UNIVERSITÀ DEGLI STUDI DI MILANO

CORSO DI DOTTORATO IN SCIENZE PER LA SANITA' PUBBLICA.
36ESIMO CICLO

DIPARTIMENTO DI SCIENZE CLINICHE E DI COMUNITA'

TESI DI DOTTORATO DI RICERCA:
ANALYSIS OF THE DEATH EVENTS, ACCORDING TO THE HISTORICAL, DEMOGRAPHICAL AND TERRITORIAL CONTEXT OF MILAN PEOPLE. DATA FROM THE MILANO SFORZA REGISTERS OF THE DEAD (1452-1801)

MED/01

PhD Candidate: ESTER LUCONI

R12896

Tutor: Prof. ELIA BIGANZOLI

PhD Coordinator: Prof. CARLO VITANTONIO BATTISTA LA VECCHIA

Academic Year: 2022-2023

ABSTRACT (ENG): The Milano Sforza Registers of the Dead, a corpus of 366 volumes, is one of the oldest mortality registers of any large European city. Established in 1450 by Duke Francesco Sforza, they are an example of disease monitoring and prevention. The information in them reflects death events at the individual levels and includes both socio-demographical context and clinical data (causes of death and description of the illness).

The information network that compiled the registration started with the parish elder; it included health officials and ended with the notary at the Health Office who transcribed the information.

REDCap (Research Electronic Data Capture), a web-based application developed by Vanderbilt University, allowed the creation of a database to store the data of the registers and ensured their integrity. The data of 1480, a year without plague and war, was taken as an example to show the database structure for the registers. The database is essential to perform statistical analysis based on the data.

The registrations of 1480 were used to study the event distribution for different age classes and gender. They were used to study the distribution of the events in different parts of the city and the time trend. Detailed descriptions of the death in foundlings, death for childbirth, violent causes of death, and hospital events are provided.

Two plague periods were then considered: 1630 and 1485. Data about 1630 were used to perform the spatiotemporal reconstruction of a famous epidemic, known all over the world thanks to the work of Alessandro Manzoni. Data about 1485 were used to analyze the association of plague symptoms, differences between life and death for plague, and to study the time distribution between symptoms and death. Finally, Ferrario's time series, which was compiled in 1840s, and aggregated information about the number of deaths in each month, was used to find the outliers that can be connected to extraordinary events and to subdivide the series into periods with similar characteristics. With this aim, a R package for the ecological time series was used.

ABSTRACT (ITA): I Registri dei Morti di Milano degli Sforza, un corpus di 366 volumi, è uno dei più antichi registri di mortalità di qualsiasi grande città europea. Istituiti nel 1450 dal duca Francesco Sforza, sono un esempio di monitoraggio e prevenzione delle malattie. Le informazioni in essi contenute riflettono gli eventi di morte a livello individuale e includono sia il contesto socio-demografico che dati clinici (cause di morte e descrizione della malattia).

La rete informativa che ha redatto la registrazione è iniziata con l'anziano della parrocchia; comprendeva funzionari sanitari e si concludeva con il notaio dell'Ufficio di Sanità che trascriveva le informazioni.

REDCap (Research Electronic Data Capture), un'applicazione web sviluppata dalla Vanderbilt University, ha consentito la creazione di un database in cui archiviare i dati dei registri e ne ha garantito l'integrità. I dati del 1480, anno senza pestilenze e guerre, sono stati presi come esempio per mostrare la struttura del database dei registri. Il database è essenziale per eseguire analisi statistiche basate sui dati.

Le registrazioni di 1480 sono state utilizzate per studiare la distribuzione degli eventi per diverse classi di età e sesso. Sono state inoltre impiegate per studiare la distribuzione degli eventi nelle diverse parti della città e l'andamento temporale. Vengono fornite descrizioni dettagliate della morte dei trovatelli, della morte per parto, delle cause violente di morte e degli eventi ospedalieri.

Sono stati quindi considerati due periodi di peste: 1630 e 1485. I dati relativi al 1630 sono stati utilizzati per effettuare la ricostruzione spaziotemporale di una famosa epidemia, conosciuta in tutto il mondo grazie all'opera di Alessandro Manzoni. I dati relativi al 1485 furono utilizzati per analizzare l'associazione dei sintomi della peste, e la loro relazione con l’outcome (la vita o la morte per la peste) e per studiare la distribuzione temporale tra i sintomi e la morte. Infine, le serie temporali di Ferrario, compilate nel 1840, e le informazioni aggregate sul numero di decessi in ogni mese, sono state

utilizzate per trovare valori anomali che possono essere collegati a eventi straordinari e per suddividere la serie in periodi con caratteristiche simili. A questo scopo è stato utilizzato un pacchetto R per le serie temporali ecologiche**.**

To all those who have walked the soil of Milan before us, so that they are never forgotten

Summary

1 INTRODUCTION

1.1 Plague and the analysis of death events for Public health surveillance

The systematic collection of health-related data and their analysis and interpretation are crucial to planning, implementing, and evaluating public health practice (*Introduction to Public Health Surveillance|Public Health 101 Series|CDC*, 2018) Data about the causes of death are fundamentals in this context (Naghavi et al., 2010) at various levels (Byass, 2007):

- Medical and legal practitioners need individual causes to follow up on the consequences of individual deaths.

-Institutional managers and clinical auditors need patterns of deaths within institutions and health care systems for monitoring trends over time and within departments.

-Causes of death relating to specific populations and subgroups help interpret particular situations regarding mortality patterns.

-The top-ranking causes of death are essential for the local public health managers.

-The cause-specific mortality estimates are significant at national and international levels.

Historical death registrations, called Books of Deaths, *Necrologi*, or *Liber Mortuorum*, were made principally to evaluate the possible presence of plague in the city. In addition, Milanese death registers provide more excellent clinical and epidemiological information than similar records elsewhere in Europe but have not been comprehensively studied. This study illustrates how the Milanese *Necrologi* contributed new evidence about epidemics in the premodern era. They provide valuable information about a lethal infectious disease still active today.

Plague occurs cyclically in certain territories (for example, today in Madagascar, cases of bubonic plague happen nearly every year between September and April). It is endemic in the Democratic Republic of Congo, Madagascar, and Peru (WHO, 2022) Still, in history, there were three pandemics of this disease with a very high number of deaths and significant impacts in social, cultural, and economic terms (in addition to these, there have been many other plague epidemics, albeit with a minor impact):

- Justinian Plague of 541-544
- Black Death in 1347-1352
- Third Pandemic of 1894

The Justinian Plague spread from Ethiopia (Africa) to Pelusium (Egypt), Alexandria, Gaza, Jerusalem, and Constantinople (now Istanbul), where it probably caused 5000 deaths per day during the spring. Subsequently, it spread to Italy, France, the Rhine Valley, Iberia, Denmark, Ireland, Africa, the Middle East, and Asia Minor. It was estimated that 100 million people died.

The initial pandemic wave, followed by an eight-year famine, impacted society and the economy and laid the groundwork for the nations of medieval Europe as a consequence of the end of Roman rule.

The second multi-century pandemic began with the Black Death, spreading from Central Eurasia. Again, the mortality was different from city to city. Still, Giovanni Boccaccio (an Italian writer and poet), who also testified the symptoms in his *Decameron*, described that up to half of the inhabitants of his city, Florence, died; a similar outcome was probably present in Venice and Paris. From 1347 to 1350, 50 million people died.

The Black Death created a social leveling. Significant mortality led to the affection of primary and secondary production. This fact enabled the laborers to ask for higher wages due to the lack of competition, creating high costs for the landowning aristocrats; it was possible that villains could acquire properties.

The third pandemic spread from Yunnan (China) in the late 1800s and ended only in 1959; it caused 15 million deaths (Frith, 2012).

So, it is not surprising that plague was considered a great danger to the population. Many novel measures were taken to control the disease, first to prevent infected people and goods from entering the city by establishing the sanitary cordon, and, once the plague entered a town, with the separation between healthy and infected persons (makeshift camps and quarantine) (Tognotti, 2013). The analysis of the causes of death was of primary importance to establish the presence of plague in the city.

The best known, and perhaps the most studied, example of a book of the dead is the Bills of Mortality of London: from the plague of 1518, all parish priests had to report the number of plague deaths in their parishes each week. By 1560, older women called searchers were tasked to inspect bodies to determine the causes of death (Heitman, 2020). In the beginning, the assignment of those women was to distinguish if the person died from plague (but later, they were called to report other causes of death) or from other diseases. A physician was called when the diagnosis was disputed. In addition, they had to live together and carry a white wand to allow others to recognize and avoid them (Forbes, 1974).

After their inspections, older women reported to the parish clerk, who had the task of making counts that would be sent to the clerk of his Company. Finally, in the 1600s, a city-wide report based on the data was made for the Lord Mayor and aldermen and was sent to the Privy Council the next day. In 1592, another plague year, those reports became public because large-scale broadsheet copies were posted in public areas (Heitman, 2020).

Classically, Bills of Mortality were composed of a single sheet that presented on one side the mortality for each parish in London and on the other the various causes of death. For example, the Bills of Mortality from 31 October to 7 November 1665 reported that 1414 people died of “Plague”, 61 of “Consumption” and one, whose cause of death is not known, “Found dead in the Fields at St. Mary Islington”. In 1629, two sheets of weekly bills were printed: one with the Diseases and Casualties on and the Deaths by Parish on the other. The first one allowed the Bills to become a set of commercial news services, public health measures, and scientific publications (Boyce, 2020).

The Bills of Mortality were available because they were posted in public areas.

Beyond the Bills of Mortality of London, there were other registries in Europe; for example, in Barcelona, the incredibly high mortality of plague in the Mediterranean area in the fifteenth and sixteenth centuries led to the recording of the deaths. In addition to evaluating the potential presence of plague in the city, the registration of the death allowed the correct assessment of the extension of the phenomenon in the town (there was the risk of being accused of misfortune with the decline of the trade in the Mediterranean). This goal has been achieved thanks to the letter carrier (*correu*), who takes information about the number of deaths and baptisms in the city's parishes (Smith, 1936).

In some Italian cases, the registration of the deaths also had other goals apart from monitoring the presence of plague; for example, in Florence, recordings helped provide adequate provision for the city. There were two registration series: one covering the period 1385-1778 produced by the Board of the Grascia (agency for the provisioning of the city) and the latter covering the period 1450-1808 produced by the Guild of Physicians and Apothecaries. The two registrations were based on the information that the city's gravediggers reported, who also compiled a registration for the Health Magistracy (even if the registers were not preserved). Unfortunately, at least until the seventeenth century, children's recorded deaths were rarely reported (based on the discrepancy between the number of deaths and the expected for a pre-industrial period). Also, the deceased's age was usually absent, but from this period, the record of the deceased's profession appears with the names of the dead (C. M. Cipolla, 1978).

In Venice, the collection of information about death in the form of registers started in 1504, after a plague pestilence, and at least it did not end until 1805; this *Necrologi* was made before parish registers (which began in 1550). Information about the details of the identity of the deceased, the illness, the treatment, and the funerary arrangement was collected by parish priests and transmitted daily to the Health Magistrate. Each record corresponds to a subject. Each register generally corresponds to a single year, and each book starts in March and ends in February. Moreover, those registers included drawings accompanying the recorded cause of death, used to highlight sudden deaths that could be attributed to the plague. Thanks to this system, in 1630, a cluster of deaths in the parish of Sant' Agnese was rapidly identified, leading to the immediate activation of prevention protocols. Other causes of death accompanied by images include death related to water (such as drowning and falling off a boat), murders, fallings, and death of centenarians. The registers also allowed for the establishment of the demographical trend of the population, with each month beginning on a new page (Bamji, 2019).

1.2 The Libri Mortuorum of Milan

In the mid-XV century, the Duke of Milan, Francesco I Sforza (1401-1466), established a unique daily registration method for socio-demographical data and causes of death in the town of Milano, the *Libri Mortuorum*, or the Milano Sforza Registers (MiSfoRe). Like the registration in other cities and periods, they were made in response to plague epidemics: they were conceived to track the spread of infectious diseases, manage contagion dynamics, and distinguish plague from other illnesses. They were organized in one person for each entry, which made possible the reconstruction of the chain of contagion. The physician could use the information on the causes of death to help identify new cases of plague (Vaglienti, 2020). They contain not only the residents of the city but also those who were in the city temporarily (migrants, pilgrims, soldiers, refugees). In addition, until the first periods of the sixteenth century, they reported subjects died in city hospitals (Carmichael, 2016).

The *Libri Mortuorum* of Milan contains very detailed information about the socio-demographic context of mortality, and the causes of death of approximately 1.5 million individuals in Milan. Registers covered the period from 1450 to 1801, all now preserved at the *Archivio di Stato di Milano* (Milano State Archive). To our knowledge, it is one of the first European registrations with this information at the individual level. In addition, unlike other registers, the causes of death were ascertainment by a physician, and it reports the date of the death (not the date of the burial) (S. K. Cohn & Alfani, 2007).

