Population Studies

A Journal of Demography

A survival analysis of the last great European plagues: The case of Nonantola (Northern Italy) in 1630

Guido Alfani & Marco Bonetti

View supplementary material

Submit your article to this journal

View related articles

View Crossmark data

Citing articles: 7 View citing articles

A survival analysis of the last great European plagues: The case of Nonantola (Northern Italy) in 1630

Guido Alfani and Marco Bonetti
Bocconi University

This paper develops the first survival analysis of a large-scale mortality crisis caused by plague. For the time-to-event analyses we used the Cox proportional hazards regression model. Our case study is the town of Nonantola during the 1630 plague, which was probably the worst to affect Italy since the Black Death. Individual risk of death did not depend on sex, grew with age (peaking at ages 40–60 and then declining), was not affected by socio-economic status, and was positively associated with household size. We discuss these findings in light of the historical–demographic and palaeo-demographic literature on medieval and early modern plagues. Our results are compatible with the debated idea that ancient plague was able to spread directly from human to human. Our methods could be replicated in other studies of European plagues to nuance and integrate the findings of recent palaeo-biological and palaeo-demographic research on plague.

Supplementary material for this article is available at: https://doi.org/10.1080/00324728.2018.1457794

Introduction

Demographers and historians have long been aware of the fundamental role played by plague in shaping medieval and early modern European population dynamics. Having disappeared from the continent during the early Middle Ages, plague returned in 1347. It was the beginning of the terrible Black Death pandemic, which killed between 30 and 60 per cent of Europe's population (Del Panta 1980; Benedictow 2004). Such a catastrophic event had an enormous demographic impact, as well as important social, cultural, and economic consequences (Alfani and Murphy 2017). What is more, plague had come back to stay, as recently confirmed by palaeo-biological studies (Bos et al. 2016; Seifert et al. 2016) and, in the following three centuries or so, outbreaks of the disease were a fairly common occurrence (Biraben 1975; Livi Bacci 2000, 2007). By itself, plague was able to deeply alter European demography, establishing a 'high-mortality' regime (Clark 2007). Only from the late seventeenth century did plague begin to retreat from Europe and its final disappearance is usually counted among the factors explaining the observed decline in mortality at the beginning of the demographic transition (Livi Bacci 2000).

Although the importance of plague for pre-transitional demography can hardly be overstated, in many regards it continues to be a mysterious disease. Nevertheless, there is a scarcity of new studies, especially those of the demographic kind, providing new data. From a macro-demographic perspective, one such study recently focused on the last great plagues of the seventeenth century, suggesting fundamental differences across Europe, both in the epidemiology and the overall demographic impact of the disease (Alfani 2013a). The research situation is better at the micro level, where recent works have covered such diverse aspects as the household structure of plague mortality and its patterns of transmission (Cohn and Alfani 2007; Whittles and Didelot 2016), the within-city diffusion of plague and the seasonal pattern of mortality (Galanaud et al. 2015; Cummins et al. 2016), the age and sex structure of excess mortality from plague

(Manfredini et al. 2002), and the post-crisis 'economic recovery' that went hand in hand with the 'demographic recovery' (Alfani 2010a). However, research on plague has been steady in other fields, like general and global history, history of medicine, and palaeo-biology/palaeo-demography (Cohn 2002, 2010a; Henderson 2003; Green 2015).

Regarding palaeo-biology, recent research has confirmed the presence of DNA traces of *Yersinia pestis* (the pathogen responsible for contemporary plague) in the dental pulp of skeletons coming from a range of European plague cemeteries (Haensch et al. 2010; Bos et al. 2011, 2016). However, the way in which the infection was transmitted remains somewhat mysterious: the rapidity with which the Black Death and many subsequent plague waves spread across Europe does not fit well with the traditional description of a complex process involving two vectors (rats and fleas), a fact which continues to puzzle palaeo-biologists and historical demographers alike (Bolton 2013; Raoult et al. 2013). Moreover, again on the grounds of skeletal sources, palaeo-demographers have proposed a number of tentative findings about relevant matters, such as the differential susceptibility to plague according to sex, age, and frailty (DeWitte and Wood 2008; DeWitte 2010; DeWitte and Hughes-Morey 2012) and the overall impact of plague prevalence on survival, living standards, and stature (Barbiera and Dalla Zuanna 2009; DeWitte 2014).

Skeletons, however, might not be the best source for answers to questions about the demographic and epidemiological characteristics of ancient plagues. More traditional sources of historical demography can be explored in novel ways to provide evidence relevant to the aforementioned debates, while avoiding many of the limitations common to studies based on archaeological data (Séguy and Buchet 2013). This paper makes use of the largest existing micro-demographic database for any European plague, which we have built from archival sources. The database covers the rural town of Nonantola during the 1630 plague epidemic, the last affecting the area and possibly the most severe after the Black Death, killing about 2 million people in Northern Italy alone (30–35 per cent of its overall population; Alfani 2013a).

We perform a statistical analysis of the survival time of all residents of Nonantola, describing the effect of some baseline covariates, including age, sex, and socio-economic status, while taking into account the possible household-induced association among survival times. We provide original findings about the factors influencing the risk of death, as well as some insights into the possible ways in which the contagion was transmitted.

Data and methods

Based on tax records, 3,451 people lived in Nonantola in 1629. Located in the Po Plain, not far from the city of Modena, Nonantola belonged to the Estense State. The vast majority of its population was involved in agricultural activities, a fact that is also reflected in the household structure and the settlement pattern. Apart from a main settlement where the communication routes converged, its territory comprised three hamlets physically removed from the main centre: Redù, Bagazzano, and Rubiara. Nonantola had two additional distinctive characteristics. First, it hosted a Benedictine abbey, which throughout the Middle Ages had played a key role in improving the quality of the land. Second, the community had access to sizeable commons (woods and meadows), parts of which have survived until today: a decidedly uncommon feature in the densely populated and intensively used North Italian plain. These peculiarities explain why Nonantola has been the object of a significant amount of research, especially about its commons (Debbia 1990; Alfani 2011a). From the historical–demographic perspective, there have been some microanalyses of the local impact of the terrible 1590s famine (Alfani 2010b, 2011b, 2013b) and earlier studies of the effects of the 1630 plague (Alfani and Cohn 2007; Cohn and Alfani 2007). These studies, however, adopted entirely different analytical methods and were based on a database not nearly as complete as the one used here.

Our data come from two main sources: parish books of baptisms and burials, and a salt tax register (*Boccatico del sale*). The *Boccatico* (dating from 1629, the year before the plague) is a fiscal source and was used to determine the salt tax due from each household, according to the number of 'heads' (individuals). The register provides us with the complete household structure of Nonantola's population, as divided into six main neighbourhoods: Castello and Borgo, which together constituted the main settlement (Castello was a smaller neighbourhood, encircled by medieval defensive walls; Borgo was a larger built-up area immediately beyond the northeastern walls); Panara, which was a sparse group of large rural houses located on the strip of land bordering the Panaro river; and finally, the three aforementioned hamlets of Redù, Bagazzano, and Rubiara (see Figure 1). The *Boccatico* also informs us about
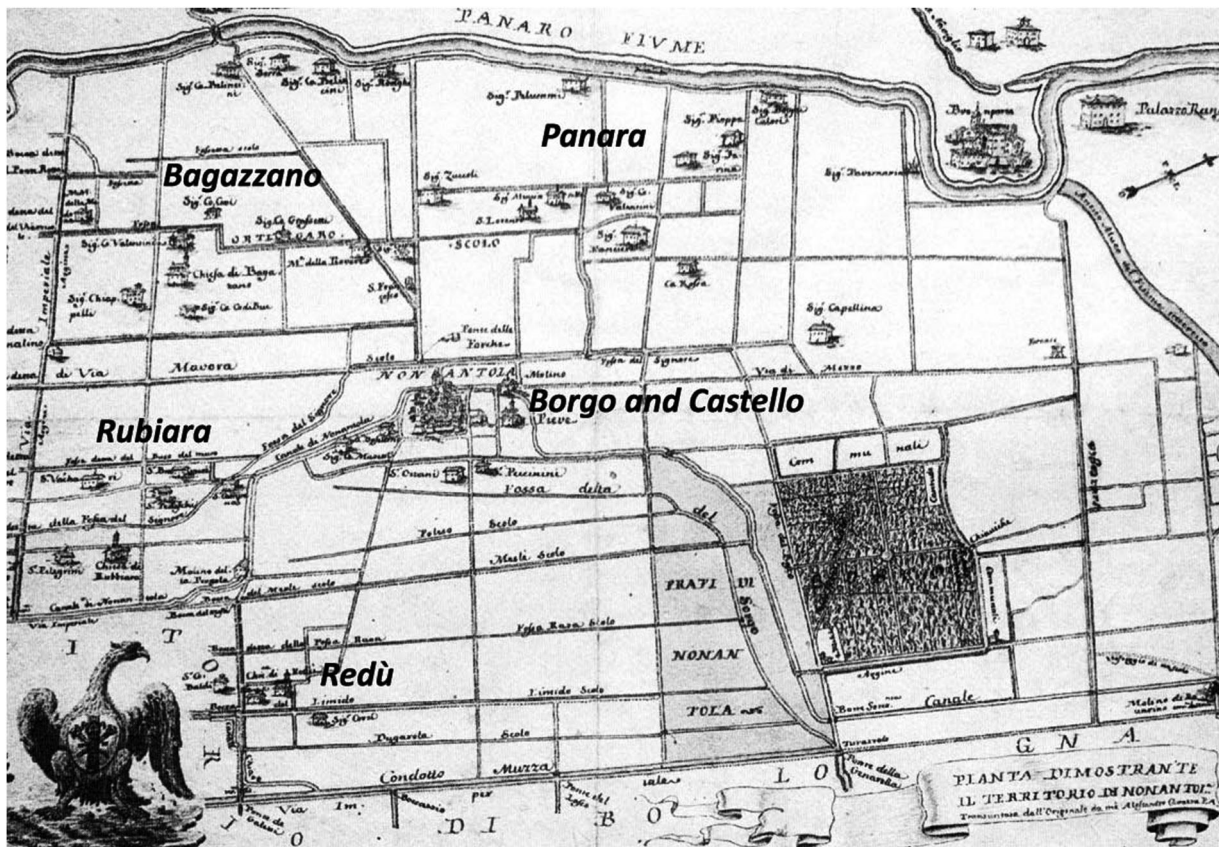
Figure 1   The territory of Nonantola, Italy, during the early modern period: main settlement and hamlets
Source: Our elaboration from an eighteenth-century map preserved at the National Archive of Modena (*Mappatorio estense*, Serie generale, n. 50).

the ages of the very young and the old (as children under five and adults aged 60 and over were exempt from the salt tax, their ages were recorded) and about the title of the head of household.

