

# Acceleration of plague outbreaks in the second pandemic

David J. D. Earn<sup>a,c,e</sup>, Junling Ma<sup>f</sup>, Hendrik N. Poinar<sup>b,c,d,e</sup>,  
Jonathan Dushoff<sup>a,c,e</sup>, and Benjamin M. Bolker<sup>a,c,e</sup>  
[earn@math.mcmaster.ca](mailto:earn@math.mcmaster.ca)

<sup>a</sup>Department of Mathematics & Statistics,

<sup>b</sup>McMaster Ancient DNA Centre, Department of Anthropology,

<sup>c</sup>Department of Biology, <sup>d</sup>Department of Biochemistry,

and <sup>e</sup>Michael G. deGroote Institute for Infectious Disease Research,

McMaster University, 1280 Main Street West, Hamilton, ON, L8S 4K1, Canada;

<sup>f</sup>Department of Mathematics & Statistics, University of Victoria, Victoria, BC, Canada

March 12, 2020 @ 13:42

Various historical records reveal the temporal patterns of a sequence of plague epidemics in London, England, from the 14th to 17th centuries. Analysis of these records shows for the first time that later epidemics spread significantly faster (“accelerated”). Between the Black Death of 1348 and the later epidemics that culminated with the Great Plague of 1665, we estimate that the epidemic growth rate increased fourfold. Currently available data do not provide enough information to infer the mode of plague transmission in any given epidemic; nevertheless, order-of-magnitude estimates of various epidemic parameters suggest that the observed slow growth rates in the 14th century are inconsistent with direct (pneumonic) transmission. We discuss the potential roles of demographic and ecological factors such as climate change or human or rat population density in driving the observed acceleration.

plague | London | epidemic growth rate

## Significance Statement:

Epidemics of plague devastated Europe’s population throughout the Medieval and Renaissance periods. Genetic studies of small numbers of skeletal remains indicate that the causative agent of all these epidemics was the bacterium *Yersinia pestis*, but such analyses cannot identify overall patterns of transmission dynamics. Analysis of thousands of archival records from London, England, reveal that plague epidemics spread much faster in the 17th century than in the 14th century.

**P**lague epidemics have afflicted human populations since at least the 6th century (1, 2). These events have had dramatic and long-lasting effects on human demography

and behaviour, especially those outbreaks associated with the “second pandemic” (14th–19th centuries) in Europe and Asia (1, 3–5), and have inspired many theoretical studies of the ecology and evolution of infectious disease (6–13). We are now in the third pandemic (Modern Plague), with outbreaks continuing to occur in some parts of the world (14–18). Plague also remains a source of concern due to the bioterror potential of the causative agent, *Yersinia pestis* (19, 20).

Recent advances in paleogenomics have definitively established that historical plague pandemics were caused by *Y. pestis* (21, 22), as proposed in the 19th century after Yersin discovered the bacterium’s link to bubonic plague (23). Researchers have reconstructed the evolutionary history of plague and other pathogens by sequencing and reconstructing nearly complete pathogen genomes from persistent DNA fragments (21, 24, 25). The strain isolated from victims of the Black Death (London 1348) is remarkably similar to extant human strains (Modern Plague): the core genomes<sup>1</sup> of these strains are  $\approx 99.99\%$  similar (21), which makes it challenging to identify important evolutionary or ecological patterns from genomic investigations alone. Here, we complement genetic studies by exploring more traditional (historical, demographic, and epidemiological) sources of information from a 300-year span of plague outbreaks in the same location (London), revealing a striking change in plague transmission dynamics over the course of the Renaissance period, namely a fourfold increase in the initial growth rate of outbreaks.

We quantify this change without making any assumptions about the underlying transmission processes, exploiting methodology that we have developed previously for this purpose (26). We then consider how this inference can contribute to the debate concerning whether plague transmission was primarily indirect (via rat fleas) or direct (pneumonic human-to-human). We argue that strictly pneumonic transmission in the 14th century is implausible, but that beyond this the best that can be done at present is to highlight the biological complexities and uncertainties that limit the potential for further inferences.

## Data

The city of London, England, is unusual in the extent to which patterns of death and disease can be reconstructed from extant documents. We have analyzed three sources of data (Figure 1):

**London Bills of Mortality** The London Bills of Mortality (27, 28) (LBoM) provide weekly counts of deaths by cause during each of the known plague epidemics in the 16th and 17th centuries. The LBoM thus give us information specifically about mortality attributable to plague for each epidemic, which we will use as a proxy for plague incidence (6, 26). When available, the LBoM are our most reliable records (except for the epidemic of 1593; see Supplemental Information).

**Parish registers** After 1538 (29, p. 54), London parish registers provide information on mortality. We use weekly counts of deaths published online by Cummins, which are estimated to cover 80% of total deaths (30). During plague epidemics, most deaths

---

<sup>1</sup>The core genome is about 80% of the full genome, shared between all sequenced strains of *Y. pestis*.

were likely due to plague; we also adjust for non-plague mortality by estimating a baseline death rate from the years before and after an epidemic (see [Methods](#)).

**Wills and testaments** Before 1538, no direct tabulations of mortality for London are available. However, we do have detailed information on last wills and testaments, which can be used as proxies for mortality (29). In particular, we have digitized and tabulated daily counts of wills recorded in the Court of Husting (31) for the 14th century; these give us data for the plague epidemics of 1348, 1361, 1368 and 1375 (monthly counts for three of these epidemics were previously published by Cohn (32)). We use online data from the Prerogative Court of Canterbury (PCC) (33) for the plague epidemics of the 16th and 17th century. Plague epidemics in London are mentioned in historical documents from the 15th century (27, Ch. IV), but during this period wills recorded in the Courts of Husting and Canterbury were too sparse to enable identification of epidemics.

Two dates are associated with each will: the date on which it was written, and the date on which it was probated (*i.e.*, accepted by the court after the death of the testator). Traditionally, demographers have used annual counts of wills by probate date, as testators are known to have died by these dates. While this approach works well for inferring long-term trends in annual mortality, probate dates are inappropriate for reconstructing detailed time series of mortality over short time scales, as there is a variable and irregular lag between the (unknown) date of the death and the probate date; most of the wills associated with 14th century epidemics were probated several months after they were written (34) (see [Figure S2](#)).

Instead, we use the counts of wills *written* in a given time interval as a proxy for plague incidence. The date on which a will was written may precede the testator's death by a long period, but during severe epidemics will dates were likely to have been correlated with the fear of infection (and hence with disease incidence), just as internet searches for influenza symptoms can predict 21st century epidemic patterns (35).

When using wills as proxies for mortality or incidence, we must keep in mind that: 1) a much smaller proportion of the population is sampled (the will counts in [Figure 1](#) are about 10 times smaller than the counts based on parish registers); 2) the subpopulation of individuals who wrote wills is strongly biased towards merchants and others who owned property (primarily males).