The Registers of the 15th century were written in Latin [1]. Each record reported (Vaglienti, 2013):

(1) The date of the event.

(2) The location of the events within the city or suburbs. The record specifies in which of the six *Porte* the death happened. The city of Milan was subdivided into six parts according to the gate's name in the Medieval walls in the corresponding area. Those areas are also called *Sestieri* (with the

---

[1] Registers were written in Latin before 1774 (Zanetti, 1976)

form of irregular triangles and with the zone of the *Palazzo della Regione* as the vertex). From north and clockwise, there were *Nuova, Orientale, Romana, Ticinese, Vercellina,* and *Comasina*.

(3) The parish where the deceased person lived, died, or would be buried. Parishes were the center of social life for the population.

(4) A specification of where the death occurred (the street, a hospital, an inn, a prison).

(5) The deceased person's name, surname, or nickname. For children, adolescents, and unmarried women, declarants reported (with few exceptions) the father's name (specifying if he was still alive or deceased) or, if the father was unknown, of their mother. Foundlings were identified as *hospitalis filius* or *filia* (son or daughter of the hospital) or received a surname, most frequently *da Milano* or *Colombo*, with modifications in the frequency of their use and variants over the centuries. Their husbands' names and surnames often accompany the names of married women. Gender is not mentioned, but with a few exceptions, it is easily inferred from the person's name or adjectives used.

(6) The social and professional status of the deceased (of the father in the case of children or the husband in the case of adult women), including titles indicating nobility, ecclesiastical or professional status, or titles deriving from military or administrative functions. The occupation (servant, merchant, craftsman or worker, etc.), the condition of poverty, or the status of a prison inmate is also declared. [2]

(7) The age at death, expressed in years for adults and in months, days, or even hours for infants. However, the ages of older adults were often rounded up to the nearest decade (Zanetti, 1976).

(8) The cause of death. Reported as plague or not-plague in the earliest registers, the causes of death became afterward described with care, in some cases with a level of detail such as that provided in the following sample, taken from the register of 1459: “Long-suffering of quartan fever, the night before he had a paroxysm, aggravated by indigestion that caused heart palpitations and blurred intellect, then evolved into syncope and apoplexy that led to the death".

(9) The information about the duration of the illness leading to death was usually collected because it was an essential indicator of the presence of plague, which was believed to lead to death within

---

[2] Social status and employment data are largely incomplete and reported only in some cases, especially for ordinary people. The registers most accurately report this kind of data are from 1485 to 1495. (Zanetti,1976)

three to four days. Moreover, the signs and symptoms, with the time of their appearance, are collected in the case of plague.

(10) The name of the official in charge of inspecting the bodies and determining the causes of death, which were then transcribed into the register along with all other information provided. This official could be a 'collegiate' doctor (indicated in this case with personal name and title) or a health official appointed to this specific duty; less frequently, the elder of the parish or, in exceptional circumstances, the gravedigger.

This excellent source is a product of health reforms that took place from the Duke of Milan Gian Galeazzo Visconti (1351-1402): having to deal with the resurgence of the plague epidemic, around 1401, he established the figure of "health police" (*offitium perquirendi et exequendi expedientia circa conservationem sanitatis civitatis our Mediolani*), and he designated his official, Giovanni Rosselli, for this function. In addition, he charged the *Anziani delle Parrocchie* (Parish Elders) to report daily new suspected cases of plague (Carmichael, 2016). They were laypersons who had various tasks in the administrative management of the parish (seen as a territorial unit rather than a building). In the last year of the seventeenth century, they had to know all their parishioners to indicate people who needed help (frail and poor people) and foreigners. In addition, they had to ensure the territory's security under their jurisdiction by reporting potential criminals and supervising activities that produce polluting waste or sewage (Antonielli, 2008). Moreover, from 1773, they were responsible for the body until the burial (Vaglienti, 2013).

From 1399-1400, each surgeon (*medicus, ciroychus),* barber (*barberius*), and apothecary (*herborarius)* of the city must make a list of the pathologies found in the patients and annotate the course of the disease (Albini, 1982). In their turn, the notaries were assigned to proceed with the registration operations of the death (previously communicated by the Parish Elder) (C. Cipolla, 1976). At the time of Gian Galeazzo Visconti, the measure to avoid plague diffusion was the separation between the healthy and sick (or potentially sick), so ascertaining the causes of death allowed this. Finally, the *Officiali delle Bollette* had the task of isolating infected areas and had to control the movement of man and goods (Zanoboni, 2020)

Under Filippo Maria Visconti (1392-1447), the once temporary position of the ducal health commissioner became regularized. In addition, starting from 1438, health professionals (doctors, surgeons, barbers, and apothecaries) had to communicate the names of the sick individuals they were treating. Furthermore, the elder was required to notify all sick and dead people within a maximum of five days from the event (C. Cipolla, 1976).

Francesco I Sforza implemented the Visconti reforms in the Duchy of Milan. At the time, the city counted 110.000 inhabitants (Chandler, 1987) and was an essential connection between northern Europe with Italy and the Mediterranean Basin (Tucci, 2011). *The Ufficio di Sanità*, constituted on 10 March 1450, aimed to maintain public health in the entire dominion, primarily monitoring the diffusion of the epidemics. It was composed of a physician who specialized in epidemic illness, a physician-surgeon (called *Catelano*), a barber (for minor surgical interventions), a carter, two buryers, a notary, two ridings, three servants, a messenger to bring the bulletins about the deaths. Commissioners and two deputies were at the top of the structure. All these professions were employees of the Duke (F. Vaglienti, 2020).

The procedures of the *Ufficio di Sanità* were not uniform, so a better organization was needed (Giovanelli Onida & Marinai, 2001). With the *Magistrato di Sanità* (Health Magistrate), Francesco II Sforza revised the office norms. Francesco Lampugnani listed some people involved in this Health Magistrate: the *Scrivano dei Morti* (scribe of the dead) to transcribe the information about who died during the day in the death registers; it was necessary to have the cause of death assessed by a physician of the *Ufficio del medico di chirurgia*. The *Ufficio del medico di chirurgia* (the office of the surgeon physician) took care of the judgment of the physician about the causes of death, and in case of uncertainty, an inspection of the body will be performed; in their task, they can be helped by the Parish Elders. Furthermore, the *Commissari di Sanità* (Health Commissioners) had the job of maintaining a certain level of sanitary conditions regarding the environment; in times of plague, they had to reconstruct the chain of contagion, close the infected houses, and get the sick out of the city, towards the appropriate places. There were also commissioners tasked with preserving and managing plague outside Milan in the duchy's territory.[3]

The health magistrate remained organized in those terms until 1749, when Empress Maria Theresa led to the reduction of the number of commissioners to one. Finally, in 1786, the Austrian government decreed the abolition of the Magistrate (Visconti, 1911).

---

[3] Lampugnani Francesco. 1578. La Erettione del Magistrato della sanità di Milano, insieme con le ordinationi à questo ufficio appartenenti : fatta per lo Illustrissimo signore Francesco secondo Sforza, duca di Milano, l'anno M. D XXXIIII. Et ora tradotta in lingua volgare. Alla quale si sono aggiunte due lettere & un'ordinatione , mandate in varii tempi da tre Illustriss. duchi di Milano à i deputati alla sanità di Pavia, intorno ad alcune occorrenze del loro ufficio, & all'autorità à loro data. Seguono poi ordinatamente tutte le provigioni & gride, fatte & publicate da gli Illustri signori conservatori della sanità di essa città di Pavia per gli sospetti & estirpatione del contagio ultimamente ivi occorso l'anno M. D. LXXVII.
Pavia: Appresso Girolamo Bartoli)

1.3 Literature review

Compared to other registers, such as the Bills of Mortality of London, the Registers of Milan were never publicly available, a decision made by governments that succeeded each other in ruling the city since the beginning of the 15th century.

One of the first scholars to use this source was Giuseppe Ferrario (1802-1870): in his "*Statistica medica di Milano dal secolo 15 fino ai nostri giorni, Volume 2*" (Ferrario, 1840), he used the Registers of Milan, to make tables that reported the number of deaths for each month for each year. In his count, he put hospital deaths in different tables. The religious (friars and nuns) were not written regularly for many years covered by the registers.

In his work in 1976 (Zanetti, 1976), Zanetti discussed the use and the limitations of the registers for demographical analysis. In addition, he compared the causes of death and the terminology used in three years (1503, 1616, and 1783), showing that the differences reflected changes in medical knowledge. For those three periods, he reported the ratio of diagnostic formulation and the number of deaths, the recurrent terms of the causes of death and the number of deaths about some groups of diagnosis, and the ratio of the two sexes concerning different causes of death and age.

Another work on the usefulness of the registers to assess the state of medicine in different periods was centered in the fifteenth century (Carmichael, 2019): the Galenic pathological processes were the center of the medical ideology; illness was seen as a sign of humor imbalance. In fact, in the registry are some descriptions of urinalysis.

Two elements were essential to diagnose plague: buboes (also seen as aphostema or abscess) and fever. The first one represents the localization and consolidation of putrefying humors, and its position is strictly associated with the outcome of the patients. However, buboes could also express the absorption of poisons in the corrupted air (among the causes of plague for Galenic practitioners). But buboes were not sufficient for a plague diagnosis; it was necessary to have a detailed clinical history of the patients and/or other signs. In the registers of 1468, there is a case of a 25-year-old man with a bubonic swelling in his right groin, in which plague was excluded. Plague fever differs from other types of fever, signs of humor imbalance, but the first can also be a sign of inspired poisons.

One study on the registers was about the elite physicians and community practitioners who certified cause of death (Bertoglio, 2021) in the year 1478 (chosen for the completeness and because it was a year free from the plague, but in which there was an epidemic year): the majority of the reports belonged to Giovanni Catellano, Dionigi da Norimberga, and Giovanni de Casetis. The first one controlled and registering the deaths, visited the sick, and served the *Ospedale Maggiore* of Milan. The second one cared for the dying who had no previous medical visits and dealt with suspected cases (he was the plague doctor). The third one was in charge of hospitalized people. In a recent article by Ann Carmichael, there is a significant emphasis on the underlying forensic tradition embedded in these registers (Carmichael, 2022).

Further information in the registry that has been a subject of study was the name of the deceased (Carmichael, 2016): it allowed social identification in civil records because it is linked with the subject's existence and personality. The proof of the existence was essential because there was the fear of anonymity during the burial. The study considered the period between 1503-1505 (because of the complete registration): of 11.156 deaths, only 86 had no name, but they were nevertheless characterized in some way (for example, "a beggar women"), 18 of those 86 there were adults with unknown age. Four hundred sixty-seven subjects on 11.156 were identified only by their first name, but 95.7% were children left at San Celso hospital.

Cohn and Alfani (S. K. Cohn & Alfani, 2007) took especially into consideration the registry of 1523, a plague year, that was rich in information about the place of death: not only the parish (and the *Sestiere*) but the house. So, they analyzed the household in addition to the duration of the symptoms until death and found that 25% of household victims died on the same day.

Historical registers also have hidden information not contained in the text. A work (D'Amato et al., 2018) about the registers of the plague of 1630 concerned 11 pages on which physicochemical analysis of the EVA diskettes was performed. Proteins of *Yersinia* were found, but also the persistent presence of proteins from *Ovis aries* and *Bos taurus*: the first animals were present in the *Lazzaretto* to feed the newborns with milk, and the second were related to the material of the pages, especially the glue. Proteins of plants allowed the authors to have some information about the diet in the place.

2 CREATION OF A REDCap DATABASE TO COLLECT INFORMATION OF THE MILANO SFORZA REGISTERS OF DEAD

“There are so many stories that are locked away in historical data” (Andrew Trant, an ecologist at the University of Waterloo in Canada).

Historical data are helpful for many researchers, for example, in studying climate change (ecological and climatological information) and the epidemiological field. The digitization of the resources makes this easier (Kwok, 2017).

Digitization can help preserve information, but creating the database can put the information in a more suitable format, in particular, to analyze the data. Examples of the possibilities through the utilization of a database are given by the work of Ulf Büntgen et al. (Būntgen et al., 2012). They constructed a database based on Biraben’s list, including data on 6929 European plague outbreaks between 1347 and 1900 AD. The authors declare that those data are helpful to the study of spatiotemporal patterns and dynamics of historical plague outbreaks. The dataset is freely available on a website (https://www.envidat.ch/#/metadata/digitizing-historical-plague).

Also, databases based on the information of one city can have significant implications for scholars: the “Death by Numbers” project involved the Bills of Mortality of London, and the aim is to transcribe and publish the information in the registers in a format suitable for computational analysis. This database can be helpful, for example, to study the transmission of the disease (change of the infection pattern over time, how the natural and human infrastructure had an impact on this, the implication of an event such as the “Great Fire”). The database will be freely available online (https://deathbynumbers.org) (Otis, 2017).

The historical death registry had interdisciplinary interests for scholars of different disciplines (historians, medical historians, anthropologists, clinicians, and epidemiologists), so creating the database about Milano Sforza’s registry is helpful to make the information accessible to them.

