ESTIMATES OF THE BASIC REPRODUCTIVE NUMBER FOR 1918 PANDEMIC INFLUENZA IN THE UNITED STATES:

IMPLICATIONS FOR POLICY

Marc Lipsitch[1,2], Christina Mills[1], James Robins[1,3*]

1. Department of Epidemiology, Harvard School of Public Health, Boston, MA USA
2. Department of Immunology and Infectious Diseases, Harvard School of Public Health, Boston, MA USA
3. Department of Biostatistics, Harvard School of Public Health, Boston, MA USA
* NOTE: DUE TO A FAMILY EMERGENCY DR ROBINS HAS NOT HAD A CHANCE TO APPROVE THE FINAL MS. WE WILL PROVIDE SUCH APPROVAL AS SOON AS POSSIBLE.

ABSTRACT

We have estimated that the basic reproduction number (R_0) – the mean number of secondary infections created by a single case in a fully susceptible population – for pandemic influenza in US cities in 1918 was in the approximate range 2-3 -- less than previous estimates, less than that of many other respiratory infections, and less than or equal to that of SARS in 2003. We discuss the main findings of this study here. Even with such a modest value of R_0, containment of pandemic influenza would be extremely challenging, since the methods successful in controlling SARS, which depended heavily on identification and isolation of cases, would be much less effective against pandemic influenza, in which transmission occurs more rapidly after infection and before or just at the onset of symptoms. Moreover, biological and probabilistic considerations suggest that, if an influenza pandemic is likely in a given period of time, multiple introductions of strains capable of pandemic spread must also be anticipated. Hence, preparation for a pandemic must proceed on the assumption that multiple introductions are possible. Because vaccines are the only major transmission-blocking measure that is effective against repeated introductions, we argue that the capacity to produce vaccines rapidly and in sufficient quantities to immunize a large majority of the population is the most urgent priority in pandemic preparedness. The modest values of R_0 estimated for pandemic influenza, furthermore, make it likely that such an effort could make it possible to block transmission in a population immunized with achievable levels of coverage.

INTRODUCTION

The 1918 influenza H1N1 pandemic is often considered a worst-case scenario for planning for future pandemics. Estimates of deaths from that pandemic range upwards from 550,000 in the United States and from 20 to 100 million worldwide (20). The rapid geographic spread of the infection and the explosive growth of the epidemic within localities, combined with an extraordinarily high risk of death among infected persons (varying from 0.8% to 9% depending on the demographic group, about 1-5% in healthy young adults (9), about an order of magnitude higher than that seen even in other pandemics) to produce the most deadly pandemic of influenza in documented history.

The basic reproductive number, R_0, is a single number that summarizes the infectiousness of a pathogen in a given population. Defined as the average number of secondary cases infected by a typical infected person in a population that is almost fully susceptible, R_0 contains several key pieces of information relevant to the spread and control of an infection in a population. If an infection has a value of R_0 less than 1, it cannot have sustained transmission in the population and will eventually go extinct, though there may be isolated episodes and even brief chains of transmission. If R_0 exceeds 1, the infection can spread in the population, because each case is expected to cause more than one additional case, leading to an initially exponential growth in the number of cases. If R_0 is greater than 1 but still close to 1, (roughly speaking, less than 2 or 3, depending on some details of transmission), individual introductions of the infection into a population may well die out by chance alone, but the likelihood of such chance fadeouts declines exponentially with the number of introductions (12). From the perspective of control, R_0 determines the proportion of all transmission events that must be blocked – via vaccination, treatment, isolation, reductions in contacts, or in other ways – to halt transmission of the infection. The higher the value of R_0, the greater the effort required for containment, since the average number of people infected by each case must be reduced from its initial value (i.e., R_0) down to below one to achieve control of the epidemic. Specifically, the proportion of such transmissions that must be blocked is given by the relation $p_c = 1-1/R_0$: half the transmissions if R_0=2, 80% if R_0=5, and 95% if R_0=20.