The books of burials provide us with a list of people who died in Nonantola during the plague. Nonantola's burial records are exceptional in many respects. First, they are particularly well-preserved and start very early for a rural community—in 1574. Second, the priest noted the first plague victim (buried on 9 June 1630) with the annotation *qui comincia il contaggio* ('here starts the contagion'): a fairly uncommon kind of clarification, which he was not required to give. Third, from 1587 the age at death was also recorded: a truly exceptional feature. However, for unknown reasons, the registration of burials of children aged four or younger stopped almost entirely during 1600–31. Consequently, children under five are excluded from this study. During the plague, the number of people buried increased so much that we can safely treat them all as victims of the epidemic. Indeed, over the roughly six months of the epidemic, the number of recorded burials was about 25 times the

average total for an entire 'normal' year, as seen in Figure 2. The figure also confirms that 9 June 1630 is a reliable date for the beginning of the epidemic, as in the four preceding weeks only three burials were celebrated.

Perhaps some under-registration of burials occurred, but we can presume that this problem is not serious, considering: (1) the overall high quality of the records; (2) the fact that the recording of burials never stopped during the plague; and (3) that the records seem to follow the expected pattern of rise and decline of the contagion closely. A potentially more serious issue is the possibility that part of the population fled the plague and therefore died elsewhere. Although we have no way to fully account for this problem, we have reason to believe that out-migration from Nonantola during the epidemic was at most very limited. First, we know from historical evidence and specific studies that the community was immediately quarantined. Nonantola was prepared for such an event, following a precautionary ducal decree aimed at preventing the spread of the plague within the State. Consequently, when the epidemic began in town, all the roads
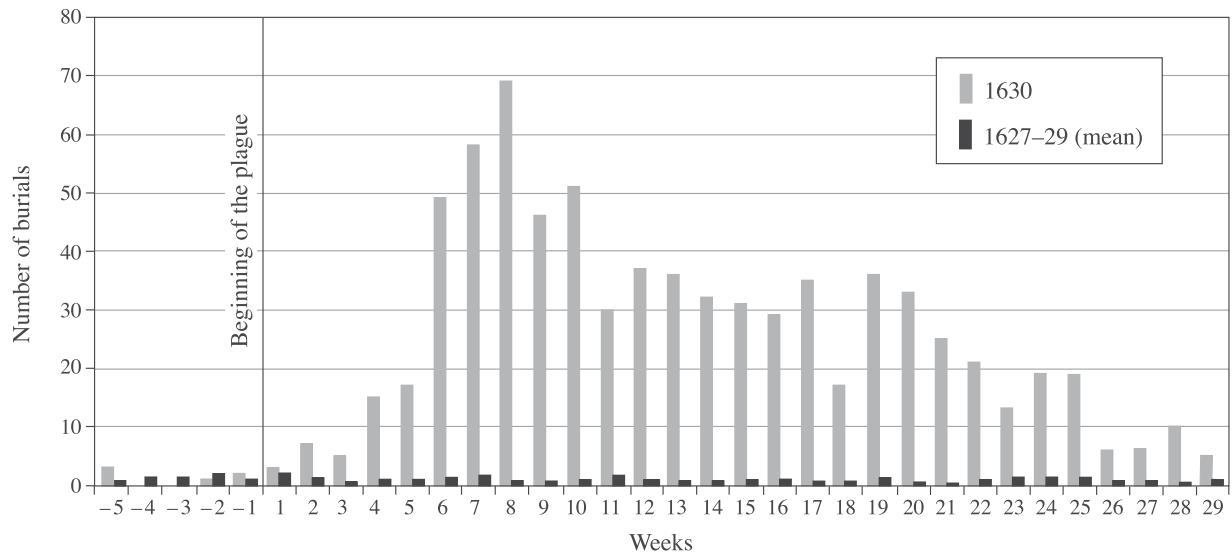
Figure 2 Weekly burials in Nonantola in 1630 compared with the preceding years
Notes: Day 0 corresponds to the start of the plague epidemic (9 June 1630), so Week 1 begins on 9 June. The mean number of deaths in the period 1627–29 can be considered indicative of a 'normal' year.
Source: Parish archive of St Michele of Nonantola, Registers of Burials.

leading out of Nonantola were already blocked by armed guards and patrols controlled its boundaries, so leaving must have been extremely difficult. A plague ward was immediately established to take care of the infected (Sperandini 1997). Additionally, there is evidence that population flights at the very beginning of the epidemic in the Estense State, when some sort of movement was still possible, were in a different direction: from the city (Modena) to the countryside (Nonantola included).

We were able to link 503 out of the 760 burials (66.2 per cent) that took place between 9 June and 31 December 1630 to entries in the *Boccatico*. Some of the burials that we could not link relate to families of which there was no trace in earlier sources (75 burials at least, related to 16 distinct families), presumably refugees from Modena who were blocked in town by the quarantine (an appeal to the duke to remove the quarantine, dated 12 February 1631, explicitly mentioned the presence of refugees who wanted to return to their homes; Sperandini 1997). As we include in the analysis only the burials of Nonantola residents listed in the *Boccatico*, the presence of these extra families is of no consequence. Also, of no consequence for the analysis is the presence of the Benedictine abbey, as monks were not recorded in either the *Boccatico* or the parish books.

The parish books of burials provide us with information on the ages of the plague victims, while the *Boccatico* provides information on the ages of the very young and the old. To discover the exact ages

of the individuals in the 5–59 age group who were spared by the plague, we looked forwards in time, collecting information about those who died between 1 January 1631 and 31 December 1700. We also looked backwards in time, for the dates of baptism of the people recorded in the *Boccatico*. We collected information on all the baptisms celebrated between 1 January 1615 and 1 January 1629, to which we added information about baptisms celebrated in the late sixteenth century from an earlier study dedicated to Nonantola's social networks (Alfani and Munno 2012). In total, we collected complete information on an additional 4,258 burials and 3,681 baptisms, to reconstruct the ages of the Nonantola residents who were spared by the plague. Note that the absence of information on the ages of plague survivors places serious constraints on all earlier works on the 1630 plague.

Our data set originally included all 3,451 individuals (from 625 households) who, in 1629, resided in one of the six neighbourhoods as recorded in the *Boccatico*. From this data set we removed 60 individuals who died between being recorded in 1629 and the beginning of the epidemic (up to 8 June 1630), and 341 individuals under the age of five (as such burials were not usually recorded). The final data set thus contains all of the 3,050 recorded individuals across 623 households who were: (1) exposed to the risk of dying of plague starting from 9 June and observed for the potential event until 31 December 1630; and (2) of an age to be recorded in the parish books of burials (also see section A1 of the

supplementary material). All (and only) the survival times of the individuals still alive at the end of 1630 were right censored. For such individuals the observed censoring time was set equal to 206 days (i.e., the exact number of days from 9 June to 31 December inclusive).

For each individual, the following information was available: neighbourhood, household of residence, household size (i.e., number of people residing together), date of burial (if the event occurred in the time frame considered), and social status. The latter, categorized as high or low, was attributed based on the title denoting status (e.g., *messere* or 'gentleman') with which some household heads were recorded. All household members were given the same status as the head. Moreover, we had almost complete information on age and sex. Information on age was missing for 148 individuals and the sex of two individuals was unknown (for one of them, age was also unknown). Consequently, an additional 149 individuals were dropped from the regression analyses, but for completeness they are included in the descriptive analyses.

For the time-to-event analyses, we use the Cox proportional hazards (PH) regression model (Cox 1972). The hazard function $h(t)$ is a function that describes the individual-specific instantaneous risk of experiencing the event of interest (here, burial) right after time t, given that one has *not* experienced the event up to time t. Such a function takes non-negative values and it provides a complete description of the distribution of the time-to-event random variable (T). Time is measured with respect to the specified origin, which in our case is 9 June 1630. From now on we will refer to the time-to-event (burial) distribution as the 'survival' distribution. Also, we will sometimes indicate the events of interest as 'deaths' even though they are really 'burials'. Deaths and burials can safely be assumed to coincide without much error (Del Panta and Rettaroli 1994).

The Cox PH model is a semi-parametric model, where the non-parametric part is a baseline hazard function $h_0(t)$ with a set of k baseline covariates $(X_1, X_2, \ldots, X_k)$ that modify it multiplicatively. Parameters are estimated by maximization of the so-called partial likelihood function (Cox 1972, 1975). The predictors that we consider here are all baseline covariates. We performed graphical tests for the hypothesis of proportionality of the hazards functions across values of the covariates (results not shown) and, overall, we are comfortable with the support that they provide to our analyses.

We fit a series of models with increasing complexity. Specifically, we apply both the basic model just described and stratified versions of the same, which seems appropriate given the rather different patterns of burial distributions across the neighbourhoods. In addition, it seems important to analyse the survival times while taking into account their possible dependence within the same household. We apply a robust method for consistent estimation of the standard errors of the regression parameters, which makes no specific hypotheses about the dependence structure of the survival times (Lin and Wei 1989). We also explored a gamma frailty model, in which the association among failure times is modelled through a frailty term, that is, a random-effect term, shared by all members of a household (see, e.g., Vaupel et al. 1979).