The database is a project of *Università degli Studi di Milano* (University of Milan) in collaboration with the Direction of the State Archives of Milan. The project is not funded. It will contain all the data of the registry that will be transcribed by an expert paleographer and translated from Latin.

Data are entered by a manual transcription of the information contained in the registers, including, when available, information about household organization, professional positions, and the social status of the deceased person. In addition to the information available in the registry, additional data from external sources will be included: the coordinate's location of the parishes, external links toward literature and archival sources, and a grouping of the causes of death (for analysis purposes). The latter will be added after an accurate interpretation of the terms used in the registers and a careful study of their meaning based on the consultation of medical treatises of the 14th to 16th centuries.

REDCap (Research Electronic Data Capture) (Harris et al., 2009, 2019) is used to build this database; it is a Web-based application developed by Vanderbilt University in 2004 primarily for clinical research because it allows for capturing data and creating databases and projects. In addition, it is free for universities. The pros of using REDCap are:

- long-term reduction of the research cost
- knowledge of a programming language is not required
- the possibility of utilization on both smart devices and desktop computers (in particular, for smart devices, an application with a QR code for the project enables the data collection or the view of the project)
- the software implemented rapid data entry, review, and descriptive statistics.

In addition, REDCap is compliant with FISMA (Federal Information Security Management Act), GDPR (General Data Protection Regulation), HIPAA (Health Insurance Portability and Accountability Act), and 21 CFR Part 11 (Part 11 of Title 21 of the Code of Federal Regulation). Therefore, it is possible to indicate variables that contain sensitive data, and it is possible to exclude them from the data export not to allow the personal identity of a subject.

The cons of using REDCap are the necessity of a computer technician to perform maintenance and the API configuration (an interface that allows external applications to connect to REDCap remotely), the fact that the interface is not easily intuitive for new users (Garcia & Abrahão, 2021; Patridge & Bardyn, 2018).

The structure of the database in a period free from plague is described in Figures 1 and 2.

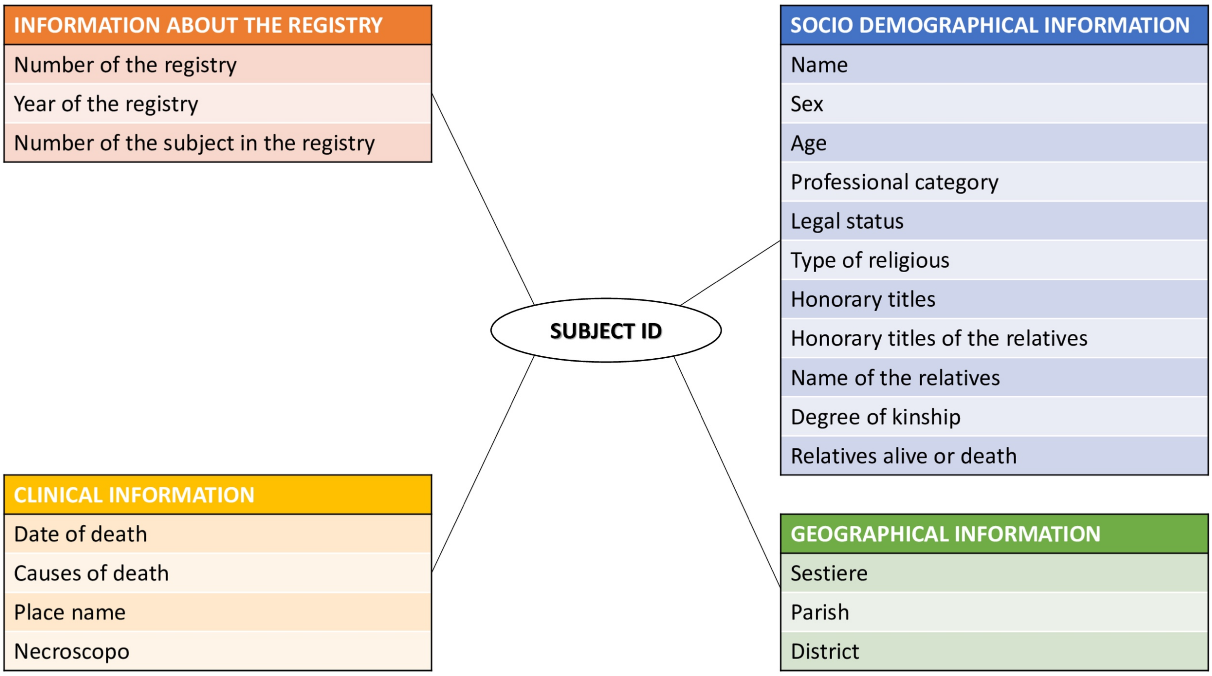


Figure 1: database structure with the four-area linked to the subject's ID and the variables in each area.

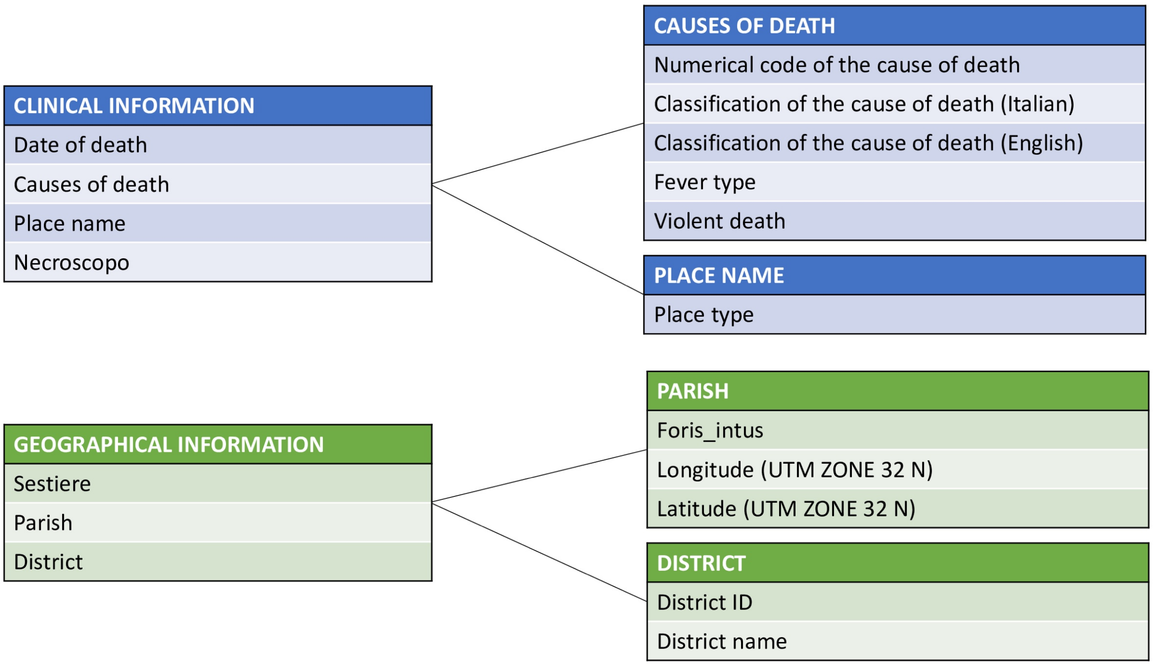


Figure 2: variables in strict relationship with each other

Here is presented the data dictionary of the project, which represents the structure of the database in a concise form. Each row corresponds to a field in the data entry (Queen Mary University of London, n.d.) (University of Nebraska Medical Center, 2019):

-Variable/Field name: the name of the variable

-Form Name: grouping of variables within the database (it can be seen that in this database, there is none)

- Section Header is used to visually separate items within a form, primarily to aid data entry (in this database, there are:

  -General information about the registry

  -Socio Demographical information about the subject

  -Geographical information about the *Sestiere* and the Parish

  -Clinical information about the death event)

-Field type: determines what types of responses are allowed and how they will be displayed.

- Text: containing a short text, a number, or a date. It is possible to validate the entry, for example, by specifying the minimum and the maximum values allowed.
- Notes: it contained a long text.
- Radio: the possibility of answers will be displayed as radio buttons with a single possible answer for an entry. The possible answers will be identified as numbers, which is important for the data import

-Field Label: describe the field

- Choices, Calculations, OR Slider Labels: the response options for the radio buttons

-Field Note: information to assist in data entry.

-Text Validation Type OR Show Slider Number: other information to assist in data entry. If it is not satisfied, a message error will be displayed.

-Identifier: if the field contains sensible information about the person (in this case, being a historical register, no information was deemed sensitive).

-Branching logic: Branching logic can be applied to a field to specify whether or not it will be displayed, depending on values in previous fields. For example, “Type of Religious” can be specified only if the Legal status of the subject is “*Religioso*”.

-Required Field: if the field is left blank, an error is displayed, and moving on to the following field will be impossible. If the field is left blank, an error is displayed, and it will be impossible to move on to the following field.

-Custom Alignment: The location of text boxes or categorical responses (dropdown, radio, checkbox)

-Matrix Group Name: a tag that is used to group all the fields in a single matrix group

-Matrix Ranking

Variable / Field Name	Form Name	Section Header	Field Type	Field Label	Choices, Calculations, OR Slider Labels	Field Note	Text Validation Type OR Show Slider Number	Text Validation Min	Text Validation Max	Identifier?	Branching Logic (Show field only if...)	Required Field?	Custom Alignment	Question Number (surveys only)	Matrix Group Name	Matrix Ranking?	Field Annotation
Id	Daae		text	Subject ID		Unique number	integer	1	1900								
registry_number	Daae	GENERAL INFORMATION	text	Number of the registry								y					
number_in_registry	Daae		text	Number of the subject in the registry		If Present	integer	1	1900								
Name	Daae	SOCIO-DEMOGRAPHICAL INFORMATION	notes	Name of the subject													
Sex	Daae		radio	Sex of the subject	1, F | 2, M												
age_days	Daae		text	Age in days		Age have to be expressed in days or in years.	integer										
age_years	Daae		text	Age in years		Age have to be expressed in days or in years	integer										
profession_subj_cat	Daae		radio	Professional category of the subject	1, canevario | 2, giurisperito												
subject_legal_status	Daae		radio	Legal status of the subject	1, vedova | 2, religioso | 3, servo												
subject_religious	Daae		radio	Type of religious	1, presbitero | 2, suora | 3, frate | 4, canonico | 5, monaca						[subject_legal_status] = '2'						
honorary_subject	Daae		radio	Hororary titles of the subject	1, dominus | 2, maestro | 3, domina												
honorary_relatives	Daae		radio	Honorary titles of relatives	1, maestro												
relatives_name	Daae		notes	Name of the relatives		Raccomandabile farlo col text mining											
degree_kinship	daae		radio	Degree of kinship	1, _MARITO_												

					| 2, _PADRE_												
relatives_alive_death	Daae		radio	Is the relatives alive or death?	1, _DEFUNTO_	Death: present in the year of the registry											
Sestiere	Daae	GEOGRAPHICAL INFORMATION	radio	Porta (door)	1, Romana | 2, Orientale | 3, Vercellina | 4, Comasina | 5, Nuova | 6, Ticinese												
Parish	Daae		notes	Parish													
foris_intus	Daae		radio	Is the parish in a foris area or in an intus area?	1, foris | 2, intus												
X	Daae		text	Longitude of the parish			number										
Y	Daae		text	Latitude of the parish													
district_id	Daae		text	Id of the district			number										
district_name	Daae		notes	Name of the district													
Date	Daae	CLINICAL INFORMATION	text	Date of death			date_ymd	1480-01-01	1480-12-31								
cause_death_registry	Daae		notes	Causes of death		Please report the causes of death as described in the registry											
cause_death_code	Daae		text	Cause of death code			number										
cause_classification_italian	Daae		notes	Classification of the cause of death (italian)													
classification_death_eng	Daae		notes	Classification of the cause of death (english)													
fever_type	Daae		text	Fever type													
violent_death	Daae		radio	It was a violent death?	0, 0 | 1, 1												
place_type	Daae		radio	Type of the place of death	1, _CASTELLO_ | 2, _MONASTERO_ | 3, _OSPEDALE_												
name_place	Daae		notes	Name of the place of death													
Necroscopo	Daae		notes	Necroscopo													

Table 1: data dictionary of the project

For plague years, some information on causes of death could be added, for example, the symptoms (Figure 3).

Figure 3: additional information on the causes of death in case of plague registry.

3 1480: A PEACEFUL PERIOD FREE FROM PLAGUE

3.1 Introduction

Plague causes social, demographic, and economic changes due to the rapid and high decline in the population in a particular area. In addition, the plague's selective effect led to a low number of frail individuals immediately after the event (DeWitte, 2014).

In a period with plague, most of the causes of death are attributed to this illness, so only with the study in a period free from this is it possible to know the typical causes of death in a particular historical period. In this work, 1480 was selected because it was free from plague, far from the previous plague outbreak, and characterized by peace. Therefore, it could represent other (rare) periods with this characteristic in the story of Renaissance Milan. In particular, this period's analysis depicts Milan's epidemiological and demographic profile in the last part of the 15th century.