The control of SARS in 2003 is often seen as a model for the control of an outbreak of pandemic influenza. For reasons described below – primarily the fact that identification and isolation of influenza cases is much more difficult than that of SARS cases, and will even in the best case often miss the peak of transmissibility (8) -- many of the measures that ultimately succeeded in controlling SARS will be poorly suited to control of pandemic influenza. Nonetheless, a comparison of the estimated transmissibility of the two infections, and the effectiveness of control measures, can be instructive. Several groups estimated that the value of R_0 for the SARS epidemic in Hong Kong (12, 21, 24) and elsewhere (24) was approximately 3. Importantly, following the imposition of control measures in several countries, the infectiousness of SARS patients was reduced by about fourfold, so that for every 4 infected individuals, there were about 3 secondary cases once control measures were in place (24). While this extraordinary effort resulted

in control of the SARS epidemic in each of the centers worldwide where it had taken hold, it also demonstrated that our capacity for control of an infection is limited, and that in the case of SARS only perhaps 75% of transmissions could be blocked.

Despite intense interest in the epidemiology of influenza, estimates of the basic reproductive number for influenza have been made infrequently and have ranged more than tenfold, from less than 2 to as high as 20 (8, 10, 22, 23). Technical difficulties (mainly the difficulty in attributing illness or death to influenza based on clinical diagnosis alone), and difficulties in ascertaining the degree of susceptibility in the population in all but pandemic years partly explain why few such estimates had been made. We set out to provide such an estimate for the autumn, 1918 pandemic in cities in the United States. Because of the extraordinarily high numbers of deaths caused by pandemic influenza in 1918, we were able to use statistics on pneumonia and influenza mortality to perform this estimate, circumventing some of the difficulties typical for such estimations in nonpandemic years. Because the strain was novel to the population (with a possible caveat described below), it was possible to estimate R_0 directly from the rate at which such deaths increased over time in urban US populations.

METHODS AND KEY ASSUMPTIONS:

A detailed description of our scientific methods is given in (17). We analyzed data on pneumonia and influenza mortality from 45 United States cities in autumn, 1918. By fitting the early, growing phase of the epidemic to a mathematical model of the transmission of influenza, we estimated the value of the effective reproductive number, R, for each city during this early phase. A number of assumptions went into this analysis. The most important was the assumption that the serial interval – the time from when one individual is infected to the time when that individual infects another individual – follows a distribution with a mean of approximately 4.1 days and a range between 1 and 8 days (15). This assumption is consistent with previous models and with one interpretation of viral shedding data (11); however, a recent study (7) has plausibly suggested that the serial interval may be shorter than we have assumed, resulting in an even lower estimate of R. The other key assumption of this analysis was that models should be fit to the upward portion of the epidemic curve, but not to the peak and declining portion of the curve. The rationale for this assumption is that many factors, including interventions and behavioral changes, may have affected the peaking and decline of the epidemic, and we sought to use, to the extent possible, only the data that would pertain to spread unaffected by such interventions.

The effective reproductive number R is the actual number of secondary cases caused by each primary case in a given population. It may be correct to assume that the effective reproductive number at the beginning of an influenza pandemic is equal to the basic reproductive number, because pandemics by definition occur in highly susceptible populations. However, it is possible that some of the population was at least partially immune to influenza H1N1 by autumn, 1918. Inferences from age-specific mortality patterns suggest that individuals 31 or older may have been exposed to a virus circulating

prior to 1889 that offered partial protection against the 1918 pandemic strain (19). There was a severe but brief epidemic of influenza in the spring of 1918 in the United States, and epidemiologic patterns suggest that this epidemic was a "herald wave" of the larger, autumn pandemic and was itself caused by a related, H1N1 strain (4, 18). To account for this possibility, we considered that the effective reproductive number R that we estimated may have understated the basic reproductive number R_0, which would have pertained in a completely susceptible population, by as much as 30%.

Extensive sensitivity analyses indicated that other assumptions – for example, the choice of baseline rates of P&I death (due to nonpandemic causes), the assumption that P&I deaths reflected ongoing transmission via a time-to-death distribution and case-fatality proportion, and other such assumptions – could be varied substantially without making a major difference to our findings.

Because underestimation of R_0 could lead to a false sense of security with regard to the ability to control the spread of influenza, we performed a separate analysis that estimated R for each city from only the two consecutive weekly data points showing the largest proportional increase in P&I deaths over a weeklong period. The resulting "extreme R" estimate was intended to avoid possible downward bias in the estimate of R from incorporating periods of time including the imposition of control measures, but was expected to result in an overestimate of the true R due to stochastic factors (e.g. reporting errors). Hence these "extreme R" estimates were intended to be upper bounds.

We attempted to correlate estimates of R in each city with a number of demographic variables, including latitude, longitude, population size or density, and age or sex distribution.