Results

Descriptive statistics and overall event rates

Out of a population of 3,050 individuals, 503 (16.5 per cent) died during the 1630 plague. The epidemic spread rapidly, peaking at Week 8 (between 28 July and 3 August) when 48 people died. Table 1 shows that plague struck with particular intensity in the main settlement (Borgo), where most of the trades and crafts of the town were clustered. In Castello, located similarly in the main settlement but possibly protected to a degree by the defensive walls encircling it, the percentage dying was nearly half that observed in Borgo.

The cumulative distribution function (CDF) of the death (burial) times over the follow-up period shows the progression of the infection by neighbourhood. Borgo experienced a sharp increase in the number of deaths between the 50th day (28 July) and the 75th day (22 August), and a slower but steady increase thereafter (Figure 3). In Bagazzano the infection spread with a similar pattern, but here the initial outbreak occurred towards the end of August. In Castello and Panara the epidemic spread at a more constant rate; however, by the end of the period the proportion who had died was similar to that of Bagazzano. Redù and Rubiara experienced late outbreaks, with a slower spread of the infection, resulting in lower overall mortality.

The proportions of males and females who died were similar: 252 females out of 1,546 (16.3 per cent) and 251 males out of 1,502 (16.7 per cent). Regarding age, at the onset of the epidemic the average age for the individuals included in our data set was 27.4 years, whereas the median age was 23 years (children aged 0–4 are excluded).

Table 1   Descriptive statistics for individuals and overall percentage dying by neighbourhood (percentages shown in parentheses), Nonantola, 1630

		Bagazzano	Borgo	Castello	Panara	Redù	Rubiara	Total
Gender	Missing	0	0	0	1	1	0	2
	Female	196 (50.3)	279 (52.8)	46 (53.5)	599 (50.6)	284 (49.1)	142 (50.5)	1,546 (50.7)
	Male	194 (49.7)	249 (47.2)	40 (46.5)	586 (49.5)	294 (50.9)	139 (49.5)	1,502 (49.3)
Social status	High	10 (2.6)	88 (16.7)	40 (46.5)	38 (3.2)	27 (4.7)	4 (1.42)	207 (6.8)
Age[1]	Missing	9	15	9	43	51	21	148
	Mean	27.74	28.42	27.86	27.25	27.28	26.09	27.44
	Standard deviation	17.71	18.02	18.19	18.17	18.48	17.69	18.09
	Median	24	25	23	22	22	22	23
Dead by 31 December		73 (18.7)	162 (30.7)	14 (16.3)	198 (16.7)	40 (6.9)	16 (5.7)	503 (16.5)
Total		390	528	86	1,186	579	281	3,050

[1]The ANOVA test for age across neighbourhoods returns a *p*-value of 0.659, thus indicating non-rejection. The table includes all individuals alive at the onset of the plague and aged five or above.
Source: Parish archive of St Michele of Nonantola, Registers of Baptisms and Burials. Town Archive of Nonantola, *Ruolo di popolazione/ Boccatico del Sale* (salt tax register), 1629.

The average age did not vary significantly across neighbourhoods (*p*-value = 0.659). The greatest absolute number of deaths was reported among those aged 10–14, but the proportion of people of that age group who died was not particularly high (Figure 4). The highest mortality can be observed among individuals aged around 45–59 years. Notice that the population age structure implicit in Figure 4(a) closely corresponds to Nonantola's population history pre-1630. The fact that there were fewer children aged 5–9 than 10–14 reflects the small cohorts born in the famine-stricken early 1620s (839 baptisms were celebrated in 1616–20 vs. just 762 in 1621–25). Similarly, the sharp drop in the size of the age groups 40–44 and higher reflects the fact that all those aged 41 or above in 1630 had had to survive the terrible famine of 1589–91 (moreover, the cohorts born during the famine, aged 39–41 in 1630,
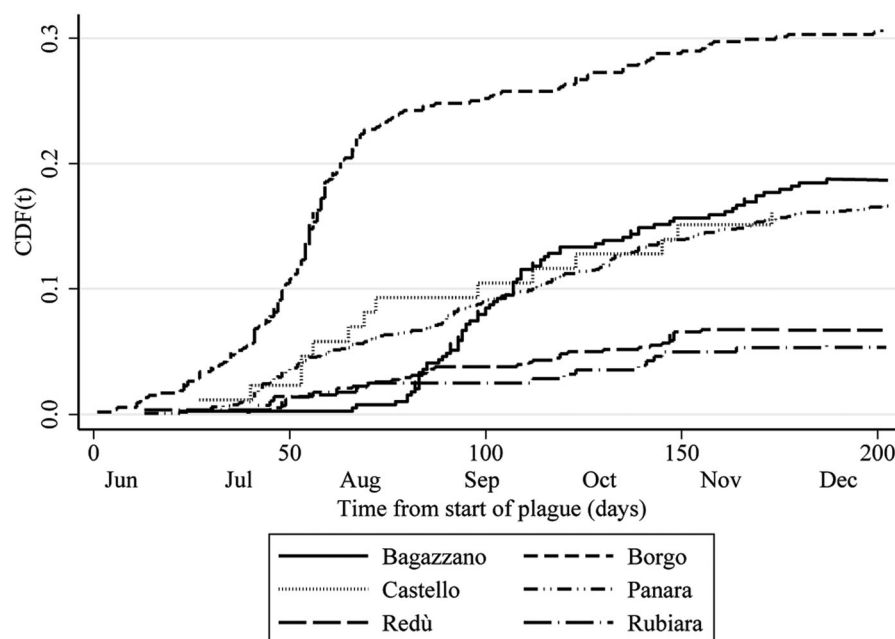
Figure 3   Cumulative distribution function (CDF) of burial time from the beginning of the epidemic, neighbourhoods in Nonantola, 1630
Notes: The value of the CDF at time *t* indicates the observed proportion of individuals who have died by time *t*. Day 0 corresponds to the start of the plague epidemic (9 June 1630).
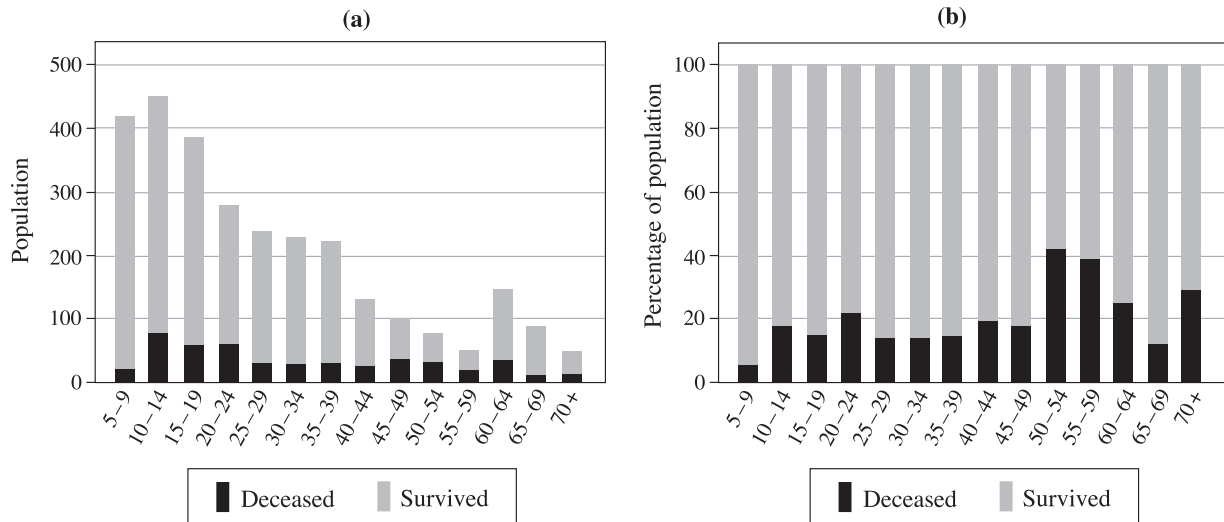*Source*: As for Table 1.

Figure 4   Absolute frequencies (a) and proportions (b) of deaths and survivors by age group, Nonantola, 1630
Source: As for Table 1.

were exceptionally small in number). Instead, the relatively large size of the age groups 60–64 (and older) results mostly from the fact that the *Boccatico* gives us the ages of all those aged 60+ who were exempt from the salt tax. Consequently, linkage attrition is zero. Some age inflation may also have occurred, possibly fuelled by the fiscal exemptions given to those aged 60+. Finally, we cannot exclude that the individuals listed in the *Boccatico* and whose age is unknown (*n* = 149) came predominantly from the age groups 40–44 to 55–59; however, this would not change our main conclusions (see later, and section A1 of the supplementary material).

A total of 207 individuals (6.8 per cent) were of high social status (Table 1). The residual 2,843 (93.2 per cent) fell in the low social status group, as expected in a pre-industrial rural community. The neighbourhood with the greatest prevalence of high-status individuals was Castello (46.5 per cent), followed by Borgo (16.7 per cent).

In Table 2 we provide additional household-level descriptive information. Households tended to be smaller in the main settlement (Borgo and Castello) than elsewhere: mean household size was 6.8 in Panara, 5.8 in Bagazzano, 5.3 in Redù, 4.8 in Rubiara, 4.3 in Borgo, and 3.0 in Castello (differences are statistically significant). Table 2 also reports on the observed proportions of households affected by plague. Among the 623 households, 224 (36.0 per cent) experienced at least one death.

Social status is a household-level variable, being defined here from the title of the head of the household. As seen in Table 2, the highest prevalences of high-status households were found in Borgo

and Castello (13.2 and 38.7 per cent, respectively). Moreover, high status is associated with smaller household size. The average high-status household size was 5.3 compared with 5.5 for low-status households, and none of the high-status households consisted of more than eleven members, while 8 per cent of low-status households did (with up to 27 members). Household infection rates show that 40 per cent of high-status households experienced at least one event, a slightly higher proportion than for low-status households (35.6 per cent).