The analyses were made for all causes of death, including comments about the ones not directly due to illness (for example, homicides). The research on the pathological classification of the causes of death is still in progress.

Children's death (subjects under eight years old) is an essential indication of the health conditions of an area; perinatal and infant mortalities are associated with biological and socio-economic factors (Younis, 2014). Child mortality was high during the Renaissance: in Milan, in 1470, it was estimated that 5% of newborns died on the first day of existence. In Florence, during the fifteenth century, it was estimated that 45% of children died before the age of twelve, and in Pistoia, for the same period, 18% of children died in the age class 1-4 years, 11% between 5-9 years and 11% from 10 to 14 years old. Those deaths were due to epidemic diseases, chronic malnutrition, poverty, and violence (King, 2008).

Dante Zanetti, as presented before, took into consideration 1503, 1610, and 1783 for his study and reported that for the first year, the death in the neonatal period was 4.3%; for the second, the 5.9%, and for the third, 5.4% (the denominators referred to the people died for those causes: childbirth and complications of pregnancy, childbirth and the puerperium, neonatal death, accidental or violent causes, age of the mother, and undefined morbid states, morbid states defined by symptomatic or etiological diagnoses). However, there is probably an underestimation of the data for 1503 because 500 subjects did not have a cause of death written in the registry (Zanetti, 1976).

In the analyses of our work, gender was considered: generally, in the Renaissance, a woman's destiny was to become a mother and have children to give offspring to the family. Becoming pregnant and giving birth was an honor for a female. Childbirth, nursing, and childbirth again were an endless, natural cycle for a woman, except for the rich ones who trust other women in nursing. The greater number of pregnancies for rich women was because sex was avoided during the nursing period. After all, milk was believed to affect a pregnancy negatively. It was estimated that 10% of mothers died because of childbirth (King, 2008). Zanetti estimated that 1.7% in 1503, 1.1% in 1616, and 0.9% in 1783 deaths were attributed to childbirth, complications of pregnancy, and the puerperium. So, there is a decline over time. Despite everything, there were more deaths in males than in women (Zanetti, 1976).

In this study, the analysis of death data from a peaceful year examines the geographical distribution of reported deaths by broad age categories. The use of GIS (Geographical Information System) has multiple implications in public health: for example, a significant understanding of the socio-cultural determinants of health outcomes (with the spatial analysis, it could be assessed whether there is a heterogeneity of the disease distribution), the disease surveillance and early warning system (the latter is possible with the study of the geographical distribution of the incidence and the time evolution of the situation), the assessment of the impact of the public health intervention, the analysis of a better allocation of the resources. A GIS system comprises georeferenced data (as a vector, points, polygons, lines, or raster), hardware, software, algorithms, and users. It does not require knowledge of a programming language (Fletcher-Lartey & Caprarelli, 2016).

Beyond the possibilities of GIS in modern public health, it can be used by historians and epidemiologists to understand better historical disease outbreaks, for example, the diffusion of the Black Death in Sweden (Skog & Hauska, 2013). In addition, the system could be used for research and teaching tools about the spatial, kinetic, and sensory dimensions of early modern urban societies: a project involved the city of Florence and led to the creation of a GIS tool, the "DECIMA" (Digitally Encoded Census Information and Mapping Archive), and a mobile app, the "Hidden Florence" (Terpstra & Rose, 2016).

We performed a spatial analysis on the density of the death events in the *Contrade* of Milan in three ways: with maps, Moran's I (Manfredi & Getis, 2010), and LISA (Local Indication on Spatial Autocorrelation) (Anselin, 1995). Mapping represented the distribution of the event in the city; Moran's I was used to assess the presence of a statistically significant spatial autocorrelation; LISA was used to detect the position of the spatial clusters.

Moran's I was previously used to assess the presence of spatial autocorrelation in a work about the social geography of London in different periods (Cummins et al., 2016). Considering evidence from the 16th and early 17th centuries, the authors aggregated data from prior studies, finding that the social distribution of wealth in the city had stayed the same over time. Wealthy parishes were clustered in the city’s center, and the socioeconomic conditions of residents declined, approaching the urban periphery.

Moran's I was used in our analyses because of the lack of information about the socio-economic situation in different parts of the city in 1480: the density of death events, particularly in children, could strictly relate to the socio-economic conditions. The cluster analysis (LISA) was used to assess the position of spatial clusters in the territory (clusters are divided into three categories: significant local spatial autocorrelation that is positive, significant local spatial autocorrelation that is negative, and non-significant spatial autocorrelation).

Temporal analysis of death events can be used to study whether the socio-economic conditions of a specific area or other conditions that connect with the event improved over time or were instead associated with seasonality. With this aim, the Durbin-Watson test to assess temporal autocorrelation and the representation of heatmaps were used. The latter represents the peaks of deaths in different *Contrade*. Here again, a prior study with historical data used heatmaps to assess the different times of peak plague deaths in Venice during 1630 (Lazzari et al., 2020)

3.2 Materials and methods

Data from the register of 1480 were translated from Latin and put into an Excel spreadsheet by an expert paleographer.

Data were rendered in an appropriate format to put into REDCap, to make them readable by software such as R (R Core Team, 2021), and to conduct statistical and territorial analysis. The following steps were performed on the first input of the data:

The potential repetition of the same record was checked and removed.

The missing data in the field were codified with the same code: the empty cell, NA, was chosen for this purpose. If the information about the missing data could be deduced from other fields, a correction has been made (i.e., a subject had an NA for Sex; the name of this subject was *Giovannina*, so that the sex could be coded as female).

A spelling or typing error correction was performed when it could be deduced from other fields (i.e., A. *Maria al Circolo*, in the field about the parish, was corrected as *S. Maria al Circolo* because the first was an exception concerning the rest). Those cases where the inconsistencies or errors could not be attributed to typing were solved after an expert checked the manuscript.

The descriptive analysis of deaths regarding age and sex was made using R (R Core Team, 2021).

Based on material found on the internet (MacMoreno), it was performed a digitization and a reconstruction of the distribution of the *Contrade* and the *Sestieri* in the city of Milan, along with the position of the Spanish walls and the Medieval Walls, with the utilization of the Qgis ver. 3.13.3 (Qgis Developer Team, 2021). The digitization was made to obtain the items with the spatial coordinates to perform the spatial analysis. The coordinates system used in the project was UTM WGS 84 ZONE 32 N (which refers to Italy).

Parishes can be described as polygons in the territory because they refer to a territorial unit but, to our knowledge, there is no documentation about the borders of the parishes in the studied period, so we chose to represent them by a point geometry in the map, in which each point is the position of the church to which the parish refers.

Some churches contained the registry of 1480 no longer exist, so their positions in a modern map have been found thanks to online materials (Rotta, 1891) (Di Bello, 2017). The locations for those churches are approximative: for example, it was possible to associate the data with the name of the modern street in which they were, but not the precise number. However, the position on a modern map is not approximative for the still present churches (for example, San Babila).

Several parishes were divided into two parts, one inside (*intus*) and one outside (*foris)* the medieval ring of walls (for example, *San Lorenzo foris* and *San Lorenzo intus*). After the position of the churches was found, the presence of each subject of the specification, whether the event occurred in the *foris* or *intus* part of the parish, was considered. For the subject without this specification, the event was considered as happened in the *intus* part if the reference church was situated into the medieval walls and as *foris* otherwise. The impossibility of having the borders of the parishes brought to refer to a unique area for all the indications of *foris*: the one between the Spanish and the Medieval walls. The *Contrade* were present inside the medieval walls.

Analyses and representations were based on the *Contrade* and the *foris* areas because most of the churches were concentrated in the central part of the city and were very near each other, so the representation of the distribution of death in them would potentially lead to a result that is difficult to interpret. On the other hand, the *Sestieri* were relatively large, and the territory of each of them started from the center to the extreme periphery of the city, thus included a heterogeneous population. On the contrary, it was tentatively conceivable that any *Contrada* had a relatively homogeneous population from a socio-economic point of view. The territorial density of deaths that occurred in the area of competence of each *Contrada* and the entire area between the medieval ring and the Spanish walls was expressed as (number of events/areas of the *Contrada* express in square meters) * 1000. The analyses were performed using Qgis (vers. 3.13.3) (Qgis Developer Team, 2021).

So, the number of the events in one *Contrada* was obtained as the sum of the deaths in the churches in the *Contrada*, and the number of the events in the *foris* area was obtained as the number of the events in churches outside the medieval walls.

As presented before, Moran I (Manfredi & Getis, 2010) allows us to evaluate the presence of the spatial autocorrelation of the death's events. Spatial autocorrelation represents the relationship between a spatial matrix and a non-spatial matrix (values of the variable).

In particular, the spatial matrix represents the spatial relationship using spatial weights. The values on the spatial matrix strictly depend on the distance between spatial units. In this matrix, in our case, each *Contrada* was a row and the number of columns was equal to the number of features. Each cell contains the spatial weight. To define spatial weights, distance is needed, and it can be different from the Euclidean distance. In this context, we chose the spatially contiguous neighbors with queen adjacency of the *Contrade* and the *foris* area. This was made because the *Contrade* and the space between the medieval and Spanish walls are polygons. So, in this context, weights were expressed as binary numbers (for a particular feature, the value was 1 for all neighboring parts and 0 otherwise. Neighbors, in queen adjacency, are features that share at least one corner) (Wong, 1999).

As recommended with polygon features, a row standardization of the spatial matrix was performed. This row standardization was made to create proportional weights where features have an unequal number of neighbors (the *foris* area is likely to have many neighbors).

The non-spatial matrix was made by the density of the death event in the *Contrade* and the *foris* area. The non-spatial matrix is expressed as a covariance matrix to calculate Moran's I.

The value of the Moran's I is calculated as follows (R. S. Bivand & Wong, 2018):

$$I = \frac{n \sum_{(2)} w_{ij} z_i z_j}{S_0 \sum_{i=1}^{n} z_i^2},$$

With:

$S_0 = \sum_{(2)} w_{ij}$

n is the number of spatial units

w_{ij} are the spatial weights.

zi is the value for the observation I of the considered variable.

zj is the value for the observation j of the considered variable.

A normality and randomization assumption can be chosen to obtain the distribution of the variable, but for this work the randomization is chosen because it is less restrictive.

For the two assumption the expected value of I is:

$$E(I) = -\frac{1}{(n-1)},$$

The analytical variance is calculated as presented:

$$E_{\mathrm{R}}(I^2) = \frac{n\left[(n^2-3n+3)S_1 - nS_2 + 3S_0^2\right] - b_2\left[(n^2-n)S_1 - 2nS_2 + 6S_0^2\right]}{(n-1)(n-2)(n-3)S_0^2}$$

where $b_2 = \frac{m_4}{m_2^2}$, $m_4 = \sum_{i=1}^{n} z_i^4$ and $m_2 = \sum_{i=1}^{n} z_i^2$

The variance is calculated as:

$$\mathrm{Var}_*(I) = E_*(I^2) - [E(I)]^2.$$

And finally, we have:

$$Z_*(I) = \frac{I - E(I)}{\sqrt{\mathrm{Var}_*(I)}}.$$

That is the normal distribution.

It can be interesting that under the randomization assumption, the test's null hypothesis can be seen as the equiprobability of all the combinations of the data, and our observed situation is only one of those possible permutations (Salima & Bellefon, 2018).

Moran's I is affected by the outliers: the presence of the geographical ones was excluded because all polygons were very near each other. A square root transformation was performed in the case of outliers and asymmetry on the territorial density variable.

Moran's I (Moran, 1948) is the most used to evaluate spatial autocorrelation, and it is essentially a cross-product statistic that involves the spatial and non-spatial matrix. It gives only a value about the totality of the pattern, but it is helpful to assess the presence of spatial clusters. The hypothesis of Moran's I test is about spatial randomness: if the statistic is positive and the test is significant, there is an indication of the presence of clusters. If the statistic is negative and the test is significant, there is a spatial randomness to a greater extent than expected from a random distribution.

So, Moran's I indicate the presence of spatial autocorrelation, taking into consideration all the units.

LISA (Local Indication of Spatial Autocorrelation) is the decomposition of the global I in the contribution of different observations; the decomposed value for each observation is connected with the extension of the cluster around it. The LISA is a function of the observed values in a location and the values in the neighborhood, so it is possible to have the position of clusters and outliers (Anselin, 1995). The algorithm is properly performed only if the Global Moran's I test is significant. The result is expressed by the choropleth map showing where the p-value is significant (the null hypothesis is about the casualty distribution of values across the territory). Spatial clusters can be high-high (a zone characterized by high values of the local I) or low-low (a zone characterized by low values of the local I), but the algorithm also reveals spatial outliers (there is a zone with high value surrounded by zones with low values and vice versa). Generally, the high-high clusters are colored red, the low-low in blue, and the outliers are light blue (low-high) and pink (high-low).