KEY FINDINGS

Figure 1 shows the distribution of estimated reproductive numbers for 45 US cities. Our baseline estimates gave a median value of 2.0 for R, with an interquartile range of 1.7 to 2.3. Using the extreme R approach, we found a median of 2.7 with an interquartile range of 2.3 to 3.4 Inflating the upper estimates by up to 1/3 to account for the possibility that some immunity was present in the population at the start of autumn 1918, we conclude that the median R_0 for the 1918 influenza in US cities was between 2 and 4.

The variation in R estimates was considerably greater than that expected from sampling error (which should only contribute variations of about +- 0.1 to the estimates). This suggests that there was true heterogeneity in the R in different cities, and/or that sources of random variation other than sampling error, such as reporting errors or temporal changes in reporting rates, were present. While this heterogeneity may be due to differences in social conditions, differences in population composition, or differences in opportunities for transmission in the early phases of the epidemic, we did not identify any strong correlates of the variation in R estimates. Extreme R estimates were weakly correlated with population density in 1920.

IMPLICATIONS

Our estimates of the R_0 for pandemic influenza in 1918 suggest that to contain a similar pandemic, control measures would have to block at least 50% (if R_0=2) or 75% (if R_0=4) of transmission events. In contrast to some previous estimates of pandemic R_0, which ranged as high as 20, these estimates suggest that, with sufficient resources, it might be possible to contain the spread of a pandemic strain with transmission similar to that observed in 1918. Moreover, our estimates of R are broadly consistent with, though slightly higher than, estimates from other pandemics. If the estimate of R is a good approximation of R_0, then we may conclude that the exceptionally high level of mortality in the 1918 pandemic was due to an elevated case-fatality proportion, rather than due to extraordinary transmissibility.

While the relatively modest level of R_0 estimated for this pandemic suggests the possibility of controlling a similar pandemic, the details of how this might accomplished are critical. First, one should emphasize that R_0 cannot be assumed to be the same from one pandemic to the next. Biological differences in the influenza strain could certainly change R_0, and indeed R_0 may evolve (most likely upwards, evolutionary theory predicts (2)) during the course of a pandemic as the influenza strain becomes increasingly adapted to human-to-human transmission. Moreover, demographic and social changes since 1918 will certainly have some impact on the R_0 observed in particular settings. Control of pandemic influenza will require control in multiple settings simultaneously, so control measures will have to be successful where the R_0 for pandemic influenza is the highest, not only where it takes on its typical (e.g. median) value.

The rapid spread of pandemic influenza in 1918 and in other pandemics, we conclude, was primarily due to its brief serial interval or generation time (on the order of 4 days), rather than to a high R_0. Moreover, influenza infection is thought to be contagious before an individual is symptomatic, and especially before the onset of symptoms that would strongly indicate influenza infection. As a consequence, measures to identify infectious persons and block transmission from them will be especially difficult with influenza (8), much more than in the case of SARS, where transmission did not typically occur until more than a week after infection and several days after symptom onset (12). Indeed, for this reason the main measures used to control SARS – isolation of symptomatic cases and quarantine of their contacts – are not seriously contemplated to block transmission of pandemic influenza (8). The difficulty of blocking transmission by case identification helps to explain the findings of Longini and colleagues (15), who analyzed a mathematical model of the effects of providing antiviral prophylaxis to contacts of cases to contain pandemic influenza in a developed country. They found that extremely high coverage rates and long durations of prophylaxis (8 weeks) would be required, even assuming a value of R_0 (based on the 1957 pandemic) that was lower than that estimated in our study.

Since our work was completed, two groups have considered the possibility of implementing a strategy that depends on identification of infectious persons and blocking

ongoing transmission for control of the earliest phases of a pandemic as it arises, presumably in a rural part of East or Southeast Asia. These analyses conclude that antiviral prophylaxis in a geographic area around each case could contain an incipient pandemic if several conditions were met: most notably, if the strain had a basic reproductive number $R_0 < 1.8$ (7) or 1.6 (16), if it were detected early enough, if prophylaxis could be delivered with high efficiency to a geographic area around each case within 2 days, if the strain remained susceptible to oseltamivir, and if adequate antiviral supplies were available – here the estimates range considerably from 100,000 to 3 million courses, depending on the model and its assumptions. In summary, a combination of good luck (emergence of the pandemic with a modest R_0 in a place where it could be detected early), advance preparation (to have in place adequate surveillance to detect the emergent human-to-human transmission rapidly, adequate antiviral supplies, and the infrastructure to detect cases and prophylax around them with high efficiency), and excellent execution would be required for this strategy to work. Again, the fact that such a strategy is thinkable reflects the modest level of R_0 expected for pandemic flu, and the list of caveats reflects the basic difficulty of controlling influenza by measures that require contact identification.