Results from the survival models

In the initial models, the conditional hazard function for the *j*th individual is equal to:

$$h(t|\mathbf{X}_j) = h_0(t)\exp(\mathbf{X}_j^T\boldsymbol{\beta}),$$

where $\boldsymbol{\beta}$ is the vector of the coefficients associated with the vector $\mathbf{X}_j$, which contains the values of the *k* covariates for individual *j*, and $h_0(t)$ is the unspecified baseline hazard function. We included the covariates sex, household size, age, and social status in the models as modifiers of the hazard function. We allowed age to enter the model with both a linear and a quadratic term. To ease interpretation of the estimated effects, household size was included either as a numerical variable (models (1), (2), and (4)) or as a categorical variable after classification in five categories (1–3, 4–5, 6–7, 8–12, and >12 members). Indicators for the different neighbourhoods were also included (models (2) and (3)), with Panara being the baseline (reference) category.

Table 2 Descriptive statistics for households (percentages shown in parentheses), Nonantola, 1630

		Bagazzano	Borgo	Castello	Panara	Redù	Rubiara	Total
Social status	High	2	18	12	7	5	1	45
		(2.7)	(13.2)	(38.7)	(3.6)	(4.2)	(1.5)	(7.2)
Number of	Mean	5.8	4.3	3.0	6.8	5.3	4.8	5.4
household members[1]	Standard deviation	3.7	2.5	2.1	4.3	3.1	2.6	3.6
	Median	5	4	2	6	5	4	5
	<4	(26.7)	(46.3)	(64.5)	(19.5)	(30.8)	(31.8)	(31.9)
	4–5	(25.3)	(25.7)	(29.0)	(24.6)	(31.7)	(37.9)	(27.9)
	6–7	(22.7)	(17.7)	(3.2)	(25.6)	(20.8)	(18.2)	(20.7)
	8–12	(17.3)	(9.6)	(3.2)	(18.0)	(12.5)	(10.6)	(13.5)
	>12	(8.0)	(0.7)	(0.0)	(12.3)	(4.2)	(1.5)	(5.9)
At least one death by		34	73	12	76	20	9	224
31 December 1630		(45.3)	(53.7)	(38.7)	(39.0)	(16.7)	(13.6)	(36.0)
Total		75	136	31	195	120	66	623

[1]The *p*-value from the ANOVA test performed on the logarithm of the number of members across neighbourhoods is smaller than 0.0001. The table includes all the households existing in Nonantola at the onset of the plague. Note that we calculated household sizes including children below five years of age, as they potentially affected within-household plague transmission.
Source: As for Table 1.

Table 3 shows the results obtained from the estimation procedure applied to three such models (with 95 per cent confidence intervals for the individual hazard ratios), obtained by clustering the individuals by household (the original parameters are reported in section A4 of the supplementary material). We use robust estimation, since failure (death) of one individual might be linked to an event experienced by another member of the same household, meaning that death times within a group (cluster) are likely to be dependent. This fact was also confirmed by fitting a multiplicative

Table 3 Estimated parameters from Cox proportional hazards regression models (1) to (3): hazard ratios (with 95 per cent confidence intervals in parentheses)

Model		(1)	(2)	(3)
Sample size		2,901	2,901	2,901
Sex	Female	*Ref*	*Ref*	*Ref*
	Male	1.010 (0.86, 1.18)	1.043 (0.89, 1.22)	1.048 (0.89, 1.23)
Social status	Low	*Ref*	*Ref*	*Ref*
	High	1.484 (0.84, 2.63)	1.075 (0.63, 1.84)	1.107 (0.65, 1.89)
Household size		1.026 (0.99, 1.06)	1.037** (1.01, 1.07)	
Age		1.051*** (1.03, 1.07)	1.048*** (1.03, 1.07)	1.048*** (1.03, 1.07)
Age squared		0.9995*** (0.9993, 0.9998)	0.9996*** (0.9993, 0.9998)	0.9996*** (0.9993, 0.9998)
Neighbourhood	Panara	–	*Ref*	*Ref*
	Bagazzano	–	1.131 (0.69, 1.84)	1.117 (0.69, 1.80)
	Borgo	–	2.402*** (1.70, 3.39)	2.439*** (1.72, 3.45)
	Castello	–	1.251 (0.67, 2.35)	1.233 (0.66, 2.31)
	Redù	–	0.454*** (0.27, 0.77)	0.457*** (0.27, 0.78)
	Rubiara	–	0.385** (0.18, 0.82)	0.396** (0.18, 0.85)
Household size	<4	–	–	*Ref*
	4–5	–	–	0.947 (0.65, 1.38)
	6–7	–	–	0.929 (0.61, 1.42)
	8–12	–	–	1.197 (0.76, 1.89)
	>12	–	–	1.683** (1.04, 2.73)
Wald chi² (df)		*58.27 (5)*	*110.08 (10)*	*109.44 (13)*
Log pseudolik (p-value)		*–3,923.44 (<0.0001)*	*–3,852.91 (<0.0001)*	*–3,848.83 (<0.0001)*

Notes: Dependent variable is time from start of plague to burial, Nonantola, 1630. Models elaborated using Stata 12MP statistical software, using the command stcox with option cluster(household). Asterisks indicate significance levels: $0.1 \geq p > 0.05$ (*); $0.05 \geq p > 0.01$ (**); $p \leq 0.01$ (***). *Ref* indicates the reference category, *df* stands for degrees of freedom, and *pseudolik* stands for pseudo-likelihood.
Source: As for Table 1.

gamma shared frailty model that included sex, household size, age, and social status, as well as neighbourhood modifiers (results not shown). In that model, the standard deviation of the frailty term was found to be significantly greater than zero (1.45; *p*-value < 0.001), thus confirming the presence of dependence among survival times within the household structure.

The sex effect is not found to be statistically significant in any of the models in Table 3, so the epidemic apparently struck men and women equally. This is true also if we focus on people in the age group 17–35, which, for women, is characterized by intense childbearing, as discussed later. Instead, the size of the household matters when we adjust for a neighbourhood effect (models (2) and (3)), but its effect seems to be relevant only for households with more than seven members (baseline = 1–3 members). For example, living in a household with more than twelve members increases the hazard by 68 per cent when compared with the baseline group (model (3)). Living in Borgo significantly increases the risk, by more than 2.4 times when compared with Panara (the baseline neighbourhood) in both models (2) and (3), while residing in Redù and Rubiara seems to offer some protection from the risk of death. High social status fails to have a significant effect on the hazard in any of the models.

The hazard ratios of both age and age squared are significant at the 0.01 level in all three models.

Figure 5 shows the estimated hazard functions from model (2) for different ages. The highest hazards are observed for baseline ages between 40 and 65, with a peak in all the (proportional) hazard functions at about 60 days from the beginning of the epidemic.

Note that clustering the observations by household does not change the point estimates of the model parameter; it only affects the estimated variance–covariance matrix of the parameter estimator. Thus, this option can only produce a change in the statistical significance of the estimated coefficients. We have also performed these same analyses (not shown) with non-robust estimation of the variance–covariance matrix, and both social status and household size were (incorrectly) detected as being statistically significant at the 0.05 level when using model (1). This is interesting because it indicates that the robust standard errors better reflect the sampling uncertainty, while suggesting and taking into account the dependence of survival times within households.

In models (2) and (3) the hazard functions for the different neighbourhoods are assumed to be proportional to one another for any given value of the covariates. However, we know from Figure 3 that different neighbourhoods were struck by the plague with different timings and intensities. Consequently, we next extend the models to their stratified versions by neighbourhood (models (4), (5), and (6)). This
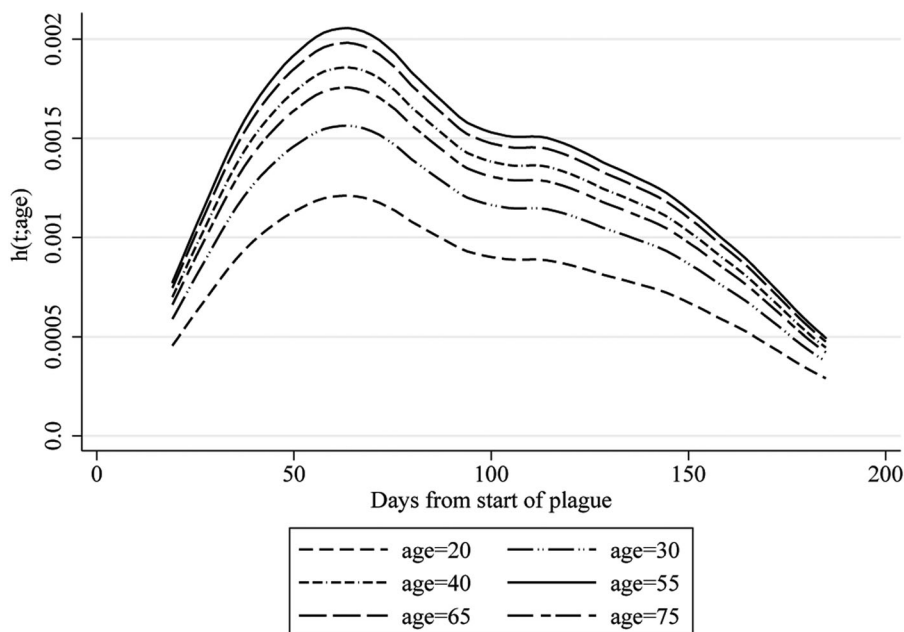

Figure 5 Model-based estimated hazard functions for burial/death, for specific ages (other covariates set to zero), from model (2), Nonantola, 1630
Note: Day 0 corresponds to the start of the plague epidemic (9 June 1630).
Source: Statistical estimates based on the same historical sources indicated for Table 1.

allows each neighbourhood to have a different, yet still unspecified, baseline hazard function, while borrowing strength from the whole data set for the estimation of the parameters of the regression part of the survival model.

The conditional hazard function for a subject in neighbourhood N is:

$$h_N(t|\mathbf{X}) = h_{0,N}(t)\exp(\mathbf{X}^T\boldsymbol{\beta}),$$

with the $\boldsymbol{\beta}$ vector parameter common across neighbourhoods and no neighbourhood indicators in $\mathbf{X}$. We still implement the robust estimation of the parameter's variance–covariance matrix by clustering the individuals by household.