For the period since 1540, the existence of multiple sources allows us to cross-check results—in particular, to test our assertion that aggregated counts of wills by date written are an adequate proxy for mortality rates—and avoid relying on any questionable or poorly sampled segments of data (36) (see [Supplemental Information](#)). In 1563 and 1593, the number of PCC wills is too small to extract a signal from the noise, but the data for the later three outbreaks verify that the numbers of wills written follow the plague epidemic patterns evident in the mortality records ([Figure 2](#)). Based on cross-correlation analysis and differences between epidemic peak times, wills appear to have lagged mortality recorded in parish registers and LBoM by approximately 3–5 weeks; our inferences about epidemic growth rates are apparently robust to this lag (see [Supplemental Information](#)).

The LBoM indicate that there were 18 other plague epidemics in London between 1563 and 1666; these were all of much smaller magnitude and are discussed only in [Supplemental Information](#). No plague epidemics have occurred in London since the end of the Great Plague in 1666: a total of 77 deaths from plague are recorded in the LBoM after 1666, never more than 5 in a single week. The last plague deaths recorded in the LBoM occurred during the week of 6 May 1679.

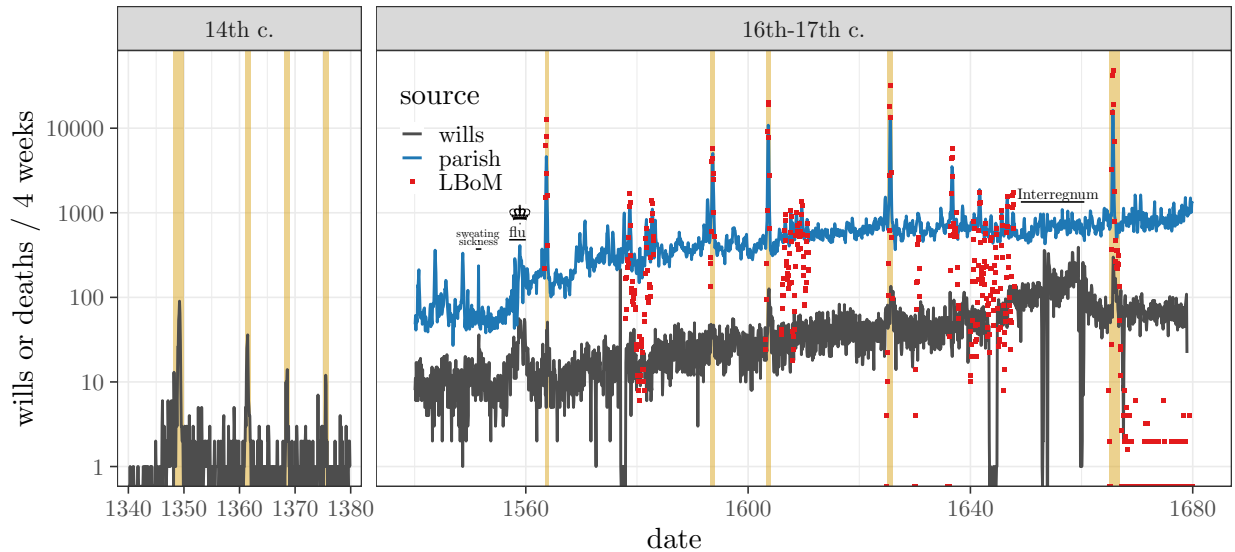
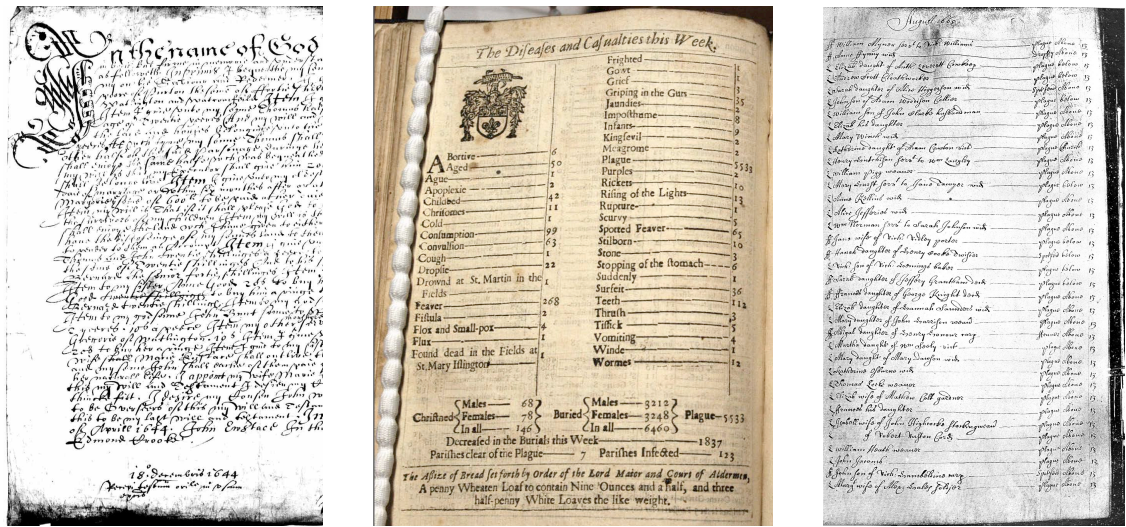
## Approach

**Estimation of epidemic growth rates** While our collection of high-resolution mortality time series during plague epidemics allows us to estimate epidemic growth rates, we are constrained by the limitations of our data—the time series for each epidemic are short and noisy, especially during the 14th century (*e.g.*, the epidemics of 1368 and 1375 peaked at a total of 5 wills written per week). Thus we chose to use simple phenomenological models with a small number of parameters that can be fitted to short time series (starting size, growth rate, and total size: see [Methods](#)), and do not attempt to separate different sources of variability or estimate the influence of various mechanistic processes (King *et al.*'s exhortation to use epidemic models that separate process from observation noise (38) applies primarily to forecasting, which we are not attempting). Furthermore, we primarily consider the initial growth rate  $r$ , which—unlike the basic reproduction number  $\mathcal{R}_0$ —can be estimated reliably from mortality data alone without any knowledge of the disease characteristics or natural history (26).

We estimate the epidemic growth rate for each epidemic as the maximum likelihood estimate of the initial growth rate of a logistic function fitted to the cumulative death curve (using differences to avoid over-confidence; see [Methods](#) for more detail), along with likelihood profile confidence intervals (CIs). To summarize the overall difference between epochs, we then use the estimates for each epidemic as observations in a linear mixed model including fixed effects of epoch (“early”/14th c. *vs.* “late”/16th-17th c.) and data source, and a random effect of outbreak year (see [below](#)), where each estimate was weighted by its certainty.

Finally, to robustly quantify the statistical significance of the difference between epochs, we ran a permutation test. We computed the null distribution of the between-epoch difference in growth rate by scrambling the estimates and associated CIs randomly across epochs. For each of the 126 possible permutations, we refitted the mixed model and extracted the estimated among-epoch difference in growth rate.

**Estimating  $\mathcal{R}_0$  and attack rate** Our analysis focuses on the epidemic growth rate  $r$ , rather than the more commonly investigated basic reproduction number  $\mathcal{R}_0$  (the expected number of secondary cases caused by a primary case in a completely susceptible population (7)), because estimates of  $\mathcal{R}_0$  depend on information about the life history of the pathogen as well as on the epidemic curve itself (39). However, if we can estimate  $\mathcal{R}_0$  then we can predict epidemic outcomes that we are unable to predict from  $r$  alone; then, by comparing observed with theoretical outcomes under different scenarios, we can make inferences about modes of transmission.