Thus, the basic reproductive number and the timing of transmission, both in absolute terms and relative to symptom onset, are critical determinants of the feasibility of control measures that rely on case identification (8). A further key determinant of the success of any local control measure is the number of introductions of viruses capable of repeated human-to-human transmission that occur within a defined time period (say, a year or two). At one extreme, one could postulate that once a virus with this capacity emerges and begins spreading in humans, it is likely that other, similar viruses will be circulating in birds, perhaps moving significant geographic distances (3, 14), and that further emergence events are likely in quick succession, possibly in areas at some distance from one another. Under this assumption, efforts to block individual introductions will have little effect and will quickly be overwhelmed by the need to repeat the difficult job required for local containment. (Richard Danzig has made a closely analogous argument about the likelihood that bioterrorist attacks, if they occur, will be made repeatedly, a phenomenon he terms "reload"; he concludes that policies to respond to bioterrorism must therefore "prepare for a campaign, not just an attack" (5)).

On the other extreme, one could argue that the emergence of a pandemic-capable strain will most likely arise through genetic reassortment in a human infected simultaneously with a novel subtype and with an already human-adapted (H1N1 or H3N2) strain, and that this event is likely to occur very rarely. Any plausible biological scenario for the emergence of a new pandemic strain involves chance processes for which the probability of another emergence event does not go down (and may go up) following a given introduction. Hence, even under this scenario, the premise that the probability of a pandemic is high right now entails that the probability of a second introduction, after the first, is equally high. Put another way, it is inconsistent to believe that a pandemic is likely to occur in, say, the next few years, and that two introductions in that same period are very unlikely (13).

6

POLICY CONCLUSIONS: NEED FOR ADEQUATE, RAPIDLY AVAILABLE VACCINE SUPPLIES

Given these uncertainties, we believe that it is critical to make plans for containing influenza that are robust to the possibility of multiple introductions and to the possibility that even a single introduction could not be contained at its source. Hence, the need for effective vaccines, in sufficient numbers, early in a pandemic, becomes inescapable. A pandemic strain with a basic reproductive number between 2 and 3 would be blocked by effective immunization of ½ to 2/3 of a population; given that influenza vaccines are not 100% effective, this would mean vaccinating more than this fraction of the population. The United States currently has the domestic capacity to manufacture approximately 60 million doses of normal influenza vaccine, each containing a total of 45 micrograms of antigen (6). At the time of this writing (September 2005), US researchers have reported publicly, though not published, data indicating that four times this amount of antigen would be required (90 micrograms in each of two injections) to obtain a "maximal" response to a tested H5N1 vaccine (1). At this dose, US capacity would be enough to immunize approximately 15 million persons, and as much as six months would be required even to manufacture this amount of vaccine (6). Most other industrialized countries have less capacity relative to population size, and worldwide the total manufacturing capacity – about 300 million conventional doses (6), or 75 million persons' worth of the recently described H5N1 vaccine, falls even further short of need.

It follows that, without a major change in the rate and volume of vaccine manufacture, a large country such as the United States will be unable to counter the effects of the next pandemic within its borders.. The policy options for expanding our vaccine supply are the subject for another, in depth analysis, but these include the development and testing of adjuvanted or intradermally administered vaccines to reduce the quantity of antigen required per immunization (6), development of cell-culture-based systems for growing vaccine virus, and development of "prototype" vaccines for a variety of influenza A subtypes that could be rapidly altered, using reverse genetics technology (25), to create a seed strain for pandemic vaccine production. Another idea that has been considered is the possibility of incorporating H5N1 virus into normal, annual influenza vaccine programs, to provide some priming (albeit not tailored to a specific pandemic strain) that might lead to enhanced immune responses to vaccines depolyed in response to a pandemic.

Many policy questions remain about the relative merits, feasibility, and other aspects of these options. We believe that given the modest values estimated for R_0 for previous pandemic influenza strains (7, 17), vaccines deployed rapidly in a pandemic could be used to block a substantial proportion of transmission and significantly reduce morbidity and mortality. Moreover, because of the difficulty of control using methods that require case identification, and the possibility of multiple introduction, we believe that vaccines are the *only* measure that could – if available in time and in quantity – be counted on to make a dent in the severity of the pandemic. For this reason, we urge that every effort be made to accelerate research on vaccine production technologies, clinical trials of antigen-

sparing vaccines, and growth in vaccine manufacturing capacity to maximize our ability to counter the next pandemic.