Table 4 shows the results obtained from fitting two different versions of this model, in which household size is included either as a numerical variable (model (4)) or through four indicator functions for the five categories (model (5)). In model (6) we treat age as categorical rather than continuous. Apart from some relatively small changes in the values of the covariate effects, the qualitative results do not vary much from the unstratified models (1) to (3). We recommend interpreting the more robust stratified models because of the clearly different neighbourhood-specific baseline hazard functions, thus allowing for a more flexible estimation of the survival distribution within neighbourhoods (Figure 6).

In model (6) we also include four interaction terms between age groups and sex. We performed a likelihood ratio test for the four interaction terms all being equal to zero. The p-value was 0.0463, suggesting that the overall effect of the four interaction terms is borderline significant. However, the effect is due essentially to the 40–59 age group. More importantly, the likelihood ratio test can only be performed without using robust standard errors (as it is not valid otherwise) and, as a consequence, its results should be interpreted conservatively when data are clustered. We also explored the possibility of a global differential age effect on mortality by sex by adding two interaction terms to models (4) and (5), with age and age squared (not shown). The tests did not reject the null hypothesis that both interaction terms have coefficients equal to zero ($p = 0.49$ and 0.48 for the two models, respectively).

Table 4 Estimated parameters from stratified Cox proportional hazards regression models (4) to (6): hazard ratios (with 95 per cent confidence intervals in parentheses)

Model		(4)	(5)	(6)
Sample size		2,901	2,901	2,901
Sex	Female	*Ref*	*Ref*	*Ref*
	Male	1.038 (0.89, 1.21)	1.042 (0.89, 1.22)	1.887 (0.80, 4.47)
Social status	Low	*Ref*	*Ref*	*Ref*
	High	1.054 (0.64, 1.75)	1.085 (0.66, 1.79)	1.144 (0.69, 1.91)
Household size		1.037** (1.01, 1.07)	–	–
Age		1.047*** (1.03, 1.07)	1.047*** (1.03, 1.07)	–
Age squared		0.9996*** (0.9994, 0.9998)	0.9996*** (0.9993, 0.9998)	–
Age group	5–9	–	–	*Ref*
	10–19	–	–	4.504*** (2.15, 9.44)
	20–39	–	–	4.591*** (2.22, 9.49)
	40–59	–	–	11.566*** (5.49, 24.35)
	60+	–	–	5.403*** (2.51, 11.64)
Interactions age–sex	(10–19) × (sex)	–	–	0.546 (0.23, 1.32)
	(20–39) × (sex)	–	–	0.563 (0.22, 1.44)
	(40–59) × (sex)	–	–	0.353** (0.14, 0.91)
	(60+) × (sex)	–	–	0.790 (0.30, 2.08)
Household size	<4	–	*Ref*	*Ref*
	4–5	–	0.953 (0.66, 1.37)	0.965 (0.67, 1.39)
	6–7	–	0.934 (0.62, 1.41)	0.934 (0.61, 1.42)
	8–12	–	1.187 (0.76, 1.85)	1.175 (0.75, 1.83)
	>12	–	1.690** (1.04, 2.73)	1.653** (1.02, 2.68)
Wald chi² (df)		*60.09 (5)*	*59.80 (8)*	*99.45 (14)*
Log pseudolik (p-value)		*–3,144.14 (<0.0001)*	*–3,140.15 (<0.0001)*	*–3,111.72 (<0.0001)*

Notes: Dependent variable is time from start of plague to burial, Nonantola, 1630. Models elaborated using Stata 12MP statistical software, using the command stcox with options cluster(household) and strata(neighbourhood). Asterisks indicate significance levels: $0.1 \geq p > 0.05$ (*); $0.05 \geq p > 0.01$ (**); $p \leq 0.01$ (***). *Ref* indicates the reference category, *df* stands for degrees of freedom, and *pseudolik* stands for pseudo-likelihood.
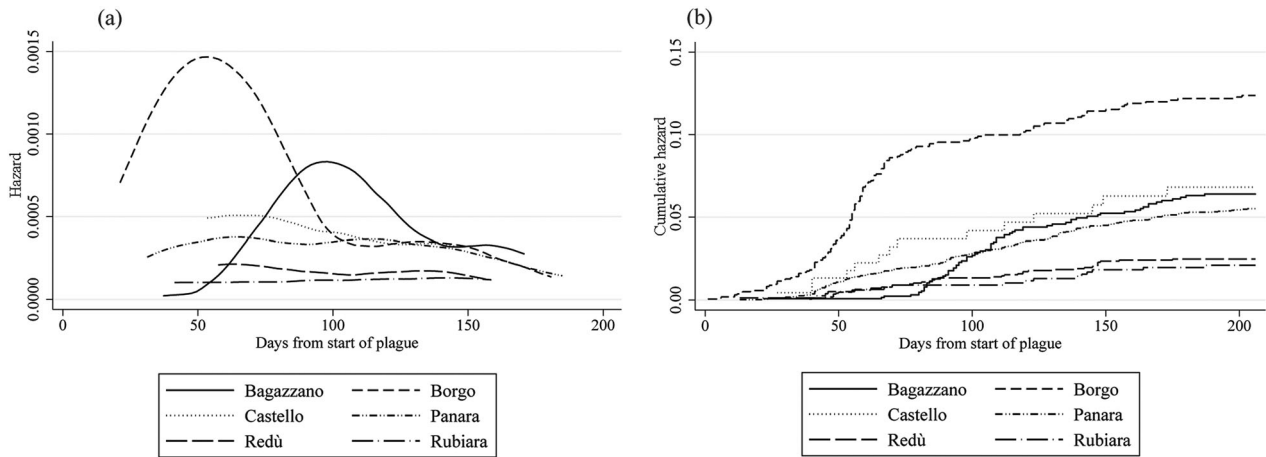*Source*: As for Table 1.

Figure 6 Predicted neighbourhood-specific baseline hazard functions (a) and cumulative baseline hazard functions (b) for burial/death, from model (4), Nonantola, 1630

Notes: Day 0 corresponds to the start of the plague epidemic (9 June 1630). From these curves, using the relationship $S(t|N, X) = \exp(-\Lambda(t|N, X)) = \exp(-\Lambda_{0,N}(t)\exp(\mathbf{X}^T\boldsymbol{\beta}))$, which expresses for each neighbourhood (N) the survival function $S(t|N, X)$ in terms of the cumulative baseline hazard function and the linear term $\mathbf{X}^T\boldsymbol{\beta}$, the survival probabilities can be obtained for any value of the covariates. See section A5 of the supplementary material for an example.

Source: Statistical estimates based on the same historical sources indicated for Table 1.

Adjusted neighbourhood-specific hazard functions, which more closely follow the mortality patterns of the different neighbourhoods, can now be estimated. Figure 6(a) shows the estimated baseline hazard functions for the different neighbourhoods when setting all covariates to zero. Figure 6(b) shows the corresponding estimated neighbourhood-specific cumulative baseline hazard functions $\Lambda_{0,N}(t)$, that is, the baseline hazard functions $\Lambda_{0,N}(t)$ integrated from zero until the argument t. The differential intensity of the death process across the different neighbourhoods is clear.

The stratified models (4) and (5) allow for a common age effect to be quantified on the complete data set, while taking into account different baseline trends for mortality across different neighbourhoods. Such a pooling of the age effects across neighbourhoods is justified by the results that we obtain by fitting individual neighbourhood-specific models involving sex, age, and household size. Figure 7 shows the age modifier $\exp(\beta_{\text{Age},1}\, \text{Age} + \beta_{\text{Age},2}\, \text{Age}^2)$ in such models, which now also allow for the effects of household size and sex to vary across neighbourhoods (social status was excluded from neighbourhood-specific models due to the scant nature of that variable).

These models confirm rather consistent effects of age in Bagazzano, Borgo, and Redù. As a consequence, pooling the information across neighbourhoods in model (4) seems convincing and produces the overall age effect represented in Figure 8. The trend and the values of age group effects in model (6) also support the smooth shape shown in Figure 8.

As a final consideration, note that the shape of the overall age effect is not the result of known faults in our database. In particular:

- It is not the result of a greater prevalence of younger individuals among the unlinked deaths, as the age structure of the unlinked deaths follows that of the linked deaths quite closely (see Figure A1 in the supplementary material).
- Neither is it the result of under-representation of adults aged 40–59 in the complete database (which might lead to artificially higher mortality rates). To test this, we calculated age-specific mortality rates for the complete database (not reported) and compared them with those resulting from an 'extreme' hypothesis according to which all 149 individuals listed in the *Boccatico* whose age was unknown were attributed proportionally to the age groups 40–44 to 55–59. Even under this extreme hypothesis, mortality rates peaked around age 50–54 and then declined (consistent with the shape in Figure 8).