**Figure 1:** *Top left:* Part of a will proved in the Prerogative Court of Canterbury (PCC), dated 18 December 1644 (33). *Top centre:* One of the London Bills of Mortality (LBoM), for the week beginning 26 September 1665 (photo by Claire Lees). *Top right:* A parish register page from St Giles without Cripplegate, August 1665 (37, courtesy of the Wellcome Collection: <https://wellcomecollection.org/works/uutd7scc>). *Bottom:* Mortality in London, England, 1340–1380 and 1540–1680, aggregated 4-weekly, plotted on a log scale. The three distinct sources of data (see Table S2) are: weekly aggregations of burials listed in extant parish registers (29), weekly plague deaths listed in the LBoM (27), and last wills and testaments of Londoners whose wills were probated in the the Court of Husting (14th c.) or the PCC (16th-17th c.). Major plague years are highlighted in yellow. Aside from these and various minor plague years, there are notably unusual patterns during an epidemic of “sweating sickness” in 1551 (29, p. 70), the “influenza” epidemic of 1557–9 (29, p. 70) (which coincided with the end of the reign of “Bloody” Mary I and the ascension of Elizabeth I in 1558 [indicated by a crown icon]), and the absence of a monarch during the Interregnum from 30 January 1649 to 29 May 1660.



In particular, if the host population is homogeneously mixed, we can use  $\mathcal{R}_0$  to estimate the expected final size  $Z$  (the proportion of the population infected by the end of an epidemic, also called the *attack rate*) (6, 40). The *observed attack rate* based on mortality records depends in turn on the fraction of infected individuals who die from disease. We can thus compare theoretical expectations of observed attack rates under different transmission modes with historical information about observed attack rates (see [Implications](#)).

We can use the distribution of the generation interval (the times elapsed between the moment a host is infected and the times at which they infect others) to compute  $\mathcal{R}_0$  from  $r$  (39). If we only know the *mean* generation interval  $T_g$ , we can still approximate  $\mathcal{R}_0 \approx rT_g + 1$ , by assuming that the generation intervals are exponentially distributed; this approximation underestimates  $\mathcal{R}_0$  in the typical case where the generation interval distribution is less variable than an exponential distribution with the same mean.

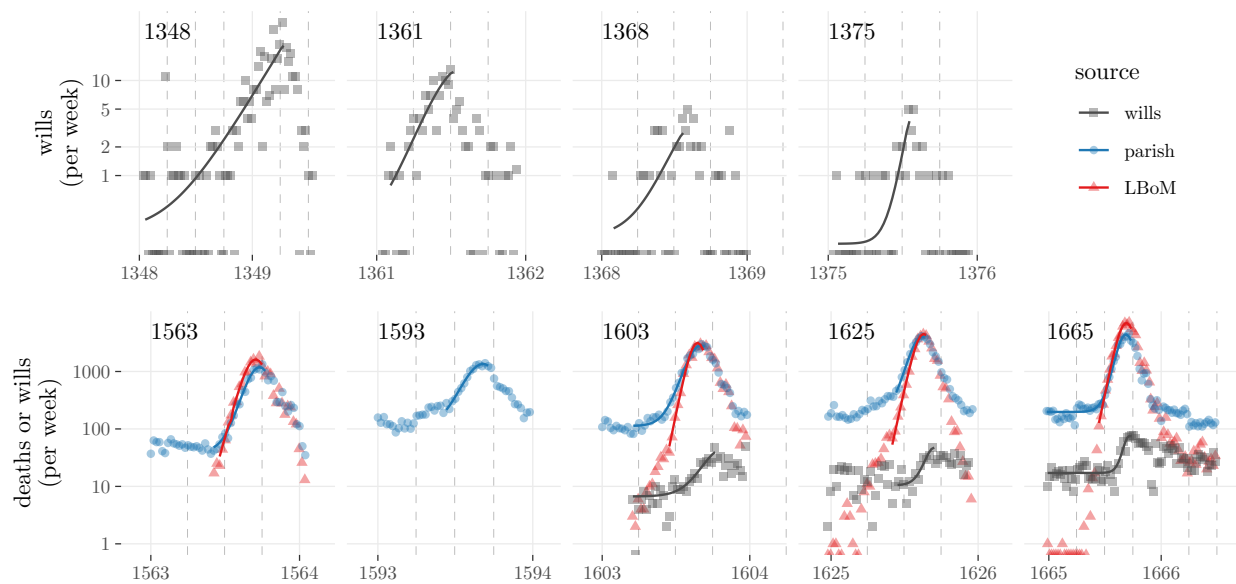
Beyond the information on generation interval that we use to derive  $\mathcal{R}_0$  from  $r$ , the observed attack rate depends on additional characteristics of an epidemic. If the population does not mix homogeneously, the final size will typically decrease (*e.g.*, (41, Fig. 6)). Since the *case fatality proportion* (CFP, the proportion of plague cases who die) is not 100%, the observed attack rate is less than the final size. Observed attack rates also decrease if some individuals are infected without showing symptoms (asymptomatic infections) (42), or without symptoms identified as plague (sub-clinical infections). Finally, the observed attack rate depends on the initial proportion of the population susceptible, which will be lower if some individuals have previously acquired immunity after previously non-fatal plague infections.

We can use independent estimates of the generation interval to estimate  $\mathcal{R}_0$  (and hence the theoretical attack rate) for London epidemics under different transmission modes. By considering the additional (extreme) assumptions that (1) the population was homogeneously mixed; (2) the population was initially 100% susceptible; (3) there were no asymptomatic infections; and (4) all infected individuals died (*i.e.* the CFP was 100%), all of which increase the observed attack rate, we can calculate an upper bound on the total mortality we would expect to observe in an epidemic with a particular transmission mode.

## Results

The fitted models do as good a job as one could hope for in capturing the initial patterns of epidemic growth evident in the noisy will counts, and agree very well with the patterns of growth in the mortality time series that are available in the later epoch ([Figure 2](#)).

The growth rates for the 14th century epidemics are lower than those for the 16th-17th centuries ([Figure 3](#)); they may also be more variable than in the later period, although some of the appearance of variability is due to larger uncertainties (presumably due to the relatively poor sampling in the early epoch). The mixed model quantifies the dramatic increase in growth rate (“acceleration”) of epidemics between epochs. The growth rates increased fourfold in the late *vs.* early epochs [ $r$  increased  $3.9\times$ , 95% CI:  $1.5\times$  to  $10.3\times$ ] (see [Methods](#)). The permutation test results support the statistical significance of this difference; the observed order of outbreaks yields the most extreme difference in growth rates out of all 126 possible orderings, giving a two-tailed  $p$  value of  $2/126 = 0.016$ .