REFERENCES

1. **Altman, L. K., and K. Bradsher.** 2005. Vaccine Alone Won't Stem Avian Flu, Experts Warn, The New York Times.
2. **Anderson, R. M., and R. M. May.** 1991. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford.
3. **Chen, H., G. J. Smith, S. Y. Zhang, K. Qin, J. Wang, K. S. Li, R. G. Webster, J. S. Peiris, and Y. Guan.** 2005. Avian flu: H5N1 virus outbreak in migratory waterfowl. Nature **436:**191-2.
4. **Crosby, A.** 2003. America's Forgotten Pandemic: The Influenza of 1918, 2nd Ed. Cambridge University Press, Cambridge.
5. **Danzig, R.** 2003. Catastrophic Bioterrorism: What Is to Be Done? Center for Technology and National Security Policy, National Defense University, Washington, DC.
6. **Fedson, D. S.** 2005. Preparing for pandemic vaccination: an international policy agenda for vaccine development. J Public Health Policy **26:**4-29.
7. **Ferguson, N. M., D. A. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, and D. S. Burke.** 2005. Strategies for containing an emerging influenza pandemic in Southeast Asia. Nature.
8. **Fraser, C., S. Riley, R. M. Anderson, and N. M. Ferguson.** 2004. Factors that make an infectious disease outbreak controllable. Proc Natl Acad Sci U S A **101:**6146-51.
9. **Frost, W. H.** 1920. Statistics of influenza morbidity with special reference to certain factors in case incidence and case fatality. Public Health Reports **35:**584-97.
10. **Gog, J. R., and B. T. Grenfell.** 2002. Dynamics and selection of many-strain pathogens. Proc Natl Acad Sci U S A **99:**17209-14.
11. **Hayden, F. G., R. Fritz, M. C. Lobo, W. Alvord, W. Strober, and S. E. Straus.** 1998. Local and systemic cytokine responses during experimental human influenza A virus infection. Relation to symptom formation and host defense. J Clin Invest **101:**643-9.
12. **Lipsitch, M., T. Cohen, B. Cooper, J. M. Robins, S. Ma, L. James, G. Gopalakrishna, S. K. Chew, C. C. Tan, M. H. Samore, D. Fisman, and M. Murray.** 2003. Transmission dynamics and control of severe acute respiratory syndrome. Science **300:**1966-70.
13. **Lipsitch, M., and C. E. Mills.** in preparation.
14. **Liu, J., H. Xiao, F. Lei, Q. Zhu, K. Qin, X. Zhang, D. Zhao, G. Wang, Y. Feng, J. Ma, W. Liu, J. Wang, and G. F. Gao.** 2005. Highly Pathogenic H5N1 Influenza Virus Infection in Migratory Birds. Science.
15. **Longini, I. M., Jr., M. E. Halloran, A. Nizam, and Y. Yang.** 2004. Containing pandemic influenza with antiviral agents. Am J Epidemiol **159:**623-33.