Discussion

Our findings are relevant for many ongoing debates about medieval and early modern plague. A first aspect to consider is the effects of sex and age on plague mortality. In a seminal paper based on the London 'Bills of Mortality' for the 1603 and 1625
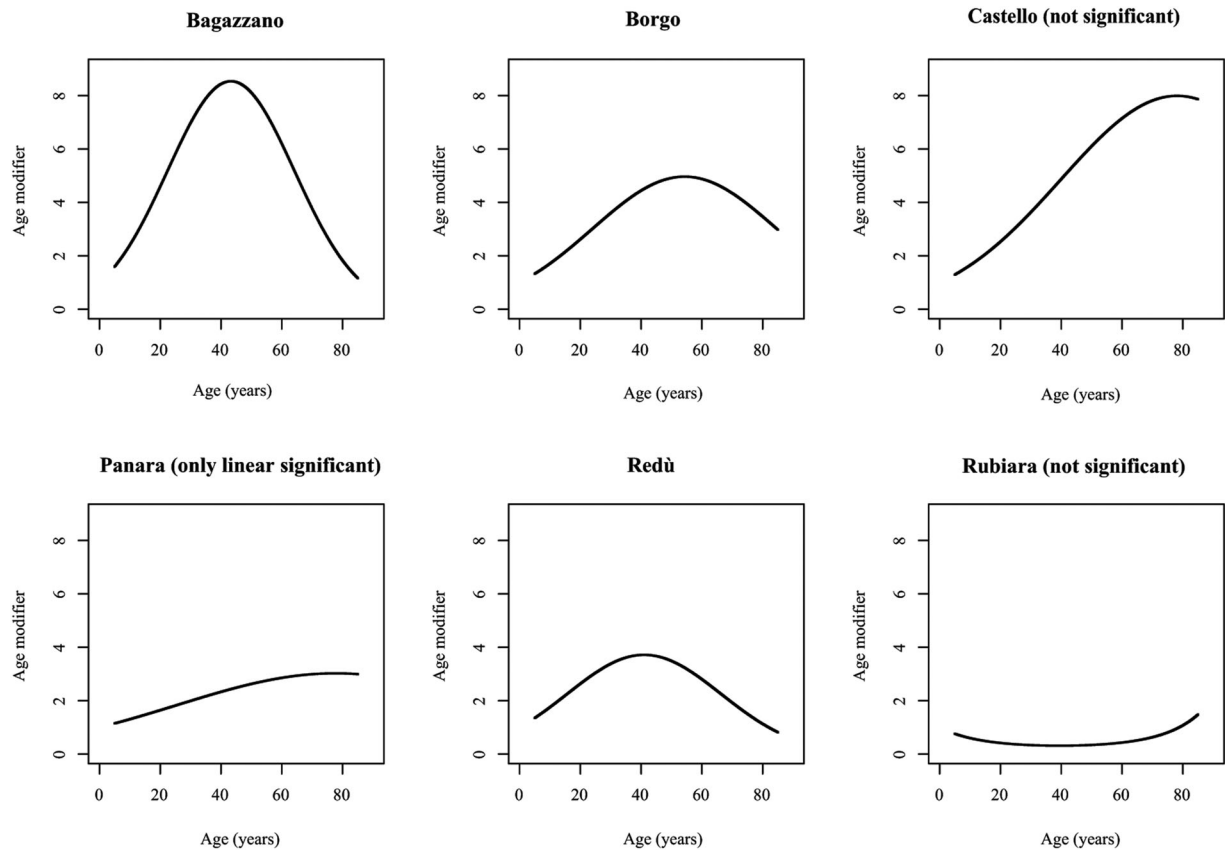
Figure 7   Age effects from fitting model (4) separately on each neighbourhood, Nonantola, 1630

Notes: The coefficients for the linear and the quadratic terms in age are both significant for Bagazzano, Borgo, and Redù. The linear term only is significant for Panara, and for Castello and Rubiara neither the linear nor the quadratic term is significant. Note that the original parameter estimates for $\beta_{\text{Age},1}$ and $\beta_{\text{Age},2}$ can be obtained by taking the natural logarithm of the corresponding hazard ratios (the parameters are also reported in section A5 of the supplementary material).

Source: Statistical estimates based on the same historical sources indicated for Table 1.
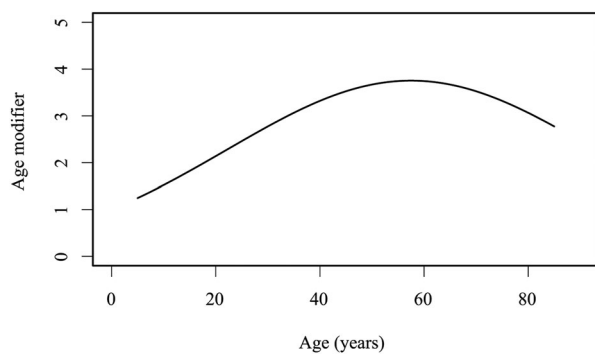

Figure 8   Overall age effect on the hazard function for burial/death, from model (4), Nonantola, 1630

Source: Statistical estimates based on the same historical sources indicated for Table 1.

plagues, Hollingsworth and Hollingsworth (1971, p. 144) concluded that it was 'very likely that the men were much more affected by plague than the women'. In contrast, a study of the six plagues occurring in Milan from 1452 to 1523 showed a marked prevalence of women among the victims, although this might have resulted from the concentration of immigrant women and poor widows in the overcrowded neighbourhoods that tended to be affected with particular severity by epidemics of any kind (Alfani and Cohn 2007). However, the clearly prevalent view in the literature is that plagues affected both sexes in about the same way. This was the case for the famous plagues of Colyton, Eyam, and Penrith in sixteenth- and seventeenth-century England (Bradley 1977; Schofield 1977; Scott and Duncan 2001; Whittles and Didelot 2016), as well as for Mantua (Italy) in 1575–76 (Del Panta 1980). Additionally, skeletal sources have been used to assess whether sex affected an individual's risk of death during the fourteenth-century Black Death, but no significant effects have been found (DeWitte 2009). Overall, our findings for Nonantola are consistent with such a conclusion, as in all models the sex variable was not found to affect mortality significantly. Other studies of the 1630 plague in Italy have

reached similar conclusions (Abrate 1972; Manfredini et al. 2002).

Regarding the effect of age on mortality, our study adds significant nuance to the conventional wisdom, as it includes a reconstruction of the complete age structure of the population exposed to the risk of dying of the plague—not just the age structure of the deceased, as is usually the case in the literature. This led to somewhat surprising results. In fact, an earlier study on Nonantola that analysed excess mortality by age group concluded that the plague affected the group aged 11–20 with the greatest intensity, followed at a distance by those aged 21–30 (Alfani and Cohn 2007 and section A2 of the supplementary material). Such findings are in agreement with the aforementioned studies of the 1630 plague that also analysed excess mortality (Abrate 1972; Manfredini et al. 2002), as well as with a study of the 1575–76 plague in Mantua (Belfanti 1981). Our study demonstrates that: (1) the aforementioned literature detected correctly the fact that the young (<10 years of age) were less affected by the plague; but that (2) the risk of death increased with age, peaking much later than earlier believed, that is, between ages 40 and 60 (Table 4 and Figure 8). As our data are more complete and precise than those used by other studies, it seems reasonable to conclude that we have detected a recurrent distortion in earlier analyses, determined by the kind of information they used.

The same problem could have occurred in studies relating to other areas of Europe, which also tend to suggest that plague struck the young preferentially from the fifteenth century onwards. Schofield (1977) tried to overcome the limitations implicit in analyses of excess mortality by comparing the actual number of plague victims to model populations obtained from stable population tables. He compared London (St Botolph), Eyam, and Colyton and found that the risk of dying of plague seemed to decline with age—contrary to what had been suggested by earlier works on excess mortality from plague, like Hollingsworth and Hollingsworth (1971). Schofield (1977), however, admitted that '[the] simple method we have employed … makes a number of assumptions that may or may not be true in reality. Our confidence in the validity of our conclusions would therefore be increased if we could show that a direct calculation of age-specific mortality during the epidemic produces essentially the same results' (p. 115). Our study provides exactly this kind of direct calculation of age-specific mortality, but unfortunately our results are in clear contrast with Schofield's (although not for the very

old ages). We know of only one other study where age-specific mortality rates were computed from real data: the work of Séguy et al. (2006) on Martigues in Provence (France), during the so-called 'Great Plague of Marseille' of 1720–21. These authors' results resemble our own quite closely as, overall, plague mortality rates increased with age. However, they did not report a decline in the risk of dying of plague at the older ages; this may be partly due to differences in the age classifications used (they clustered together all deaths occurring at ages 70+).

Additional evidence that the very old were less susceptible to dying of plague comes from plague treatises of the sixteenth and seventeenth centuries. Doctors suggested that old and 'dry' bodies were less susceptible to the disease—indeed, in sixteenth-century Venice, older people were elected as officers of the plague wards due to their presumed resistance to the disease (Alfani and Cohn 2007). Another hypothesis is that the reduced impact of plague mortality on the older population resulted from the selection process that had seen an exceptionally severe famine affecting the area 35–40 years earlier (Alfani 2011b, 2013b). This may have left only the stronger individuals, who constituted the older population in 1630. Such stronger individuals may also have been more resistant to plague. The age pattern of mortality may also have been influenced by age-specific behaviours. It seems plausible that middle-aged adults, especially women, were more likely to visit other households and help with nursing care, whereas the older and maybe also the younger adults stayed inside more, looking after infants and children, thus experiencing a reduced risk of contracting the disease outside their household. We found some empirical support for this, as the interaction term between sex (with female as the reference category) and the age group 40–59 in Table 4 shows a statistically significant greater risk of dying for women. As a final point, note that early modern plagues differed from the Black Death, which reportedly affected all age groups in a similar way (Biraben 1975; Scott and Duncan 2001; Alfani 2009).

A different perspective on the relationship between age, sex, and plague mortality is provided by studies of frailty based on skeletal sources, which have concluded that plague (in particular, the Black Death) was probably selective with respect to pre-existing health conditions, more easily killing people who had been in poor health since before the epidemic. These 'frail' people were identified by a variety of skeletal lesions (DeWitte

and Wood 2008) or by shorter-than-average stature (DeWitte and Hughes-Morey 2012). As the probability of being in poor health increases with age, this might help to explain the rising part of our age modifier curve (Figure 8). Note that this may be different when one looks at older subjects *who are still alive* at a given age, due to selection (see earlier). Frailty has also been studied with respect to sex, as in modern populations women are favoured in terms of morbidity and survival. The results of such studies are, however, unclear—plague excess mortality of frail men was found to be greater than that of frail women, but this could have resulted from: (1) a biologically determined greater resistance of women; or (2) a greater ability of plague to kill healthy women than healthy men (DeWitte 2010).

We could not test for frailty in a way similar to the aforementioned literature, due to the lack of information about pre-epidemic health status. However, women aged 17–35 were more likely than younger or older women to be pregnant at the time of the epidemic (mean age at first marriage was 17 for women in the years immediately preceding the plague), and earlier works have suggested that this might explain some of the reported excess mortality for young women compared with men (Alfani and Cohn 2007). Particularly high mortality among pregnant women has also been reported for some English plagues (Scott and Duncan 2001). We tested for a possible higher risk of dying of plague for young women aged 17–35 but the results were not significant. However, it is possible that the models did not have sufficient power to identify such interactions.