The LISA is calculated as follows (R. S. Bivand & Wong, 2018):

$$I_i = \frac{z_i \sum_{j=1}^{n} w_{ij} z_j}{m_2}$$

where

$m_2 = n^{-1} \sum_{i=1}^{n} z_i^2$

Temporal autocorrelation is the relationship between successive values of the same variable, in this case, the number of deaths weekly. Temporal autocorrelation was assessed using the Durbin-Watson test (Durbin & Watson, 1950) (DURBIN & WATSON, 1951). The number of deaths each week was considered the response variable, and the time was the independent variable.

The Durbin-Watson statistic is based on the errors of linear regression, and it is calculated as follows:

$$d = \frac{\sum_{t=2}^{T} (e_t - e_{t-1})^2}{\sum_{t=1}^{T} e_t^2}$$

The assumptions are normally distributed errors with a mean of 0 and the stationarity of the errors. The first assumption was checked by a normal quantile quantile plot, and the second by the Dickey-Fuller test. The Durbin-Watson test has the null hypothesis of the absence of the first order autocorrelation; a value of d statistic that is 2 relates to the non-evidence of the rejection of this hypothesis, a value between 0 and 2 indicates a positive autocorrelation, and between 2 and 4 a negative autocorrelation.

The subdivision of the time was the same for all subjects, children, and adults, to evaluate the presence of temporal autocorrelation in the same time intervals.

The relationship between the number of deaths each week and the time was not linear (based on graphical representations), so a restricted cubic spline with four knots was used. For the analysis, it was considered that our data started on 23 January because, in the registry, there was no indication of the deaths that occurred before that date.

To analyze if the *Contrade* has a peak of the density of the event in same periods, heatmaps were used, and the temporal intervals were in months.

All the analyses were made using R (ver. 4.0.4) (R Core Team, 2021). The proximity analysis was made using the library "rgdal" (R. Bivand et al., 2022), "GISTools" (Brunsdon & Chen, 2014), and "mapview" (Appelhans et al., 2021). Library "spdep" (R. S. Bivand & Wong, 2018) was used to perform the Moran'I.

For the spatial local autocorrelation were used the packages "maptools" (R. Bivand & Lewin-Koh, 2022), "classInt"(R. Bivand, 2020) and "gstat" (Gräler et al., 2016; Pebesma, 2004).

Heatmaps were performed using the package "gplots" (Warnes et al., 2022), Durbin Watson test was performed using the package "lmtest" (Zeileis & Hothorn, 2002). Splines were performed using the packages "rms" (Harrell, 2017).

3.3 Results

The registry of 1480, one of the rare periods of peace and without epidemics, in which the registration book started on January 23rd, included information about 1813 subjects. Although it was impossible to find the recordings of the first 22 days, the data were apparently accessible in the 19th century to Ferrario (Ferrario, 1840), who reported for January 1480 a total of 170 deaths and a total of 1948 deaths in the entire year. Table 2 compares the monthly deaths in 1480 of our data to those reported by Ferrario. Only minimal differences were evidenced, apart from the deaths in January.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Tot
Ferrario	170	190	144	194	162	106	191	192	130	159	164	146	1948
Present study	50	178	150	194	161	107	190	192	131	159	153	148	1813

Table 2: Monthly deaths in 1480 according to Ferrario and the data in our disposition.

Table 3 presents the distribution of age at death in years by gender of the 1,811 subjects for whom age was reported. Females had a slightly older age at death (26,78 years in mean and 18 years in median) than males (25,55 years in mean and 14 years in median). Other measures (minimum, maximum, first, and third quartile) are very similar for females and males, so the dispersion of the distribution of age between the two genders is very similar. Women have a longer life expectancy than men in modern times and conditions of famine, slavery, and epidemics (Zarulli et al., 2018).

	Mean	**Median**	**Min**	**Max**	**First quartile**	**Third Quartile**	**Standard deviation**	**Includes subjects**
Overall	26,17	17,00	0	100	1	50	27,57	1811
Females	26,78	18,00	0	100	1	50	27,58	910
Males	25,55	14,00	0	100	0	50	27,56	901

Table 3: Age (years) at death and distribution by gender.

The distribution of age classes (5 years interval in each class) of death between females and males is represented in Figure 4. It can be observed that, for each age class, the distribution is similar between the sexes, but the majority of deaths occurred in the class [0,5]: about 41% of the total deaths in males and 39% of those in females occurred in children aged 0-5 years.

Children's mortality is closely associated to the life expectancy in society: it has been estimated that if the latter is less than 30 years, then 40% of mortality will occur in children; if the life expectancy is 50 years, the children's mortality will be 20%, and if it is above 60 years, children mortality will be 10% (Woods, 2006).

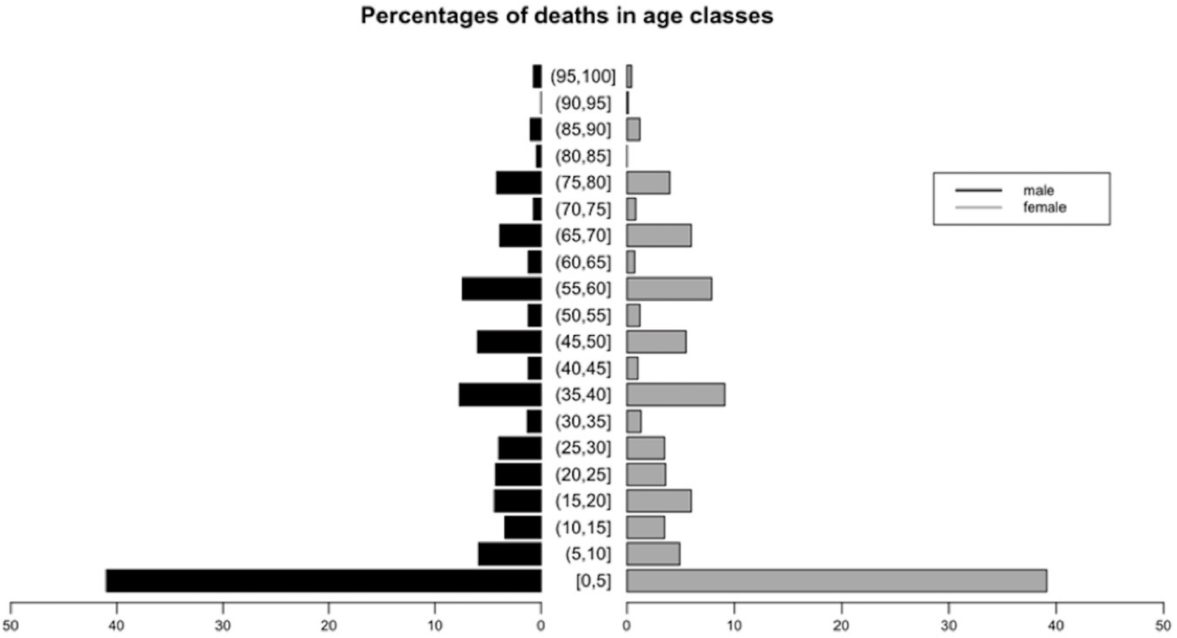


Figure 4: Comparison of percentages of deaths in the different age classes in males and females.

According to this historical source, in the Milan of 1480, the median age at death for adults was 40 years, as shown below.

Interestingly, among the 1046 subjects remaining in the analysis after excluding children (younger than eight years old), no other differences in the age at death between males and females were observed (Table 4). For females, the mean was 44,32 years, and the median was 40 years; for males, the mean was 44,60 years, and the median was 40 years. Once again, the dispersion measures testify that the dispersion of the distribution of age of the two genders is very similar.

	Mean (years)	**Median (years)**	**Min (years)**	**Max (years)**	**First quartile (years)**	**Third Quartile (years)**	**Standard deviation (years)**	**Included subjects**
All subjects	44,45	40,00	8,00	100,00	24,00	60,00	22,84	1046
Female	44,32	40,00	8,00	100,00	24,00	60,00	22,85	540
Male	44,60	40,00	8,00	100,00	25,00	60,00	22,84	506

Table 4: Distribution of age (years) at the death after excluding infant mortality.

The over mortality of males during childhood could be explained by genetic causes (biological hypothesis): having XY chromosomes makes males more susceptible to recessive diseases on the x chromosome because the second x chromosome is not present as a counterpart, making recessive traits non-expressible. Some environmental factors affect the baby's health both in the uterus and after birth: parental stress, medical conditions during conception, occupation, and exposure to environmental risks; these conditions mainly affect males. An analysis involving 50.000 twins in sub-Saharan African countries tested these two hypotheses on male over mortality and gave results affirming the role of environmental factors. Twins were chosen to analyze because they are exposed to the same condition (Pongou et al., 2017). However, other research indicates the fact that the sex ratio at conception is unbiased and that the female mortality during pregnancy exceeds that of male one (Orzack et al., 2015).

The register included 104 parishes, a large number for a relatively small territory, mainly concentrated primarily in the central part of the city. The number of deaths observed in each parish was variable, probably reflecting their different sizes, populousness, and the social conditions of parishioners. It was necessary to assume an equal distribution of the population in different parts of the city and parishes because, to our knowledge, no documentation exists about the number of people in the parishes in this time period. Ancient maps in the medieval and Renaissance periods were few and highly idealized. Based on the information in the 16th and 17th centuries from D'Amico (D'Amico, 1994) ,the poorest segments of the population were distributed throughout the city, with greater density in the more peripheral parts and some areas of the center. No information is available in this regard on the situation in the 15th century. However, it can be assumed that the majority of the poorest sections of the population resided outside the medieval walls.

Figure 5 shows the location on the map of Milan of each church and corresponding parish in which a larger than usual number of deaths occurred (defined as deaths higher than 1% of the total deaths observed).

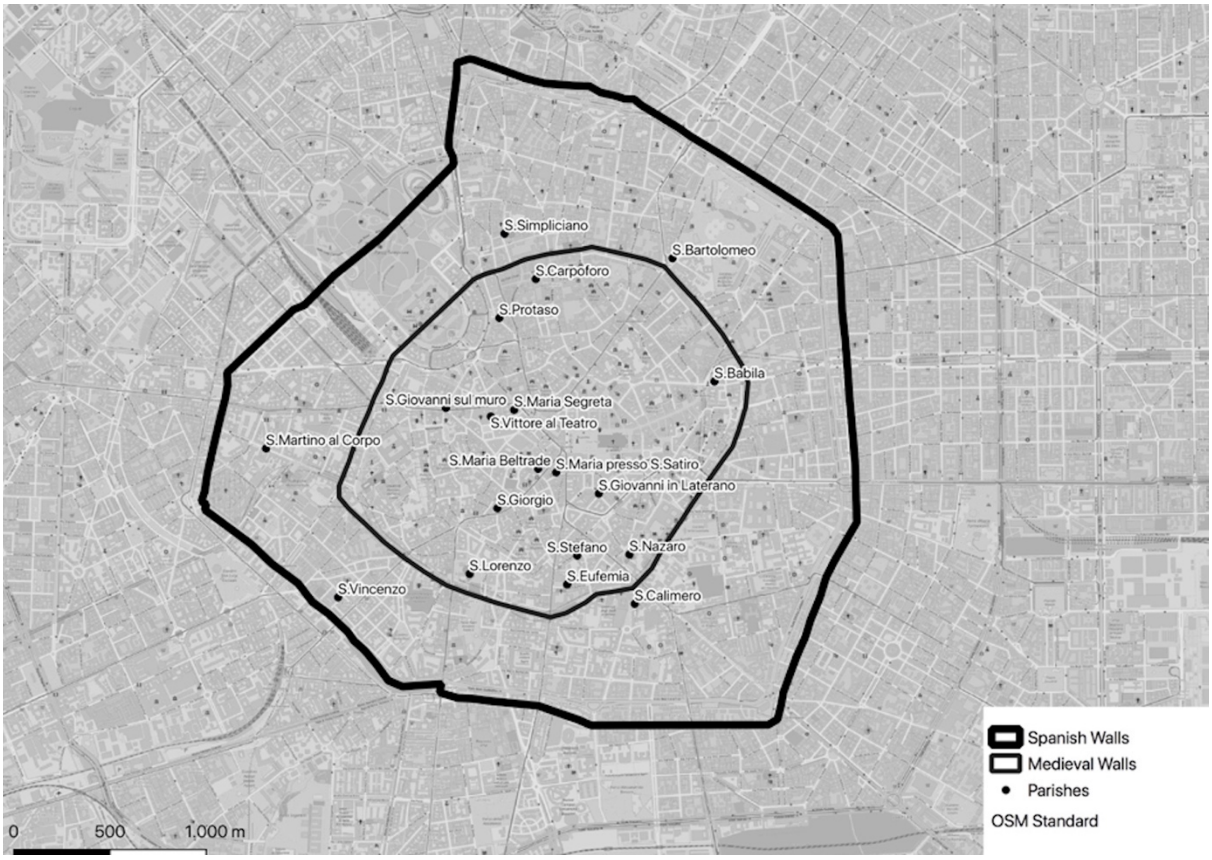


Figure 5: Distribution on the Milan map of the parishes with more than 1% of the total deaths.