**Figure 2:** Observed time series (points) and [phenomenological model](#) fits (lines; see [Methods](#)) that yield the estimated growth rates (listed in [Table S4](#) and plotted in [Figure 3](#)). The data sources ([Table S2](#)) are described in the main text and the caption to [Figure 1](#). For visual comparison, wills were aggregated weekly to match the frequency of mortality observations (fits to wills were based on the original daily counts). [Figures S5](#) and [S6](#) display the data during the major plague epidemics on linear rather than logarithmic scales. Vertical dashed lines show 1 April, 1 July, and 1 October of each year. In 14th c. data (top row), zero-count weeks are shown along the bottom edge of the graph.

## Discussion

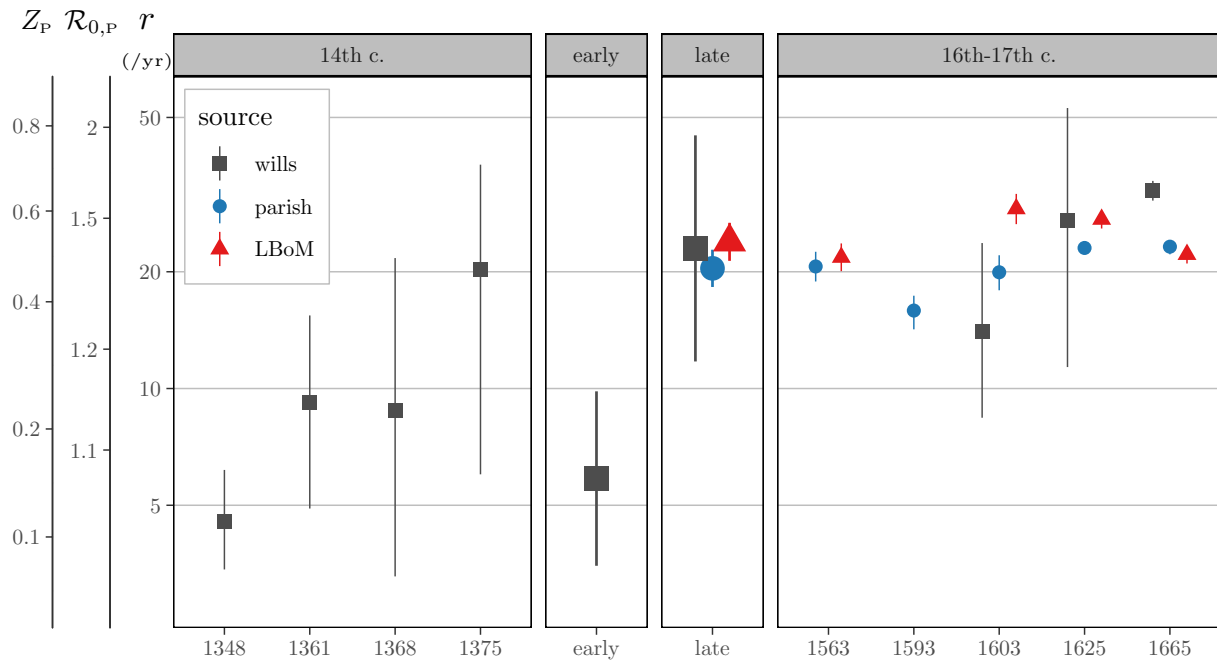
Plague epidemics in London appear to have been faster in the 16th-17th centuries than in the 14th. Our analysis shows that this difference is very likely real, rather than driven by changes in the types of documentary evidence that are available; in particular, we have shown that inferences from will counts are consistent with those from death counts when both are available, in the 17th century.

Why the later plague epidemics were faster than the earlier ones is not clear. Nevertheless, to provide some context, we consider below a number of mechanisms which *could* have contributed to increases in epidemic growth rates. We then consider what our estimated epidemic growth rates suggest about the mode of transmission during the various London plague epidemics.

### Possible causes of acceleration

What factors could have caused plague epidemics in London in the 16th and 17th centuries to accelerate relative to the earlier plagues in the 14th century?

**Pathogen or host evolution** Faster growth could be a consequence of the pathogen evolving to infect hosts more effectively, or to remain infectious for longer. However, such



**Figure 3:** Estimated initial growth rates ( $r$ ) and 95% profile CIs on a logarithmic scale, for each of the epidemics shown in Figure 2 (see also Table S4). Center panels show mixed-model estimates (see Methods) of overall average growth rate for each epoch (early: 14th c. *vs.* late: 16th-17th c.) and data source. To aid interpretation, the additional vertical scales show the implied intrinsic reproductive number ( $\mathcal{R}_{0,P}$ ) and epidemic final size ( $Z_P$ ) under the assumption that the generation interval distribution (39) was the same as that estimated from 20th century pneumonic plague epidemics (43).

evolution may be challenging to demonstrate, particularly in light of the strong genetic similarity between the Second Pandemic and Modern Plague strains (21). Evolution of host resistance has also been suggested as a cause of changes in plague dynamics (11), though we have no particular mechanisms in mind for how the evolution of resistance over the course of centuries could accelerate epidemics.

**Shifts in transmission mode** Modes of transmission may have differed between epochs, which could account for differences in epidemic growth rates. The reservoir host of *Y. pestis* in London was rat populations where it was primarily spread by flea vectors such as *X. cheopis* (14, p. 54). This transmission mode gives rise to bubonic plague in humans, as a consequence of zoonotic spillover; epidemic growth rate is driven by spread among rats and fleas, with human infection and mortality as a secondary consequence. In contrast, when *Y. pestis* infection in humans spreads to the lungs, it gives rise to pneumonic plague, which can then spread directly from human to human like other severe respiratory infections. It has also been more recently hypothesized that human ectoparasites (including lice and human fleas) could have supported indirect plague transmission without involving rats (13, 44).

Pneumonic plague probably has shorter generation intervals than bubonic plague (43, 45), and is often thought to spread much faster at the population level (43); one *potential*



explanation for our observations is that 14th century epidemics were primarily bubonically transmitted while 16th-17th century epidemics were primarily pneumonic. With currently available data, we are unable to confirm or reject this hypothesis, but we consider below whether any inferences about transmission mode can be made at present.

**Ecological and demographic changes** The complete multi-species ecology of *Y. pestis* is extremely complicated (46). We focus on humans because the only data available for the epochs studied here concern humans, and not on the rest of the ecological community supporting *Y. pestis*. Human population size and population density in London increased enormously from the 14th to 17th centuries (47–49) (Table S1 and Figure S1), and density increases were exacerbated in the 17th century by the isolation of the sick and their contacts in pest houses [A. Carmichael, *pers. comm.*]. Since bubonic plague epidemics are driven by spillover from rat-flea epidemics, human population density would only have a direct effect on the speed of pneumonic epidemics. However, changes in human population (and living conditions) almost certainly affect rat, and flea, density (50). In addition, higher rat densities make it more likely that fleas departing dying rats end up on susceptible rat hosts. The magnitudes of these effects are extremely difficult to estimate, but in Supplemental Information we explore them in a (crude) quantitative manner, without attempting to model the truly complex population dynamics that occur even during a single epidemic (45, 51).