We also analysed the risk of death according to socio-economic status. The historical literature on Italian and European late medieval and early modern plagues seems to have clearly established that, in contrast to the Black Death and other early plague waves, the disease acquired a social character from the fifteenth century, striking the poor preferentially (Slack 1985; Alfani 2009, 2013b; Cohn 2010b). This has been confirmed by recent studies of specific plague outbreaks (Galanaud et al. 2015; Whittles and Didelot 2016). Doctors of the early modern period were of the same view (Cohn 2010a). According to Pietro Parisi, the Sicilian author of a famous plague tract in 1593, 'it is absolutely true that for poverty and privation, … the sticky illness, that is the contagion, which easily passes from one person to another, damages commoners more than the nobility' (cited from Alfani 2013b, p. 104). However, the Italian seventeenth-century epidemics are the notable exception to this, since the infection then spread so efficiently that the social and economic elite were also severely affected (Alfani 2013a). In Venice, 17 per cent of the members of the Great Council were killed by plague in 1630, while in Genoa, 40 per cent of the members of the Great and Low Councils died in 1656–57 (Pullan 1992). In this paper, we tested for a possible impact of high vs. low social status, but this factor proved non-significant in all models.

This provides confirmation of the exceptional character of the 1630 plague, which was so severe and widespread that it acquired the characteristics of a 'universal contagion', similar to the Black Death. Our findings also provide additional support for recent literature that suggests that the 1630 epidemic, unlike other seventeenth-century European plagues, had a deep and negative impact on the areas it affected and particularly on Italy—one possible reason being the particularly severe loss of human capital resulting from non-selectivity by socio-economic status (Alfani 2013a; Alfani and Percoco 2018).

However, socio-economic status could have affected the individual risk of death through the size of the household, as: (1) high-status individuals belonged to smaller households than peasants (as was usual in pre-industrial Europe); (2) no high-status household contained more than ten members; and (3) exceptionally large households (>12 members) experienced a 68 per cent increase in the hazard compared with small households (<4 members) (Table 3, model (3)). Another way of looking at this is to consider how the specific environment of residence may be connected to the individual risk of death. For London in 1560–1665, Cummins et al. (2016) reported a lower susceptibility to plague among the richest parishes and suggested that this was due to higher living standards. Additionally, plague outbreaks usually started in the poorest parts of the city, outside the walls. The same was found by Cohn and Alfani (2007) for Milan in 1523. For Dijon in 1400 and 1428, Galanaud et al. (2015) found that plague had more effect on the richest parts of the city, where most commercial and craft activities were concentrated (as they argue, 'In a city where garbage collection was left to the inhabitants, the proximity of food stocks, crop market, second-hand clothing dealers … and/or an open area used for slaughter … created conditions suitable for rats to pullulate and favored disease transmission by fleas' (p. 20)).

In Nonantola too, we found that the plague started in the richest part of the town, Borgo, where all communication routes converged. In this specific setting, the physical separation of the peasant hamlets from

the main settlement provided some relative protection to at least part of the poor population: as our stratified analysis suggests, areas affected later experienced a lower final plague intensity (Figure 6). This was presumably the consequence of a combination of factors: first, the arrival of winter, as we know empirically that medieval and early modern plagues tended to wane in the coldest months, possibly due to lower flea activity, reduced human interaction, or both (Cohn 2002; Alfani and Cohn 2007; Whittles and Didelot 2016); and second, the sparser settlement pattern of the hamlets compared with Borgo, and maybe even the actions taken to try to contain the spread of the epidemic from the town centre to the periphery. All in all, the distinctive layout of the town may have limited the possible advantage provided by better living standards to high-status individuals only—an advantage that in the 1630 plague was, however, probably very limited to begin with, as discussed earlier.

Our findings about the higher risk of death experienced by people living in larger households are also relevant to the long-lasting debate concerning the nature of the disease and its method of transmission. As is well known, the *Yersinia pestis* bacillus responsible for contemporary plague does not easily spread from human to human but spreads mostly through vectors (rats and rat fleas). This seems too complex a mechanism to account for the speed and effectiveness with which medieval and early modern plague spread, as has been suggested by historians (Cohn 2002, 2008; Bolton 2013; Alfani 2013b) and as is openly recognized also by the palaeo-biologists who identified DNA markers of *Yersinia pestis* in medieval plague cemeteries (Raoult et al. 2013). To solve the riddle, simpler ways of transmission have often been put forward, either through a different vector (a variety of human fleas instead of rat fleas; Biraben 1975; Del Panta 1980; Audoin-Rouzeau 2007) or borne by air, although the latter hypothesis is dubious as a general explanation, given that pneumonic plague has demonstrated an ability to spread effectively only under very specific environmental conditions (Cohn 2002, 2008; Raoult et al. 2013). Notice that both airborne transmission and transmission through human parasites can be considered 'human-to-human', as already argued by Biraben (1975) and recently by Whittles and Didelot (2016). While it is beyond the aims of this paper to provide an answer to such a complex and debated question as the means of plague transmission, our findings about household size and hazard of death nevertheless seem to be compatible

with the idea that plague may indeed have been able to spread directly from human to human.

In the case of a disease transmitted from human to human, the individual risk of contracting the disease would be directly related to household size, as the risk would be the cumulative result of: (1) the risk of all household members catching the disease outside the home; and (2) the risk of catching it from an infected household member. Indeed, under the somewhat simplistic assumption that each household member has the same probability of catching the disease outside the home, each additional household member increases the risk of the other household members becoming infected. Our findings clearly suggest that individuals living in large households (>7 members) experienced a risk of dying of plague that increased steeply with household size, compatible with the hypothesis of human-to-human transmission. This does not exclude the possibility that, during the same epidemic, the 'traditional' transmission route of rats to rat fleas to humans also occurred. A recent analysis of the Eyam plague of 1665–66 using stochastic models estimated that 73 per cent of infections may have been caused by human-to-human transmission, while just 27 per cent involved rodents (Whittles and Didelot 2016). Finally, an earlier study of Nonantola focused on the time interval between deaths within each household and found that most of them happened very close to each other: in 14 per cent of cases on the same day and in 40 per cent of cases within three days (see illustrative examples in section A3 of the supplementary material). Even greater time clustering of within-household deaths was reported for plague outbreaks in Milan from 1452 to 1523 (Cohn and Alfani 2007). This might signal transmission of pneumonic plague, whose incubation period is significantly shorter than that of bubonic plague (1–3 days instead of 2–6). However, the environmental conditions in which the 1630 plague occurred do not fit those considered typical for pneumonic plague. Consequently, we need to be cautious in making assumptions about which kind of human-to-human transmission might have occurred.

Conclusion

By using a new individual-level data set of a unique kind, this paper has developed the first detailed survival analysis of a large-scale mortality crisis caused by plague. We found that the risk of dying of plague was similar for the two sexes. The age effect followed a flat inverted U curve, with the risk of

dying peaking at ages 40–59. Socio-economic status did not affect survival significantly, although it could have played a role through household size. Members of large households (usually peasant and low-status) experienced a much higher risk of dying of plague. Stratification of the analysis by neighbourhood allowed us to highlight differences in the baseline hazard functions and to explore, to some degree, the spatial diffusion of the plague. This showed that the population residing in Borgo, where the epidemic started, was at greater risk than neighbourhoods infected later. Finally, some of our findings, particularly those related to the impact of the size of the household of residence on the individual risk of death, constitute evidence in support of the idea that medieval and early modern plague was able to spread from human to human.

The historical–demographic techniques that we used to build the data set and the bio-demographic methods that we developed could be replicated in other studies of European plagues. This would allow scholars to produce comparable and detailed information that could be vital for finding the hitherto elusive answers to crucial questions about the disease; in particular, *how* plague was able to cause the greatest mortality crises ever reported in human history.

Notes and acknowledgements

1 Guido Alfani and Marco Bonetti are at the Dondena Centre, Bocconi University, Italy. Please direct all correspondence to Guido Alfani, Department of Social and Political Sciences and Dondena Centre, Bocconi University, Via Roentgen 1, 20136 Milan, Italy; or by E-mail: guido.alfani@unibocconi.it

2 We thank Ms Diana Camerini, who has contributed to the statistical analyses. We also thank Monica Green, participants at the workshops *Peste, variole: microbiologie et modélisations* (University of Nice, 2012) and *The new history of European plague* (University of Reading, 2013), and seminar participants at Bocconi University for many useful comments.

3 Supplementary material for this article can be downloaded from the journal website as well as from http://didattica.unibocconi.eu/Alfani_database

References

Abrate, M. 1972. *Popolazione e peste del 1630 a Carmagnola* [*Population and plague of 1630 in Carmagnola*]. Turin: Biblioteca di Studi Piemontesi.

Alfani, G. 2009. Crisi demografiche, politiche di popolazione e mortalità differenziale (ca. 1400–1630) [Demographic crises, politics of population and differential mortality (ca. 1400–1630)], *Popolazione e Storia* 1/2009: 57–75.

Alfani, G. 2010a. The effects of plague on the distribution of property: Ivrea, Northern Italy 1630, *Population Studies* 64(1): 61–75.

Alfani, G. 2010b. Climate, population and famine in Northern Italy: general tendencies and Malthusian crisis, ca. 1450–1800, *Annales de Démographie Historique* 120: 23–53.

Alfani, G. 2011a. Le partecipanze: il caso di Nonantola, [The *partecipanze*: the case of Nonantola] in G. Alfani and R. Rao (eds), *La gestione delle risorse collettive. Italia settentrionale, secoli XII–XVIII*. Milan: Franco Angeli, pp. 48–62.

Alfani, G. 2011b. The famine of the 1590s in northern Italy. An analysis of the greatest "system shock" of sixteenth century, *Histoire & Mesure* 26(1): 17–50.

Alfani, G. 2013a. Plague in seventeenth century Europe and the decline of Italy: an epidemiological hypothesis, *European Review of Economic History* 17: 408–430.

Alfani, G. 2013b. *Calamities and the Economy in Renaissance Italy. The Grand Tour of the Horsemen of the Apocalypse*. Basingstoke: Palgrave.