16.   **Longini, I. M., Jr., A. Nizam, S. Xu, K. Ungchusak, W. Hanshaoworakul, D. A. Cummings, and M. E. Halloran.** 2005. Containing pandemic influenza at the source. Science **309:**1083-7.
17.   **Mills, C. E., J. M. Robins, and M. Lipsitch.** 2004. Transmissibility of 1918 pandemic influenza. Nature **432:**904-6.
18.   **Olson, D. R., L. Simonsen, P. J. Edelson, and S. S. Morse.** 2005. Epidemiological evidence of an early wave of the 1918 influenza pandemic in New York City. Proc Natl Acad Sci U S A.
19.   **Palese, P.** 2004. Influenza: old and new threats. Nat Med **10:**S82-7.
20.   **Patterson, K. D., and G. F. Pyle.** 1991. The geography and mortality of the 1918 influenza pandemic. Bull Hist Med **65:**4-21.
21.   **Riley, S., C. Fraser, C. A. Donnelly, A. C. Ghani, L. J. Abu-Raddad, A. J. Hedley, G. M. Leung, L. M. Ho, T. H. Lam, T. Q. Thach, P. Chau, K. P. Chan, S. V. Lo, P. Y. Leung, T. Tsang, W. Ho, K. H. Lee, E. M. Lau, N. M. Ferguson, and R. M. Anderson.** 2003. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. Science **300:**1961-6.
22.   **Rvachev, L. A., and I. M. Longini.** 1985. A mathematica model for the global spread of influenza. Math Biosci **75:**1-22.
23.   **Spicer, C. C., and C. J. Lawrence.** 1984. Epidemic influenza in Greater London. J Hyg (Lond) **93:**105-12.
24.   **Wallinga, J., and P. Teunis.** 2004. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol **160:**509-16.
25.   **Wood, J. M., and J. S. Robertson.** 2004. From lethal virus to life-saving vaccine: developing inactivated vaccines for pandemic influenza. Nat Rev Microbiol **2:**842-7.
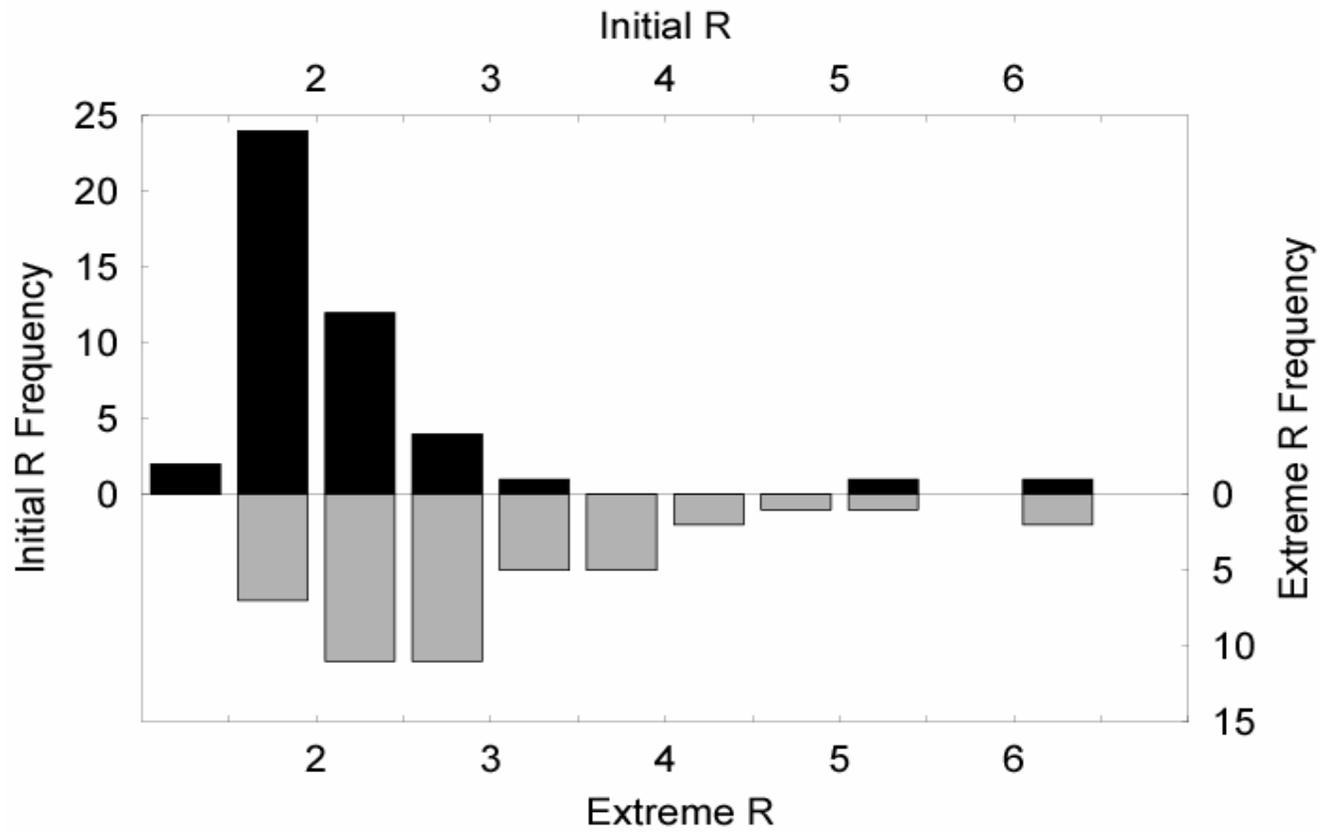
FIGURE 1: Histogram of initial and extreme estimated R values for 45 cities during the 1918 influenza pandemic. Dark bars show initial R estimates, grey bars show extreme R estimates. (Reproduced with permission from Ref. 10)