**Environmental change** The external environment represents the third side of the host-parasite-environment “disease triangle” (52). Northern European climates changed significantly between the 14th and 16th–17th centuries (the coldest period of the “Little Ice Age” occurred in the 17th century (53) (54, Figure 1)). Furthermore, late-epoch epidemics occurred later in the calendar year, and hence in different seasons (Figure 2). A variety of studies link climate (*e.g.*, overall wetness/dryness) or weather (*e.g.*, annual precipitation or temperature) to plague incidence or transmission (55–61). Climatic changes may have been responsible for the changes in plague epidemics in London between 1348 and 1666, but it is challenging to make reliable inferences due to the lack of consensus about climate and weather variations in Medieval/Renaissance Europe (62).

## Implications of growth rate estimates for transmission mode

One might hope to exploit ancient DNA (2, 21) to identify the transmission mode during London plague epidemics. Unfortunately, there are no apparent genetic differences between strains that have caused indirect or direct transmission, so it will be difficult to determine from genetic analyses whether one mode or the other was dominant in any given Medieval/Renaissance epidemic.

Our growth rate estimates provide a starting point from which we can begin to explore the extents to which direct or indirect transmission played a role in the various London plague epidemics. Current data are too sparse and limited to make precise predictions (26, 63), but we consider the extremes of estimated parameter ranges that would favour each mode of transmission and ask whether the resulting predictions of attack rates are consistent with the rough knowledge of epidemic sizes available from the literature.

**In the early epoch, primarily pneumonic transmission is implausible.** If we assume pneumonic transmission, and use generation-interval estimates from 20th century outbreaks (43) (*i.e.*, assuming that the natural history of infection in London was similar to that of Modern pneumonic plague), in the early epoch we find  $\mathcal{R}_{0,P} = 1.1$  [CI (1.04, 1.16)], which implies an attack rate of  $Z_P = 15\%$  [CI (7.5%, 27%)] (left axes in Figure 3). This estimate of total mortality is the largest possible for the early-epoch epidemics (violations of any of the [four extreme assumptions discussed in Approach](#) would lower the total mortality). Total mortality of 15% is implausibly low in comparison to typical estimates for 14th century plague epidemics (*e.g.*, “[b]etween the years 1346 and 1352, [plague] caused the death of . . . one-third of the world’s population at that time” (19)) and certainly inconsistent with the careful and comprehensive analysis of Creighton for the 1348–49 epidemic in London specifically (“the mortality was about one-half the population” (27, p. 128)). It is therefore unlikely that 14th century plague epidemics in London were pneumonic.

**In the early epoch, growth rates appear to be consistent with bubonic plague.** What if we assume bubonic transmission? Unfortunately, the expected total human mortality from a bubonic plague epidemic depends on several poorly known aspects of rat ecology. For the generation time  $T_g$  for bubonic plague we use a rough estimate of 18 days ( $\approx 0.05$  yr) based on the flea incubation period of 9–26 days (45), which dominates the time scales of other processes involved in rat-flea plague dynamics (see [Supplemental Information](#) for further detail). In combination with a growth rate estimate of  $r \approx 6/\text{yr}$ , we obtain  $\mathcal{R}_{0,B} \approx 1.3$  and  $Z_B \approx 42\%$ . However, this estimate of the final size gives the fraction of rats, not humans, infected over the course of the epidemic. Translating this rat final size to the final size of the human epidemic requires us to quantify the amount of spillover from rats to humans. In particular, we need to know the rat-to-human ratio and the number of infected humans per infected rat, which can both be *crudely* estimated as 1:1 (51) (though these ratios will vary during the course of plague epidemics). Given an estimated CFP of 70–80% for early epoch plagues (64), our (very rough) mortality predictions are reasonably consistent with the observed human mortality rates.

**In the late epoch, primarily pneumonic transmission is unlikely but cannot be ruled out.** Now suppose that the late-epoch plagues were pneumonic with a generation interval similar to that of Modern pneumonic plague. The implied reproduction number (Figure 3) would be  $\mathcal{R}_{0,P} = 1.4$  [CI (1.2, 2.1)], which yields a maximum attack rate of 51% [CI (27%, 81%)], compared to the estimated  $\approx 20\%$  of the population who died in each of the late epoch epidemics (30, p. 4). Given that the CFP for (Modern) pneumonic plague is nearly 100% (20, 65–67), reconciling the predicted attack rates with observed mortality would require strong effects of heterogeneity; a large fraction of individuals resistant to plague due to prior non-fatal plague infections; a large proportion of asymptomatic or sub-clinical infections; or a decrease in transmission rate as the plague spread (perhaps due to behavioural avoidance of transmission). The effects of heterogeneity and behavioural avoidance may indeed have been stronger in the late epoch, as a result of concentration of the wealthy in central London and tendency for them to flee London during plague epidemics (30, p. 4). While we still consider direct (pneumonic) transmission in late-epoch London

plague epidemics unlikely because of the strong effects that would be needed to reduce the expected mortality rates to the observed levels, we cannot rule it out.

**In the late epoch (as in the early epoch), growth rates appear to be consistent with bubonic plague.** Similar calculations with the estimated  $r$  for the late epoch and the approximate generation time for bubonic plague give estimated  $\mathcal{R}_{0,B} = 2.1 [1.5, 3.5]$  and predicted rat attack rates of  $Z_B = 83\% [61\%, 96\%]$ . The human  $Z$  would be identical, if we make the same assumptions as above. This attack rate is even higher than that the 51% predicted for pneumonic plague, but the CFP for bubonic plague was probably only 30–60% in the 16th–17th centuries (*e.g.*, (19, 67)), which could bring the total mortality roughly in line with the observed  $\approx 20\%$ . The lower CFP of bubonic plague would also lead to a build-up in the population of individuals who had survived previous plague epidemics and retained immunity (reducing the susceptible population available to be attacked), which in combination with heterogeneity and asymptomatic infections makes bubonic transmission plausible for the late epoch.

**Other possibilities.** As mentioned above, it is possible that human ectoparasites could have supported indirect plague transmission (13, 44). While the biological details of this transmission mode are still uncertain, we can guess that the generation time of this transmission mode would be intermediate between direct and indirect rat-flea transmission, giving rise to estimates of  $\mathcal{R}_0$  and  $Z$  intermediate between those discussed above. Investigations of plague biology in fleas have also raised the possibility that rat transmission may have shorter generation intervals than previously thought, which would lower estimated  $\mathcal{R}_0$  values for bubonic plague (68, 69) (but cf. (70)).

## Conclusion

We have estimated epidemic growth rates for all the recorded plague epidemics in London, England, during the second pandemic (1348–1666). The major plagues of the 16th and 17th centuries grew much faster than the early plagues of the 14th century when the Black Death first invaded human populations. This conclusion is based solely on estimated growth rates—in particular, our estimates do not rely on assumptions about the mode of transmission or natural history of infection. The cause of this substantial “acceleration” is currently unclear.

We considered the implications of our estimated growth rates with respect to the mode of transmission and concluded that direct transmission (pneumonic plague) was unlikely to have been the primary mode in the early epoch. Definite conclusions about indirect transmission (bubonic plague) are difficult because of the substantial uncertainties about rat and flea ecology during this period.