Table 5 shows the percentage of deaths in parishes with high percentages of events. The highest number of deaths was seen in the parishes of *S. Lorenzo foris, S. Protaso foris, S. Simpliciano foris, S. Nazaro intus*, and *S. Calimero foris.*

Parish	% (N=1813)
S.Lorenzo foris	10,4
S.Protaso foris	5,6
S.Simpliciano	5,4
S.Nazaro	5,2
S.Calimero	5,1
S.Bartolomeo	4,3
S.Stefano foris	3,9
S.Stefano intus	3,5
S.Babila intus	3,3
S.Babila foris	3,1
S.Martino al Corpo	2,9
S.Protaso intus	2,9
S.Carpoforo intus	2,2
S.Eufemia intus	2,2
S.Maria Beltrade intus	1,8
S.Vincenzo foris	1,7
S.Giorgio	1,5
S.Lorenzo intus	1,5
S.Eufemia foris	1,4
S.Vittore al Teatro	1,3
S.Giovanni in Laterano	1,2
S.Giovanni sul muro	1,2
S.Maria presso S.Satiro	1,2
S.Maria Segreta	1,1

Table 5: Percent distribution of deaths higher than 1% of the number of deaths by the parish.

Figure 6 illustrates the distribution of dead people in the *Contrade*, considering each *Contrada*'s area and the area between the Spanish and medieval walls.

The number and the percentage of dead people in the *Contrade* and the *Foris* area are in Table 6. The high density of events in *Contrade* 21, 16, 12, 11, 10, and 3 could reflect their larger population than the others.

The high density of events in parishes outside the medieval walls may reflect that the poorest part of the population was in this area. For parishes inside the medieval walls, the high density of events may be due to a high population density. Overall, there were 994 (54,8%) deaths occurred in the areas inside the medieval walls, and 819 (45,2%) outside (Table 6). This suggests that death events were higher in the suburban (*foris*) area since it was less populous.

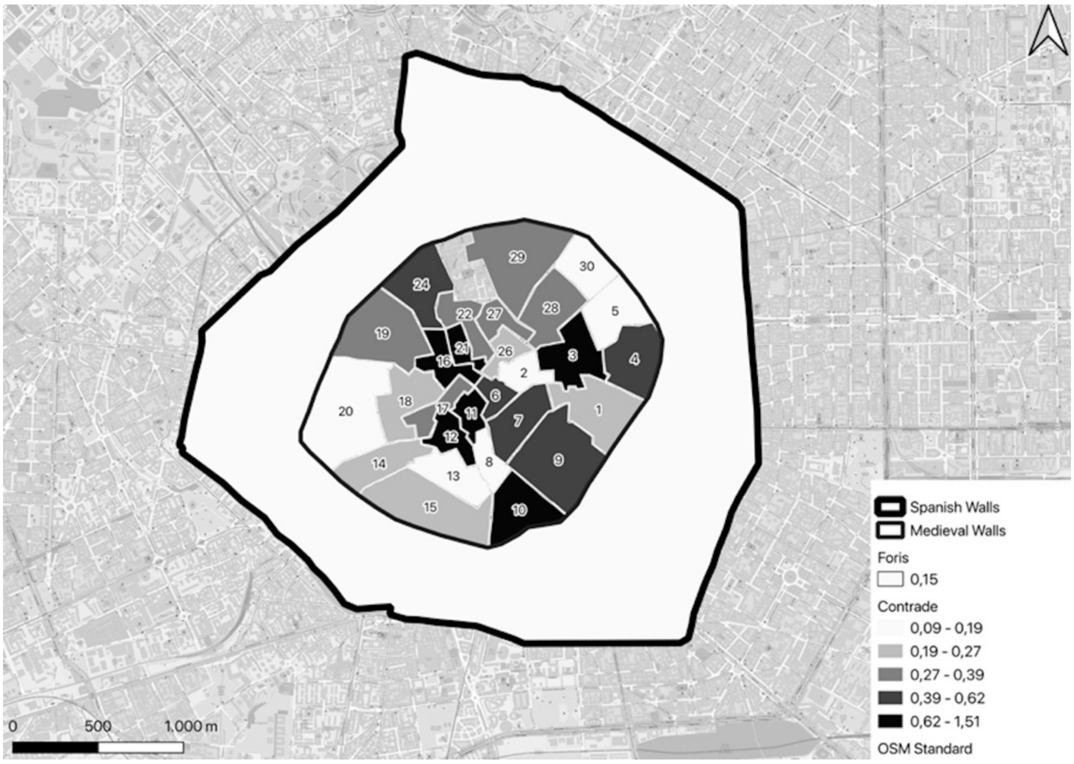


Figure 6: Distribution of the deceased people in the *Contrade*, considering each *Contrada*'s area and the area between the Spanish and medieval walls.

ID (Name)	**N= 1813 (%)**
INTUS	54.8
9 (Contrada del Brolo)	6,5
10 (Contrada delle Capre)	5,6
3 (Contrada dell'Agnello)	3,5
4 (Contrada della Cerva)	3,3
24 (Contrada del Campo)	2,9
29 (Contrada degli Andegari)	2,9
19 (Contrada della Porta)	2,8
15 (Contrada della Vetra)	2,7
7 (Nobile Contrada della Cicogna)	2,6
11 (Contrada della Lupa)	2,6
21 (Nobile Contrada del Cordusio)	2,4
16 (Contrada della Piscina)	2,1
28 (Contrada della Mazza)	2
12 (Nobile Contrada di Sant'Ambrogio)	1,8
1 (Contrada del Verzaro)	1,5
18 (Contrada dei Morigi)	1,2
20 (Contrada del Nirone)	1,2
14 (Contrada del Torchio)	1
22 (Contrada del Rovello)	1
5 (Contrada di Bagutta)	0,8
6 (Contrada del Falcone)	0,8
17 (Nobile Contrada della Rosa)	0,8
27 (Nobile Contrada dei Bossi)	0,8
13 (Contrada delle Cornacchie)	0,4
26 (Contrada Capitana)	0,4
30 (Contrada della Spiga)	0,4
2 (Nobile Contrada delle Farine)	0,3
8 (Contrada del Fieno)	0,2
FORIS	45,2

Table 6: Number of deaths and the percentage of the total deceased persons in each *Contrada* and the area outside the *Contrade*. The denominator was the subjects in the registry.

Distribution of the causes of death

Table 7 shows the percentage of deaths by cause. The deaths attributed to a defined disease are grouped.

Causes of death	% (N=1813)
Perinatal mortality	19,2
Children deaths, no causes reported	16,2
Childbirth or related causes	1,7
Old age	1,2
Unknown/ undetermined/ unreadable	1,2
Starvation	0,4
Other causes of death	58,6
Not natural/Violent causes of death	
Accident	0,8
Homicide	0,7
Suicide	0,2
Drowning	0,1

Table 7: Deaths occurred in 1480, grouped by identified causes.

As discussed in more detail below, health authorities paid limited attention to registering the causes of infant death. Thus, perinatal mortality, the most frequently cited cause of death, was reported without further specification.

Children deaths

The total number of deaths registered for children 0 to 8 years old was 765, of which 395 (51,6%) were observed in males and 370 (48,4%) in females. Table 8 shows the distribution of age at death in children.

	Mean (years)	**Median (years)**	**Min (years)**	**Max (years)**	**First quartile (years)**	**Third Quartile (years)**	**Standard deviation (years)**	**Number of subjects**
All subjects	1,16	0,00	0,00	7,00	0,00	2,00	1,78	765
Female	1,17	0,00	0,00	7,00	0,00	2,00	1,65	370
Male	1,15	0,00	0,00	7,00	0,00	2,00	1,90	395

Table 8: Distribution of age (years) at death for children aged 0 to 7

About 50% of infant deaths occurred within the first year of life, and 75% before the end of the second. The median and the mean ages at death for females and males are very similar. To our knowledge, these data on infant mortality in Milan in 1480 are the oldest available today among those taken from death registers. It is known that the deaths observed in the first 27 days of life should be mainly attributed to 'endogenous' factors such as inborn diseases, congenital abnormality, birth trauma, or pre-term birth, while deaths occurring from the 28th day to 1 year of age are more likely to 'exogenous causes', such as infectious disease, malnutrition, and poor living conditions (Bourgeois-Pichat, 1951) (Lewis & Gowland, 2007).Seasonal variation, and the illness or death of the mother are also important factors. A higher mortality risk in males during childhood also occurs in modern times. For example, in Italy in 2016, the mortality of males was 3 for one thousand live births, males and females, was 2 (Murianni & Tinto, 2019).

Table 9 shows the age at death in infants, i.e., in children younger than 28 days.

	Mean (days)	**Median (days)**	**Min (days)**	**Max (days)**	**First quartile (days)**	**Third Quartile (days)**	**Standard deviation (days)**	**Number of subjects**
All subjects	5,80	6	0	21	1	8	4,86	233
Female	6,59	7,5	1	21	1	8	5,31	94
Male	5,27	4	0	20	1	8	4,47	139

Table 9: Age at death in infants (subjects with less than 28 days at death)

These observations are similar to those reported in recent times in disadvantaged populations in low-income countries; in a study conducted in Ghana about data between January 2003 and December 2009, 64,8% (of 424 neonatal deaths) occurred in the first week of life (Welaga et al., 2013).

Excess of mortality in male infants has been well described in contemporary and historical populations (Stinson, 1985) (Drevenstedt et al., 2008). In particular, the relative fragility of male infants was confirmed by an analysis of data from fifteen developed countries between the mid-18th century and 1970, where the excess of male mortality increased in association with an overall decrease in infant mortality. The male disadvantage could be related to the decline of the contribution of infectious disease to infant mortality; in this case, infant mortality was increasingly the result of compromised perinatal conditions, conditions that affected boys to a greater extent. The decrease in this male disadvantage after 1970 was due to the decline in the contribution of perinatal conditions to infant mortality (for example, improved obstetric practices and neonatal care) (Drevenstedt et al., 2008). Similarly, in the parish of Christ Church Spitalfield in London, significant excess males infant mortality was seen in the 1750–59 cohort, but not in later decades. The study considered the period 1750-1839. Over time, the risk of death decreased in both males and females. Still, there was significant excess of infant mortality and neonatal mortality rates in male than in females in 1750-59, but not later than this period. This can be due to the reduction of endogenous and exogenous factors. (Humphrey et al., 2012).

The distribution of deaths under the age of eight, stratified by age and parish, is reported in Table 10.

Parish	**Children (%) N=765**	**0-6 months (%) N=378**	**0-6 months (females) (%) N=170**	**0-6 months (males) (%) N=208**	**6 months-7 years (%) N=387**	**6 months-7 years females (%) N=200**	**6 months-7 years males (%) N=187**
S.Lorenzo foris	9,8	5,1	3,7	6,61	4,7	4,7	4,13
S.Protaso foris	6	2,5	1,32	3,7	3,5	3,5	4,65
S.Simpliciano	5,8	2,7	2,65	2,91	3	3	2,58
S.Bartolomeo	5,2	2,9	3,7	2,12	2,4	2,4	1,81
S.Stefano foris	5	2,6	2,65	2,65	2,4	2,4	3,36
S.Calimero	4,6	2,9	2,12	3,7	1,7	1,7	2,33
S.Nazaro	3,7	1,4	1,06	1,85	2,2	2,2	1,81
S.Protaso intus	3,3	1,4	1,85	1,06	1,8	1,8	1,03
S.Babila foris	3,1	1,8	1,59	2,12	1,3	1,3	1,29
S.Eufemia intus	2,9	0,9	1,06	0,79	2	2	1,81
S.Eufemia foris	2,7	0	0	0	2,7	2,7	1,81
S.Babila intus	2,5	1	0,26	1,85	1,4	1,4	1,29
S.Maria Beltrade	2,4	1,3	1,06	1,59	1	1	0,78
S.Martino al Corpo	2,1	1,7	1,32	2,12	0,4	0,4	0,52
S.Stefano intus	2	1,3	1,32	1,32	0,7	0,7	0,52
S.Giorgio	1,8	1,2	1,59	0,79	0,7	0,7	1,03
S.Carpoforo intus	1,6	0,8	0,26	1,32	0,8	0,8	0,52
S.Lorenzo intus	1,6	0,9	0,26	1,59	0,7	0,7	1,03
S.Giovanni in Laterano	1,4	0,9	1,32	0,53	0,5	0,5	0,26
S.Giovanni sul muro	1,4	1	1,06	1,06	0,4	0,4	0,52
S.Vincenzo	1,4	0,9	1,06	0,79	0,5	0,5	0
S.Maria Podone	1,2	0,5	0,26	0,79	0,7	0,7	0,52
S.Michele al Gallo	1,2	0,8	1,06	0,53	0,4	0,4	0,26

Table 10: Distribution of subjects from 0 to 8 years old in parishes. Only the parishes reaching at least 1% of the distribution total of deaths that occurred in children from 0 to eight years old were considered. Columns express percentages.

The parishes with the highest perinatal and infant mortality were:

- S. Lorenzo *foris* (9,8% of children mortality).
- S. Protaso f*oris* (6% of children mortality).
- S. Simpliciano (5,8% of children mortality).

