney/97(H3N2)strain) [29,30]. These observations remain unexplained.

We hope that with a rapidly expanding public database of sequences of complete genomes of influenza viruses [31], we can begin to elucidate the intricate relations between influenza antigenic changes, viral transmissibility and pathogenicity, pre-existing population immunity, and the resulting influenza mortality and morbidity impact.

Acknowledgments

We thank Ottar Bjørnstad and Benjamin Bolker for providing comments early in this work, in particular helping with the SEIR model and optimization procedure. We thank the organizers of the 3rd conference on the Ecology and Evolution of Infectious Diseases (Fort Collins, Colorado, May 2005) and the Research Program on Mathematical Biology at the Park City Mathematics Institute (Park City, UT, July 2005), during which CV and AH initiated this modeling effort. This study was funded by the Fogarty Internal center, National Institutes of Health.

References