We have created and assembled—and made available with this paper—machine-readable databases comprising thousands of handwritten records spanning more than 300 years. People living in London in 1348 or 1665 could not have imagined how these records might be used hundreds of years later. The time series that have emerged have allowed us to reveal and characterize large changes in patterns of infectious disease transmission. In concert with

other forms of information (30, 71), these records should continue to contribute to a broader and more insightful view of population-level processes in early eras of history.

## Methods

**Growth rate estimates** We have previously developed methods to use mortality data to estimate initial epidemic growth rates (26). The naïve approach of fitting a straight line to the relationship between logarithmic death counts and time at the beginning of the epidemic depends sensitively on the fitting window and produces overly narrow confidence intervals. Instead, we use maximum likelihood estimation to fit a curve that resembles the shape of a deterministic epidemic model (see [Phenomenological Model](#) below) and compute likelihood profile confidence intervals (72); this approach yields robust estimates of initial growth rates and accurate assessments of uncertainty. All variation around the fitted line is assumed to arise from observation errors. See Ma *et al.* (26) for methodological details.

In contrast to Ma *et al.* (26), who assumed a Poisson error model to fit simulations that were themselves generated with that assumption, the results shown here applied a negative binomial error model, which allows for the broader range of variation found in real data. We initially tried both logistic and Richards (73) models for the epidemic curve (i.e., the expected mean of the data) and both Poisson and negative binomial models for the distribution around the mean. These variants gave qualitatively similar conclusions about the pattern of growth rates, but we found that the Richards model was too unstable to fit reliably across all epidemics, so we reverted to fitting the logistic model for all epidemics. Similarly, the negative binomial model gave unstable results when the distributions were too close to Poisson; when the negative binomial dispersion parameter was estimated as  $> 10^4$ , we reverted to a Poisson fit (see [Supplemental Information](#) for further details of the optimization).

In order to summarize the differences between epochs, we fitted a linear mixed model (using the `glmmTMB` package in R (74)) to the estimates of log growth rates. To account for differing precision of the estimates across epidemics, we approximate the standard error of the estimates as the width of the 95% confidence interval divided by 3.92 (the width in standard errors of a 95% Normal CI); square it to find the variance; and scale the residual variance for each epidemic curve by this observed variance. This procedure is equivalent to weighting each epidemic curve by the inverse of its variance. We quantified the difference in epidemic between epochs based on the predicted difference in  $\log(r)$  between wills in each epoch.

The permutation test  $p$ -value (0.016) reinforces our confidence that late-epoch epidemics were faster than early-epoch epidemics; however, since it is larger than the parametric  $p$ -values derived from the model fit (which are  $< 0.001$ ), it suggests that the confidence intervals shown in [Figure 3](#) may be too narrow.

In light of recent work investigating the possibility of initial epidemic growth that is sub-exponential, we also considered fits based on the more general growth model of Chowell and co-workers (75–77). This approach leads to slightly slower initial growth estimates, but does not change our conclusion that the later plagues were much faster than the earlier plagues.

**Phenomenological model** Following Ma *et al.* (26), we model the *cumulative* mortality curve (or will-accumulation curve; henceforth we refer to wills or death records as “mortality”) using a logistic model with an added baseline linear growth rate,

$$c(t) = bt + \frac{K}{[1 + ((K/c_0) - 1)e^{-rt}]}. \quad (1)$$

Here,  $c_0 = c(0)$  is the initial cumulative mortality,  $r$  is the initial exponential growth rate, and  $K$  is the final size of the epidemic ( $\lim_{t \rightarrow \infty} [c(t) - bt]$ ). In the fitting procedure we use the unitless parameter  $x_0 = c_0/K$ ; all parameters are fitted on a transformed, unconstrained scale (*i.e.*,  $\log(r)$ ,  $\log(K)$ , and  $\text{arctanh}(x_0)$ ). The derivative  $c'(t)$  is the instantaneous mortality rate. To avoid the inherent correlations between observations in the cumulative curve, we fit the observed data to the differences from (1):  $\Delta c(t) = c(t + \Delta t) - c(t)$ , where  $\Delta t$  is the observation interval (one day for wills and one week for mortality data). The parameter  $b$  represents baseline (non-plague) mortality. For the LBoM fits we set  $b = 0$  (since the LBoM data include only plague mortality); for the wills and parish-register fits we estimated  $b$  by computing the average observed mortality rate over the two years preceding and following the outbreak window.

As discussed above, we initially fitted a Richards model, which extends the logistic equation with a shape parameter, but found that we could not get stable fits to the smallest/noisiest epidemics; we never found qualitative differences between growth rates estimated from logistic and Richards fits. Ma *et al.* found that a *delayed* logistic curve, which adds an exponentially distributed delay between infection and mortality, estimated growth rates from mortality curves with less bias. However, the delay between infection and mortality in our data is extremely uncertain—it would represent another parameter that would have to be estimated from noisy data. For wills data, the delay could even be negative, *i.e.*, wills could have been written out of fear of *future* infection. Furthermore, we found little difference between the Richards and logistic fits—which Ma *et al.* found to have opposite biases; hence these methods are likely to bracket the growth rate estimated by the delayed logistic.

**Fitting windows** How well a given model can fit a time series may depend on the time window used for fitting (77). In order to avoid choosing a window separately for each individual epidemic, we considered window selection criteria that could be applied uniformly to all epidemics without first examining the data. In all cases we defined the end of the fitting window to be one data point past the observed peak date (the date of maximum observed deaths). For wills and parish registers, we began the fitting window at the start of the outbreak window (see below); for LBoM data, we started immediately *after* the last observation before the peak for which the mortality rate was  $\leq 1/50$  of the peak mortality rate. For two epidemics for which the data are particularly noisy (1593 and 1625), we were unable to find any simple heuristic that would automatically choose a good fitting window, and resorted to choosing the fitting window by eye.

The fitting windows used for all our fits are listed in [Supplemental Information Tables S7](#) and [S8](#). Although we experimented with different methods for determining the fitting window (e.g. using a threshold of  $1/100$  of the peak mortality rate, or using the observation after the



last local minimum before the peak), we emphasize that our choices were entirely determined by seeking “good” fits (judged by visual inspection), and not by the pattern of estimated initial growth rates. Moreover, our conclusion that later plagues were much faster than the earlier plagues is robust to choices we made about window selection.

**Outbreak years and outbreak windows** We refer to the calendar year in which an epidemic begins as the *outbreak year*. For each epidemic, we define an *outbreak window* that is the maximum time interval to be considered for fitting. For the 14th century epidemics, we defined the outbreak window to be the full outbreak year, except for outbreak year 1348 when the epidemic spanned both 1348 and 1349 (so we took the outbreak window to be the full length of both calendar years). For the late epoch epidemics, we defined the outbreak windows as sequences of consecutive weeks during which at least one plague death was listed in the LBoM. See [Supplemental Information Tables S7](#) and [S8](#).