All of them were located outside the ring of medieval walls, in the peripheral part of the city (Figure 7), where a high density of disadvantaged people and wage-earner workers can be presumed. In general, mortality is an indicator of the socioeconomic level, particularly the infant mortality rate

(Haines, 1995).

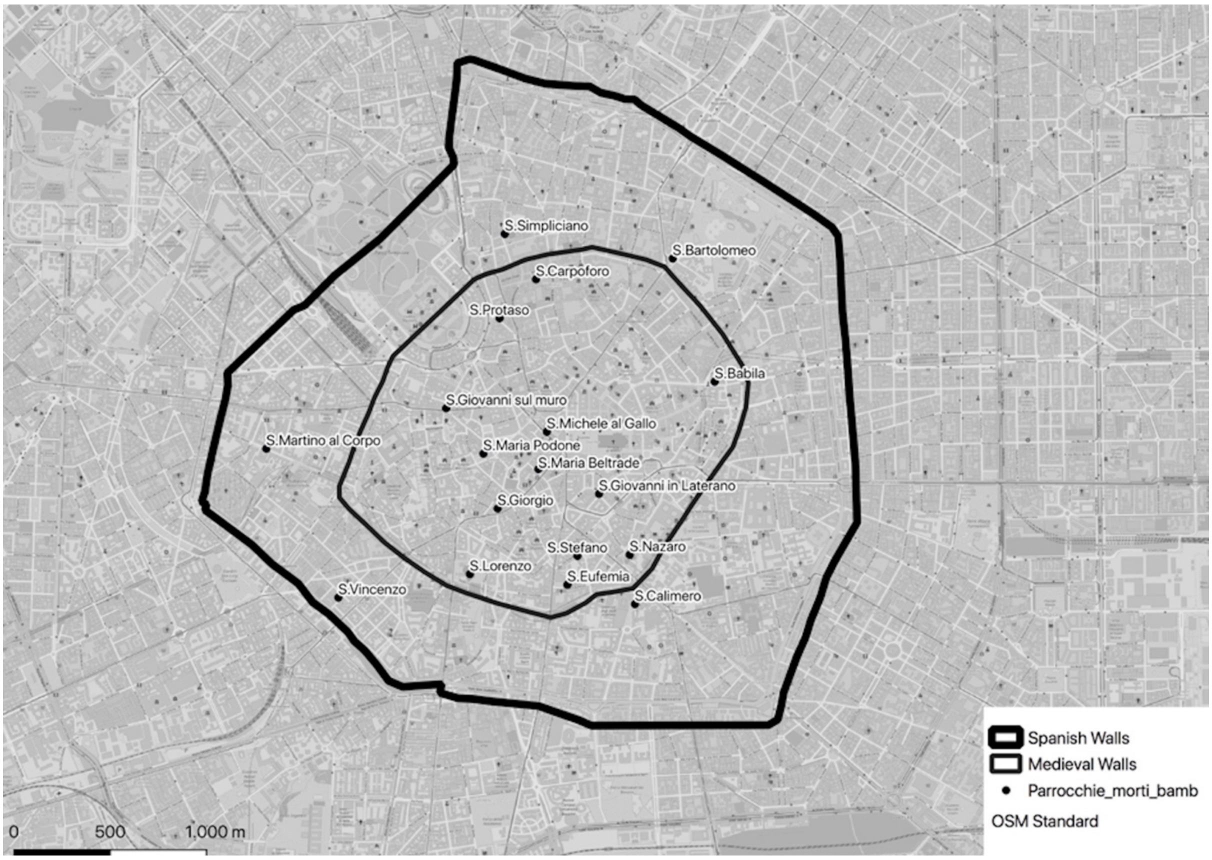


Figure 7: Distribution on the Milan map of the parishes with infant and perinatal mortality more than 1% of deaths under eight years of age.

In London, from 1580 to 1650, a higher male infant mortality was observed in poorer parishes but not the wealthier ones (Finlay, 1981).

Figure 8 shows the distribution of deaths under eight years of age in the different parts of the city.

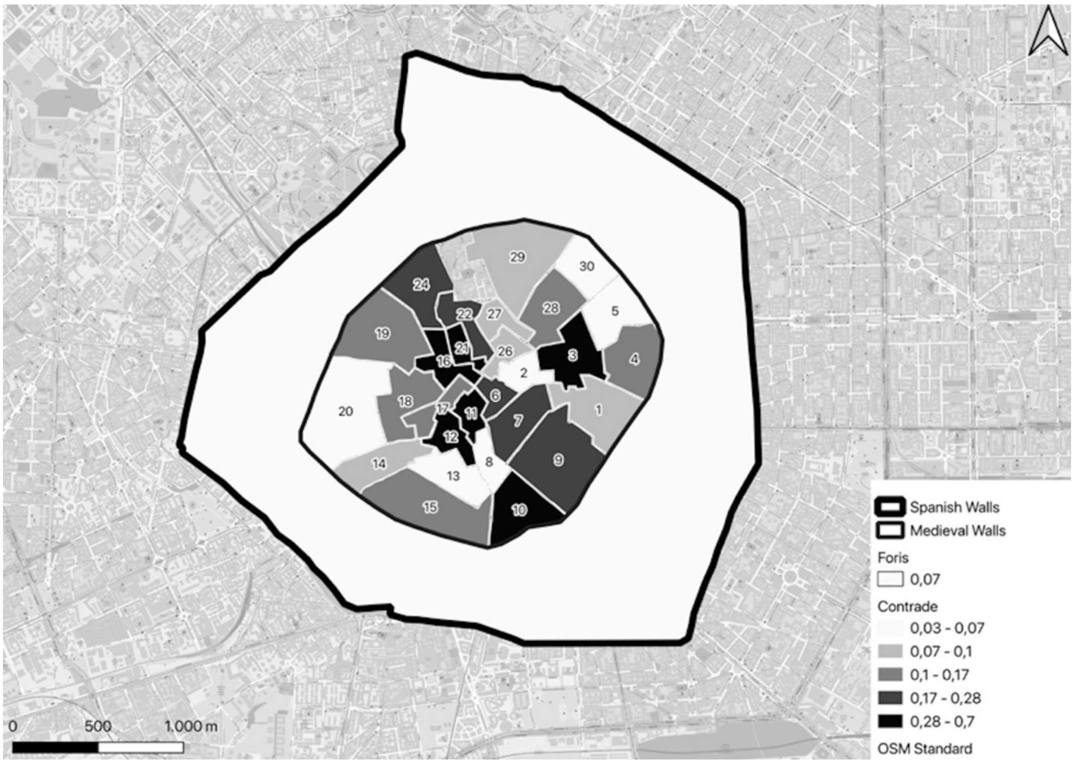


Figure 8: Distribution of the deceased from 0 to 7 years old among the *Contrade* and in the area outside the *Contrade*. The number of deaths people in each part of the city was divided by the surface in square meters of each of them and then multiplied by 1000.

The number and the percentage of child deaths in each *Contrada* and *foris* area are shown in Table 11.

	% (TOT=765)
ID *CONTRADA* (Name)	
9 (*Contrada del Brolo*)	5,1
10 (*Contrada delle Capre*)	4,8
3 (*Contrada dell'Agnello*)	4,4
11 (*Contrada della Lupa*)	3,3
24 (*Contrada del Campo*)	3,3
15 (*Contrada della Vetra*)	2,7
19 (*Contrada della Porta*)	2,6
4 (*Contrada della Cerva*)	2,5
29 (*Contrada degli Andegari*)	2,5
7 (*Nobile Contrada della Cicogna*)	2,2
12 *(Nobile Contrada di Sant'Ambrogio*)	2,2
16 (*Contrada della Piscina*)	2,1
21 (*Nobile Contrada del Cordusio*)	2
28 (*Contrada della Mazza*)	2
1 (*Contrada del Verzaro*)	1,4
18 (*Contrada dei Morigi*)	1,3
20 (*Contrada del Nirone*)	1,3
22 (*Contrada del Rovello*)	1,3
6 (*Contrada del Falcone*)	1,2
14 (*Contrada del Torchio*)	1
17 (*Nobile Contrada della Rosa*)	0,7
5 (*Contrada di Bagutta*)	0,5
13 (*Contrada delle Cornacchie*)	0,5
27 (*Nobile Contrada dei Bossi*)	0,5
26 (*Contrada Capitana*)	0,4
30 (*Contrada della Spiga*)	0,4
2 (*Nobile Contrada delle Farine*)	0,3
8 (*Contrada del Fieno*)	0,1
FORIS	47,3
Total number of children	765

Table 11: Number of subjects from 0 to 7 years old and percentage of the total observed deaths in each *Contrada* and the area outside the *Contrade* (*foris*).

Spatial and temporal analysis

Figure 9 represents the distribution of the deceased people in the *Contrade* and the *foris* area for all subjects (a), for children (subjects under eight years old) (b), and adults (c).

It can be seen that in the *foris* area there is a low density of death events, likely be due to the high area in consideration. In addition, this area was characterized by the *Cassine* (farmhouses) and craft activities; in 1412, this area was populated by modest workers, but there were also houses of prominent personalities. The high distribution of the event in the *Contrade* 3, 21, 16, 11, 10, and 12 could probably be attributed to the high density of the population. In our knowledge, Milan was characterized by social heterogeneity at the *Sestiere* and *Contrade* levels. However, it could be possible that the center of the city was a commercial area, also with homes of merchants and retailers. Moving to the peripheral area, there was probably a residential area and a production area (textiles, weapons, and armor) near the medieval walls (D'Amico, 1994) (Boucheron, 2006). Table 12 presents, for each *Contrada* and the *foris* area, the surface expressed in square meters, the *Sestiere* to which it belongs, the number and the percentage of females, the number and the percentage of children (individuals under eight years old), and the median age with the interquartile range (IQR).

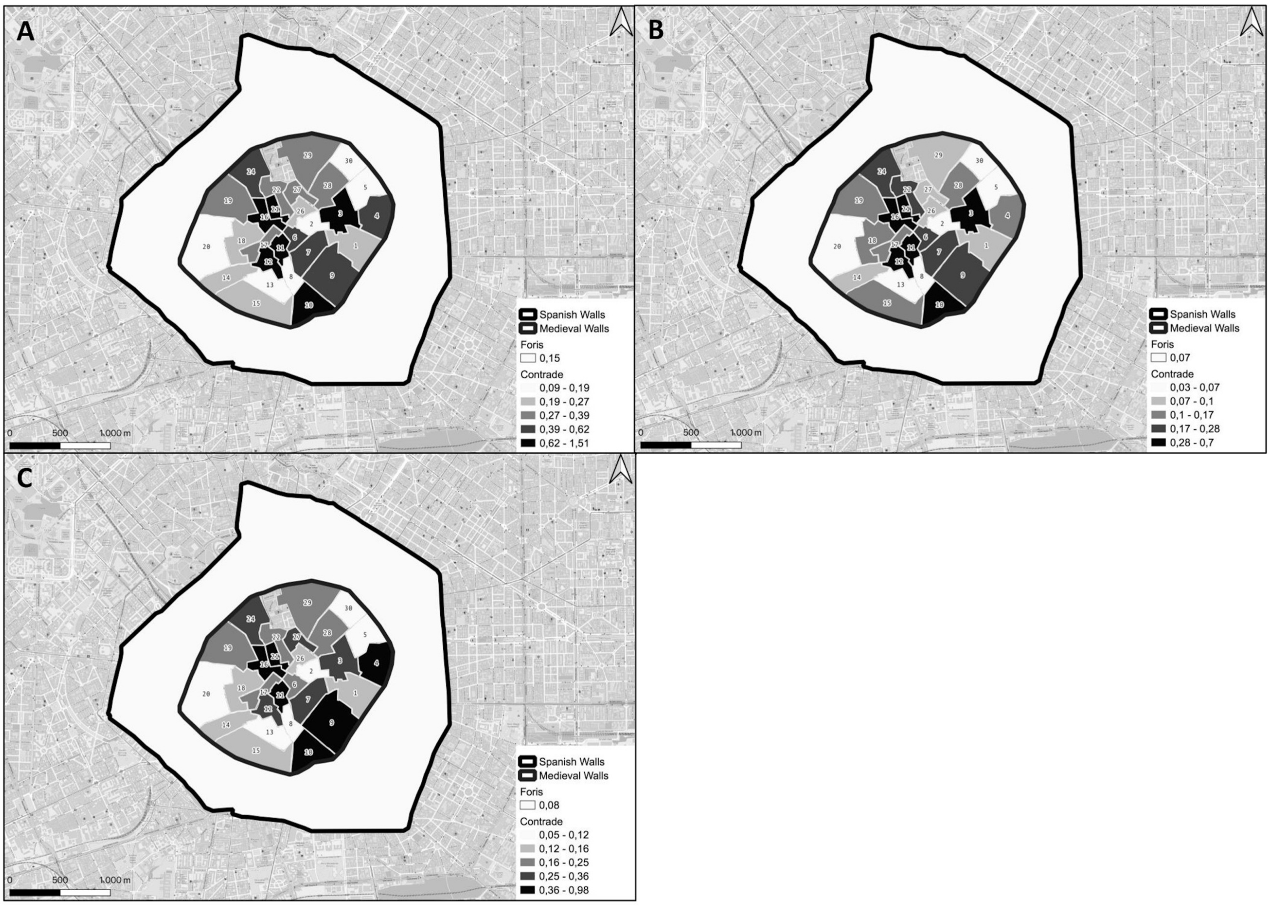


Figure 9: Distribution of the deceased people in the *Contrade* and the *foris* area for all subjects (a), for children (subjects under eight years old) (b), and adults (c).