Alfani, G. and S. K. Cohn. 2007. Nonantola 1630. Anatomia di una pestilenza e meccanismi del contagio (con riflessioni a partire dalle epidemie milanesi della prima Età moderna) [Nonantola 1630. Anatomy of a plague and the mechanism of contagion (with additional reflections from the epidemics of early modern Milan], *Popolazione e Storia* 2: 99–138.

Alfani, G. and C. Munno. 2012. Godparenthood and social networks in an Italian rural community. Nonantola, sixteenth-seventeenth centuries, in G. Alfani and V. Gourdon (eds), *Spiritual Kinship in Europe, 1500–1900*. London: Palgrave, pp. 96–123.

Alfani, G. and T. Murphy. 2017. Plague and lethal epidemics in the pre-industrial world, *The Journal of Economic History* 77(1): 314–343.

Alfani, G. and M. Percoco. 2018. Plague and long-term development: the lasting effects of the 1629–30 epidemic on the Italian cities, *The Economic History Review*, forthcoming.

Audoin-Rouzeau, F. 2007. *Les chemins de la peste* [*The paths of the plague*]. Paris: Tallandier.

Barbiera, I. and G. Dalla Zuanna. 2009. Population dynamics in Italy in the middle ages: new insights from archaeological findings, *Population and Development Review* 35(2): 367–389.

Belfanti, C. M. 1981. Mantova e la peste del 1575–76: problemi economici e condizionamenti demografici [Mantua and the plague of 1575–76: economic problems

and demographic constraints], *Studi Mantovani* 1: 57–68.

Benedictow, O. J. 2004. *The Black Death 1346–1353. The Complete History*. Woodbridge: Boydell.

Biraben, J. N. 1975. *Les hommes et la peste en France et dans les pays Européens et Méditerranées* [People and plague in France and the European and Mediterranean countries]. Paris: Mouton.

Bolton, J. 2013. Looking for Yersinia pestis: scientists, historians and the black death, in L. Clark and C. Rawcliffe (eds), *Society in an Age of Plague, The Fifteenth Century XII*. Woodbridge: Boydell.

Bos, K. I, V. J. Schuenemann, B. Golding, Hernán A. Burbano, Nicholas Waglechner, Brian K. Coombes, Joseph B. McPhee, Sharon N. DeWitte, Matthias Meyer, Sarah Schmedes, James Wood, David J. D. Earn, D. Ann Herring, Peter Bauer, Hendrik N. Poinar, and Johannes Krause. 2011. A draft genome of Yersinia pestis from victims of the black death, *Nature* 478: 506–510.

Bos, K. I., Alexander Herbig, Jason Sahl, Nicholas Waglechner, Mathieu Fourment, Stephen A Forrest, Jennifer Klunk, Verena J Schuenemann, Debi Poinar, Melanie Kuch, G Brian Golding, Olivier Dutour, Paul Keim, David M Wagner, Edward C Holmes, Johannes Krause, and Hendrik N Poinar. (2016). Eighteenth century Yersinia pestis genomes reveal the long-term persistence of an historical plague focus, *eLife*, 5, 17837. Available: https://elifesciences.org/articles/12994.

Bradley, L. 1977. The most famous of all English plagues: a detailed analysis of the plague at Eyam 1665–6, in *The Plague Reconsidered*. Cambridge: Hourdsprint, 63–94.

Clark, G. 2007. *A Farewell to Alms*. Princeton: Princeton University Press.

Cohn, S. K. 2002. *The Black Death Transformed*. London: Arnold.

Cohn, S. K. 2008. Epidemiology of the black death and successive waves of plague, in V. Nutton (ed.), *Pestilential Complexities: Understanding Medieval Plague*, Supplement Medical History 27. London: Wellcome Trust Centre for the History of Medicine at UCL, pp. 74–100.

Cohn, S. K. 2010a. *Cultures of Plague. Medical Thought at the End of the Renaissance*. Oxford: Oxford University Press.

Cohn, S. K. 2010b. The changing character of plague in Europe, 1348–1800, in S. Cavaciocchi (ed), *Le interazioni fra economia e ambiente biologico*. Florence: Florence University Press, pp. 33–56.

Cohn, S. K. and G. Alfani. 2007. Households and plague in early modern Italy, *Journal of Interdisciplinary History* 38: 177–205.

Cox, D. R. 1972. Regression models and life-tables, *J. of the Royal Statistical Society B* 34: 187–220.

Cox, D. R. 1975. Partial likelihood, *Biometrika* 62: 269–276.

Cummins, N., M. Kelly, and C. O'Grada. 2016. Living standards and plague in London, 1560–1665, *The Economic History Review* 69(1): 3–34.

Debbia, M. 1990. *Il bosco di Nonantola*. Storia medievale e moderna di una comunità della bassa modenese [The Wood of Nonantola. A Community of the Modenese Lowlands During the Medieval and Early Modern Period] Bologna: CLUEB.

Del Panta, L. 1980. *Le epidemie nella storia demografica italiana [Epidemics in Italian Demographic History]*. Turin: Loescher.

Del Panta, L. and R. Rettaroli. 1994. *Introduzione alla demografia storica [Introduction to Historical Demography]*. Bari: Laterza.

DeWitte, S. N. 2009. The effect of sex on risk of mortality during the black death in London, A.D. 1349–1350, *American Journal of Physical Anthropology* 139: 222–234.

DeWitte, S. N. 2010. Sex differences in frailty in medieval England, *American Journal of Physical Anthropology* 143: 285–297.

DeWitte, S. N. 2014. Mortality risk and survival in the aftermath of the medieval black death, *PLoS ONE* 9(5): e96513.

DeWitte, S. N. and G. Hughes-Morey. 2012. Stature and frailty during the black death: the effect of stature on risks of epidemic mortality in London, A.D. 1348–1350, *Journal of Archaeological Science* 39: 1412–1419.

DeWitte, S. N. and J. W. Wood. 2008. Selectivity of black death mortality with respect to preexisting health, *Proceedings of the National Academy of Sciences* 105: 1436–1441.

Galanaud, P., A. Galanaud, and P. Giraudoux. 2015. Historical epidemics cartography generated by spatial analysis: mapping the heterogeneity of three medieval 'plagues' in Dijon, *PLoS ONE* 10(12): e0143866.

Green, M. H. (ed). 2015. *Pandemic Disease in the Medieval World. Rethinking the Black Death*. Kalamazoo and Bradford: Arc Medieval Press.

Haensch, S., R. Bianucci, M. Signoli, M. Rajerison, M. Schultz, Sacha Kacki, Marco Vermunt, Darlene A. Weston, Derek Hurst, Mark Achtman, Elisabeth Carniel, Barbara Bramanti, and Nora J. Besansky. 2010. Distinct clones of Yersinia pestis caused the black death, *PLoS Pathogens* 6: 1–8.

Henderson, J. 2003. Historians and plagues in pre-industrial Italy over the longue durée, *History & Philosophy of the Life Sciences* 25: 481–499.

Hollingsworth, M. F. and T. H. Hollingsworth. 1971. Parish mortality rates by age and sex in the parish of St Botolph's without bishopsgate, London, 1603, *Population Studies* 25: 131–146.

Lin, D. Y. and L. J. Wei. 1989. The robust inference for the Cox proportional hazards model, *Journal of the American Statistical Association* 84: 1074–1078.

Livi Bacci, M. 2000. *The Population of Europe*. Oxford: Wiley-Blackwell.

Livi Bacci, M. 2007. *A Concise History of World Population*, 4th ed. Oxford: Blackwell.

Manfredini, M., S. De Iasio, and E. Lucchetti. 2002. The plague of 1630 in the territory of parma: outbreak and effects of a crisis, *International Journal of Anthropology* 17(1): 41–57.

Pullan, B. 1992. Plague and perceptions of the poor in early modern Italy, in T. Ranger and P. Slack (eds), *Epidemics and Ideas. Essays on the Historical Perception of Pestilence*. Cambridge: CUP, pp. 101–123.

Raoult, D., N. Mouffok, I. Bitam, R. Piarroux, and M. Drancourt. 2013. Plague: history and contemporary analysis, *Journal of Infection* 66(1): 18–26.

Schofield, R. 1977. An anatomy of an epidemic: colyton, November 1645 to November 1646, in *The Plague Reconsidered*. Cambridge: Hourdsprint, pp. 95–126.

Scott, S. and C. Duncan. 2001. *Biology of Plagues: Evidence From Historical Populations*. Cambridge: CUP.

Séguy, I. and L. Buchet. 2013. *Handbook of Paleodemography*. Heidelberg: Springer.

Séguy, I., S. Pennec, S. Tzortzis, and M. Signoli. 2006. Modélisation de l'impact de la peste à travers l'exemple de Martigues (Bouches-du-Rhône) [Modelling the impact of plague through the example of Martigues (Bouches-du-Rhône)], in L. Buchet, C. Dauphin, and I. Séguy (eds), *La paléodémographie. Mémoire d'os, mémoire d'hommes*. Antibes: APDCA, pp. 323–331.

Seifert, L., Ingrid Wiechmann, Michaela Harbeck, Astrid Thomas, Gisela Grupe, Michaela Projahn, Holger C. Scholz, Julia M. Riehm, and Xue-jie Yu. 2016. Genotyping Yersinia pestis in historical plague: evidence for long-term persistence of Y. pestis in Europe from the 14th to the 17th century, *PLoS ONE* 11(1): e0145194.

Slack, P. 1985. *The Impact of Plague in Tudor and Stuart England*. London: Routledge.

Sperandini, G. M. 1997. La peste del 1630 a Nonantola [The 1630 plague in Nonantola], in *Nonantola e la Bassa Modenese*, special issue of *Quaderni Della Bassa Modenese* 11(2): 231–247.

Vaupel, J. W., K. G. Manton, and E. Stallard. 1979. The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* 16: 439–454.

Whittles, L. K. and X. Didelot. 2016. Epidemiological analysis of the Eyam plague outbreak of 1665–1666, *Proceedings of the Royal Society B: Biological Sciences* 283: 20160618.