## Acknowledgements

The plague mortality data were photographed and/or entered by Seth Earn, Kelly Hancock, Claire Lees, James McDonald, David Price and Hannah Price. Valerie Hart facilitated work at the Guildhall Library in London. Josephine Earn provided historical context. Non-exponential comparison fits were conducted by David Champredon. Sang Woo Park assisted in finding published estimates of auxiliary epidemic parameters. Alex Bushby assisted with compiling probate dates. Sigal Balshine, Ann Carmichael and David Price made helpful comments. Early stages of this research were supported by a J. S. McDonnell Foundation Research Award to DE (<https://www.jsmf.org/grants/d.php?id=2006014>). All the authors were supported by Discovery grants from the Natural Sciences and Engineering Research Council of Canada (NSERC). HP was supported by the Social Sciences and Humanities Research Council of Canada (SSHRC). SHARCnet (<http://www.sharcnet.ca>) provided computational resources.

## References

- [1] McNeill WH (1976) *Plagues and peoples*. (Anchor Books, New York).
- [2] Wagner DM, et al. (2014) *Yersinia pestis* and the Plague of Justinian 541-543 AD: a genomic analysis. *Lancet Infectious Diseases* 14:319–326.
- [3] Pepys S (1669) *The Diary of Samuel Pepys, 1660–1669*.
- [4] Latham R, ed. (1985) *The Shorter Pepys*. (University of California Press, Berkeley).
- [5] Defoe D (1722) *A Journal of the Plague Year*.
- [6] Kermack WO, McKendrick AG (1927) A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London Series A* 115:700–721.
- [7] Anderson RM, May RM (1991) *Infectious Diseases of Humans: Dynamics and Control*. (Oxford University Press, Oxford).

- [8] Keeling MJ, Gilligan CA (2000) Metapopulation dynamics of bubonic plague. *Nature* 407:903–906.
- [9] Bacaër N (2012) The model of Kermack and McKendrick for the plague epidemic in Bombay and the type reproduction number with seasonality. *Journal of mathematical biology* 64(3):403–422.
- [10] Silva O (2016) Black Death – model and simulation. *Journal of Computational Science*.
- [11] Lewnard JA, Townsend JP (2016) Climatic and evolutionary drivers of phase shifts in the plague epidemics of colonial India. *Proceedings of the National Academy of Sciences* 113(51):14601–14608.
- [12] Yue RPH, Lee HF, Wu CYH (2016) Navigable rivers facilitated the spread and recurrence of plague in pre-industrial Europe. *Scientific Reports* 6:34867.
- [13] Dean KR, et al. (2018) Human ectoparasites and the spread of plague in Europe during the Second Pandemic. *Proceedings of the National Academy of Sciences* p. 201715640.
- [14] Perry RD, Fetherston JD (1997) *Yersinia pestis*—etiologic agent of plague. *Clinical microbiology reviews* 10(1):35–66.
- [15] Neerinx S, Bertherat E, Leirs H (2010) Human plague occurrences in Africa: an overview from 1877 to 2008. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 104(2):97–103.
- [16] World Health Organization (2010) Human plague: review of regional morbidity and mortality, 2004–2009: Introduction. *Weekly Epidemiological Record* 85(06):40–45.
- [17] European Centre for Disease Prevention and Control (2017) Plague outbreak, Madagascar, 2017. *Stockholm: ECDC* pp. 1–10.
- [18] Roberts L (2017) Echoes of Ebola as plague hits Madagascar. *Science* (6362):430–431.
- [19] Ligon BL (2006) Plague: a review of its history and potential as a biological weapon in *Seminars in pediatric infectious diseases*. (Elsevier), Vol. 17, pp. 161–170.
- [20] Centers for Disease Control and Prevention (Accessed 11 June 2018) Emergency preparedness and response: Frequently asked questions (FAQ) about plague, Technical report.
- [21] Bos KI, et al. (2011) A draft genome of *Yersinia pestis* from victims of the Black Death. *Nature* 478(7370):506–510.
- [22] Benedictow OJ (2011) *What Disease was Plague? On the Controversy over the Microbiological Identity of Plague Epidemics of the Past*, Brill’s Series in the History of the Environment. (Brill) Vol. 2.
- [23] Butler T (2014) Plague history: Yersin’s discovery of the causative bacterium in 1894 enabled, in the subsequent century, scientific progress in understanding the disease and the development of treatments and vaccines. *Clinical Microbiology and Infection* 20(3):202–209.

- [24] Schuenemann VJ, et al. (2011) Targeted enrichment of ancient pathogens yielding the pPCP1 plasmid of *Yersinia pestis* from victims of the Black Death. *PNAS – Proceedings of the National Academy of Sciences of the U.S.A.* 108(38):E746–E752.
- [25] Devault AM, et al. (2014) Second-pandemic strain of *Vibrio cholerae* from the Philadelphia cholera outbreak of 1849. *New England Journal of Medicine* 370(4):334–340.
- [26] Ma J, Dushoff J, Bolker BM, Earn DJD (2014) Estimating initial epidemic growth rates. *Bulletin of Mathematical Biology* 76(1):245–260.
- [27] Creighton C (1965) *A history of epidemics in Britain*. (Frank Cass & Co. Ltd., London and Edinburgh) Vol. 1, 2nd edition.
- [28] Tien JH, Poinar HN, Fisman DN, Earn DJD (2011) Herald waves of cholera in nineteenth century London. *Journal of the Royal Society Interface* 8(58):756–760.
- [29] Slack P (1990) *The impact of plague in Tudor and Stuart England*. (Clarendon Press, Oxford: New York), p. 443. Reprint. Originally published: London; Boston: Routledge & K. Paul, 1985.
- [30] Cummins N, Kelly M, Ó Gráda C (2016) Living standards and plague in London, 1560–1665. *The Economic History Review* 69(1):3–34.
- [31] Sharpe RR (1889) *Calendar of Wills Proved and Enrolled in the Court of Husting, London, A.D. 1258 – A.D. 1688. Preserved Among the Archives of Corporation of the City of London, at the Guildhall. Edited, with Introduction, by Reginald R. Sharpe ... Printed by Order of the Corporation of the City of London Under the Direction of the Library Committee*. (J. C. Francis).
- [32] Cohn Jr. SK (2003) *The Black Death Transformed: Disease and Culture in Early Renaissance Europe*. (Arnold, London, United Kingdom).
- [33] National Archives (UK) (accessed 2018) Research guides: Wills 1384-1858.
- [34] Bushby A (2019) *Demographic patterns in Medieval London inferred from wills probated in the Court of Husting, 1259–1689*. (MSc thesis, McMaster University, Canada).
- [35] Ginsberg J, et al. (2009) Detecting influenza epidemics using search engine query data. *Nature* 457:1012–1015.
- [36] Roosen J, Curtis DR (2018) Dangers of noncritical use of historical plague data. *Emerg Infect Dis* 24(1):103–110.
- [37] Bell WG (1924) *The great plague in London in 1665*. (John Lane; Dodd, Mead and company, London : New York). OCLC: 01941334.
- [38] King AA, de Cellès MD, Magpantay FMG, Rohani P (2015) Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. B* 282(1806):20150347.
- [39] Wallinga J, Lipsitch M (2007) How generation intervals shape the relationship between growth rates and reproductive numbers. *Proceedings of the Royal Society of London, Series B* 274:599–604.

- [40] Ma J, Earn DJD (2006) Generality of the final size formula for an epidemic of a newly invading infectious disease. *Bulletin of Mathematical Biology* 68(3):679–702.
- [41] Rolls DA, Wang P, McBryde E, Pattison P, Robins G (2015) A simulation study comparing epidemic dynamics on exponential random graph and edge-triangle configuration type contact network models. *PLOS ONE* 10(11):1–36.
- [42] Marshall Jr JD, Quy D, Gibson FL (1967) Asymptomatic pharyngeal plague infection in Vietnam. *The American Journal of Tropical Medicine and Hygiene* 16(2):175–177.
- [43] Gani R, Leach S (2004) Epidemiologic determinants for modeling pneumonic plague outbreaks. *Emerging Infectious Diseases* 10(4):608–614.
- [44] Laroche M, Raoult D, Parola P (2018) Insects and the transmission of bacterial agents. *Microbiology spectrum* 6(5).
- [45] Laperrière V, Badariotti D, Banos A, Müller JP (2009) Structural validation of an individual-based model for plague epidemics simulation. *Ecological Complexity* 6(2):102–112.
- [46] Dubyanskiy VM, Yeszhanov AB (2016) Ecology of *Yersinia pestis* and the epidemiology of plague in *Yersinia pestis: Retrospective and Perspective*. (Springer), pp. 101–170.
- [47] Finlay R (1981) *Population and metropolis: the demography of London, 1580-1650*, Cambridge Geographical Studies. (Cambridge University Press, Cambridge) Vol. 12.
- [48] Finlay R, Shearer B (1986) Population growth and suburban expansion. *London 1500(1700)*:37–57.
- [49] Krylova O (2011) *Predicting epidemiological transitions in infectious disease dynamics: Smallpox in historic London (1664-1930)*. (PhD thesis, McMaster University, Canada).
- [50] McCormick M (2003) Rats, communications, and plague: Toward an ecological history. *The Journal of Interdisciplinary History* 34(1):1–25.
- [51] Keeling MJ, Gilligan CA (2000) Bubonic plague: a metapopulation model of a zoonosis. *Proceedings of the Royal Society of London B* 267(1458):2219–2230.
- [52] Scholthof KBG (2007) The disease triangle: pathogens, the environment and society. *Nature Reviews Microbiology* 5:152–156.
- [53] Matthes FE (1939) Report of committee on glaciers, April 1939. *Eos, Transactions American Geophysical Union* 20(4):518–523.
- [54] Jones PD, Ogilvie A, Davies T, Briffa K (2013) *History and climate: memories of the future?* (Springer Science & Business Media).
- [55] Cavanaugh DC (1971) Specific effect of temperature upon transmission of the plague bacillus by the oriental rat flea, *Xenopsylla cheopis*. *The American Journal of Tropical Medicine and Hygiene* 20(2):264–273.
- [56] Xu L, et al. (2011) Nonlinear effect of climate on plague during the third pandemic in China. *Proceedings of the National Academy of Sciences* 108(25):10214–10219.

- [57] Xu L, et al. (2014) Wet climate and transportation routes accelerate spread of human plague. *Proceedings of the Royal Society B: Biological Sciences* 281(1780):20133159.
- [58] Pham HV, Dang DT, Tran Minh NN, Nguyen ND, Nguyen TV (2009) Correlates of environmental factors and human plague: an ecological study in Vietnam. *International Journal of Epidemiology* 38(6):1634–1641.
- [59] Yue RP, Lee HF (2018) Pre-industrial plague transmission is mediated by the synergistic effect of temperature and aridity index. *BMC infectious diseases* 18(1):134.
- [60] Ben Ari T, et al. (2010) Interannual variability of human plague occurrence in the Western United States explained by tropical and North Pacific Ocean climate variability. *The American journal of tropical medicine and hygiene* 83(3):624–632.
- [61] Schotthoefer AM, et al. (2011) Effects of temperature on the transmission of *Yersinia pestis* by the flea, *Xenopsylla cheopis*, in the late phase period. *Parasites & vectors* 4(1):191.
- [62] Nesje A, Dahl S, Thun T, Nordli Ø (2008) The ‘Little Ice Age’ glacial expansion in western Scandinavia: summer temperature or winter precipitation? *Climate Dynamics* 30(7-8):789–801.
- [63] Park SW, Dushoff J, Earn DJD, Poinar H, Bolker BM (2018) Human ectoparasite transmission of the plague during the second pandemic is only weakly supported by proposed mathematical models. *PNAS – Proceedings of the National Academy of Sciences of the U.S.A.* 115(34):E7892–E7893.
- [64] McEvedy C (1988) The bubonic plague. *Scientific American* 258(2):118.
- [65] Kool JL, Weinstein RA (2005) Risk of Person-to-Person Transmission of Pneumonic Plague. *Clinical Infectious Diseases* 40(8):1166–1172.
- [66] World Health Organization (Accessed 12 June 2018) International travel and health: Plague, Technical report.
- [67] World Health Organization (Accessed 16 June 2018) Plague: Key facts, Technical report.
- [68] Eisen RJ, et al. (2006) Early-phase transmission of *Yersinia pestis* by unblocked fleas as a mechanism explaining rapidly spreading plague epizootics. *Proceedings of the National Academy of Sciences* 103(42):15380–15385.
- [69] Eisen RJ, Wilder AP, Bearden SW, Montenieri JA, Gage KL (2007) Early-phase transmission of *Yersinia pestis* by unblocked *Xenopsylla cheopis* (Siphonaptera: Pulicidae) is as efficient as transmission by blocked fleas. *Journal of Medical Entomology* 44(4):678–682.
- [70] Hinnebusch BJ, Bland DM, Bosio CF, Jarrett CO (2017) Comparative ability of *Oropsylla montana* and *Xenopsylla cheopis* fleas to transmit *Yersinia pestis* by two different mechanisms. *PLOS Neglected Tropical Diseases* 11(1):e0005276.
- [71] Abbott A (2017) The ‘time machine’ reconstructing ancient Venice’s social networks. *Nature* 546:341–344.



- 605 [72] Bolker BM (2008) *Ecological models and data in R*. (Princeton University Press).
- 606 [73] Richards F (1959) A flexible growth function for empirical use. *Journal of experimental*  
607 *Botany* 10(2):290–301.
- 608 [74] Brooks ME, et al. (2017) glmmTMB balances speed and flexibility among packages for  
609 zero-inflated generalized linear mixed modeling. *The R Journal* 9(2):378–400.
- 610 [75] Viboud C, Simonsen L, Chowell G (2016) A generalized-growth model to characterize  
611 the early ascending phase of infectious disease outbreaks. *Epidemics* 15:27–37.
- 612 [76] Chowell G, Sattenspiel L, Bansal S, Viboud C (2016) Mathematical models to charac-  
613 terize early epidemic growth: A review. *Physics of Life Reviews* 18:66–97.
- 614 [77] Champredon D, Earn DJD (2016) Understanding apparently non-exponential out-  
615 breaks. Comment on “Mathematical models to characterize early epidemic growth: A  
616 review” by Gerardo Chowell et al. *Physics of Life Reviews* 18:105–108.