Table 12 represents the distribution of gender and age in each *Contrada* and in the *foris* area.

ID	**Name**	**Sestiere**	**Surface (m^2)**	**Number of females (%)**	**Number of children (%)**	**Age median (IQR)**
1	Contrada del Verzaro	Orientale	126794	16 (57,1%)	11 (39,3%)	27,0 (56,50)
2	Nobile Contrada delle Farine	Orientale	39551	4 (80%)	2 (40,0%)	36,0 (29,00)
3	Contrada dell'Agnello	Orientale	101678	31 (48,4%)	34 (53,1%)	5,0 (51,50)
4	Contrada della Cerva	Orientale	110032	31 (52,5%)	19 (32,2%)	24,0 (47,00)
5	Contrada di Bagutta	Orientale	91946	8 (53,3%)	4 (26,7%)	40,0 (62,00)
6	Contrada del Falcone	Romana	32892	11 (73,3%)	9 (60,0%)	1,0 (24,00)
7	Nobile Contrada della Cicogna	Romana	89414	21 (43,8%)	17 (35,4%)	12,0 (42,50)
8	Contrada del Fieno	Romana	35705	2 (50,0%)	1 (25,0%)	29,5 (50,75)
9	Contrada del Brolo	Romana	195466	51 (43,6%)	39 (33,3%)	25,0 (54,25)
10	Contrada delle Capre	Romana	104510	53 (52,0%)	37 (36,3%)	22,0 (48,00)
11	Contrada della Lupa	Ticinese	35742	28 (59,6%)	25 (53,2%)	5,0 (28,50)
12	Nobile Contrada di Sant'Ambrogio	Ticinese	49210	11 (33,3%)	17 (51,5%)	6,0 (36,00)
13	Contrada delle Cornacchie	Ticinese	74943	5 (62,5%)	4 (50,0%)	6,0 (34,50)
14	Contrada del Torchio	Ticinese	86363	11 (57,9%)	8 (42,1%)	16,0 (43,50)
15	Contrada della Vetra	Ticinese	184884	19 (38,8%)	21 (42,9%)	22,0 (49,00)
16	Contrada della Piscina	Vercellina	55191	21 (55,3%)	16 (42,1%)	16,5 (52,75)
17	Nobile Contrada della Rosa	Vercellina	39776	6 (40,0%)	5 (33,3%)	20,0 (31,50)
18	Contrada dei Morigi	Vercellina	96595	11 (50,0%)	10 (45,5%)	18,5 (44,75)
19	Contrada della Porta	Vercellina	166107	26 (52,0%)	20 (40,0%)	21,0 (59,00)
20	Contrada del Nirone	Vercellina	233818	10 (45,5%)	10 (45,5%)	24,0 (59,50)
21	Nobile Contrada del Cordusio	Comasina	28490	26 (60,5%)	15 (34,9%)	15,0 (54,00)
22	Contrada del Rovello	Comasina	57225	10 (52,6%)	10 (52,6%)	4,0 (45,00)
23	Contrada dell'Orso	Comasina	44126			
24	Contrada del Campo	Comasina	110958	30 (56,6%)	25 (47,2%)	10,0 (49,00)
25	Contrada dei Fiori	Comasina	28019			
26	Contrada Capitana	Nuova	36023	6 (75,0%)	3 (37,5%)	12,5 (24,00)
27	Nobile Contrada dei Bossi	Nuova	39388	7 (50,0%)	4 (28,6%)	36,5 (67,50)
28	Contrada della Mazza	Nuova	99909	16 (44,4%)	15 (41,7%)	13,0 (59,00)
29	Contrada degli Andegari	Nuova	196947	26 (49,1%)	19 (35,8%)	20,0 (49,00)
30	Contrada della Spiga	Nuova	75919	3 (37,5%)	3 (37,5%)	18,5 (43,25)
	FORIS		5452942	411 (50,2%)	362 (44,2%)	13,5 (48,00)

Table 12: distribution of gender and age in each *Contrada* and in the *foris* area.

It could be noted from Figure 9 that the distribution of the event in children and all subjects is very similar. *Contrade* 3, 10, 11, 12, 16, and 21 are characterized by a high density of death of all subjects and children; it could be noted from Table 11 that a relatively low age depicts *Contrade* 3, 11, and

12 at death. The high density of the event in other *Contrade* could be due to a high population density or a relatively low area.

The distribution of death in adults is very similar to that for children and all subjects.

In Figure 9 a, b, and c, it can be noted that there are some zones with a high density of events close to zones with a high density of the event (for example, *Contrade* 21 and 16 and *Contrade* 11 and 12. The latter not for adults' pattern) and zone with a low density of the event near zones with a low density of events (for example *Contrade* 5 and 30). So, there may be a spatial association between the events. Moran's I statistic was used to assess it.

The global Moran's I statistic for all subjects was 0,14, and the p-value was 0,03884, so there is evidence against the null hypothesis (presence of randomness of the territorial distribution of the density of death), so there is statistical evidence for the presence of spatial autocorrelation and in a positive way (presence of a zone with a high value of the variable near other zones characterized by a high value of the variable and/or presence of a zone with a low value of the variable near other zones characterized by a low value of the density of deaths).

The Moran's I test for children gave the following results: it was significant (p value=0,03185), and the value was 0,15, so there is a signal of the presence of spatial clusters.

The global Moran I test for adults is not significant (p value=0,07367), so there is no evidence of spatial autocorrelation.

Moran's I statistic is a global measure of spatial autocorrelation. To evaluate clusters at a local scale LISA (Local Index of Spatial Autocorrelation) was performed only where the global spatial autocorrelation was significant (so for all subjects and children). Figure 10 represents the clusters according to LISA for those two groups.

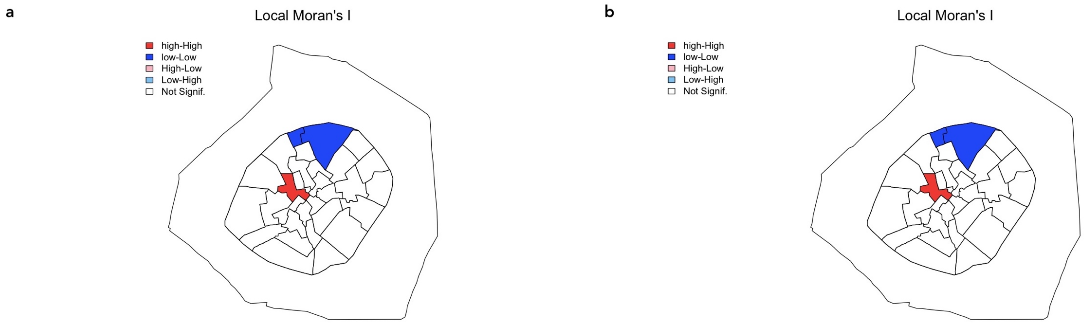


Figure 10: Position of clusters for all subjects (a) and children (b)

The *Contrade* marked by colors are statistically significant on the LISA and represents a high correlation with the local Moran I of the neighboring *Contrade*.

For all subjects, there is a “hot spot” cluster (high-high) corresponding to *Contrada* number 16, which today corresponds to the zone of Galleria Meravigli and Palazzo Turati in Milan. There is a “cold spot” (low-low) cluster in correspondence to *Contrade* 29 and 25, which show a low density value of the events; that area corresponds to the zone of Brera and Borgonuovo in Milan today.

The clusters identified for children by the analysis were the same as the previous one.

Considering those results the pattern of mortality observed for all the subjects may be mainly driven by the pattern of children.

Figure 11 represents the heatmap for all subjects (a), children (b), and adults (c). Heatmaps were made to assess the number of deaths of each *Contrade* month by month.

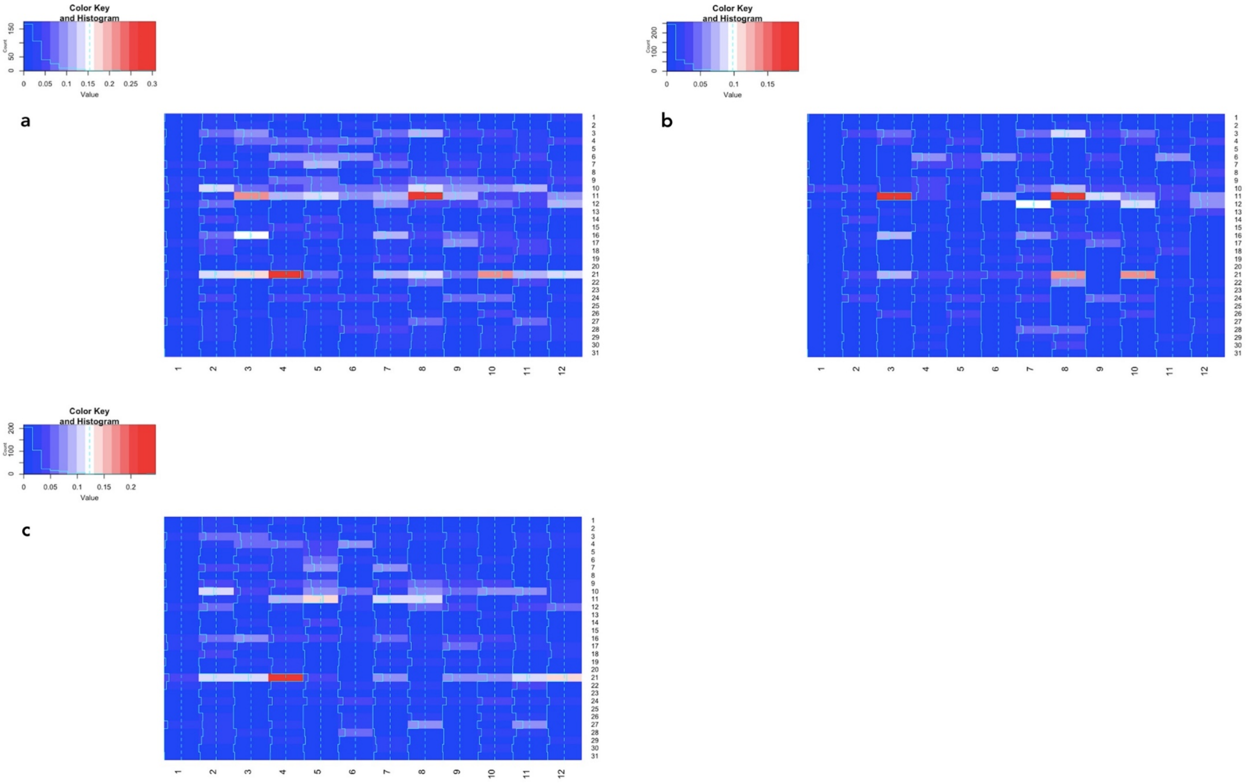


Figure 11: heatmaps for all subjects (a), children (b), and adults (c).
In the x-axis, there is the time (months). In the y-axis, there are the IDs of the *Contrade.*

Referring to all subjects, it can be observed that *Contrada* 21 has two peaks: one in April and one in October. *Contrada* 11 has two peaks: one in March and one in August. Other *Contrade* have many less pronounced peaks.

For children, *Contrada* number 21 had two peaks, one in August and one in October (so the peaks were different between children and all subjects). Also, *Contrada* number 11 had two peaks: one in

March and one in August, as for all subjects. Other *Contrade* had different peaks, but with minor intensity.

For adults, it could be noted that there is only one *Contrada* with a significant peak (number 21 in April) and other *Contrade* have a series of minor peaks.

The Durbin-Watson test for all subjects was significant (p=0,002522), so there is evidence against the null hypothesis. Furthermore, there is a presence of temporal autocorrelation. The value is 1.30, so there is a positive autocorrelation.

The Durbin-Watson test for the temporal autocorrelation for children was significant (p value<0,01), so there is evidence against the null hypothesis. Therefore, there is a presence of temporal autocorrelation. Furthermore, the value of the statistic was 0.93, so there is a positive autocorrelation.

The Durbin-Watson test was not significant (p value= 0,2039) for adults, so there was no evidence of temporal autocorrelation. The value of the statistic test was 1,78.

Considering those results, the temporal autocorrelation observed for all the subjects may be a consequence of the one in children. Positive autocorrelation indicates that the number of deaths in a previous week has a “positive” impact on the considered week.

The trends of the number of deaths in weeks for all subjects, children, and adults are shown in Figure 12: the trends for all subjects and for children are very similar.

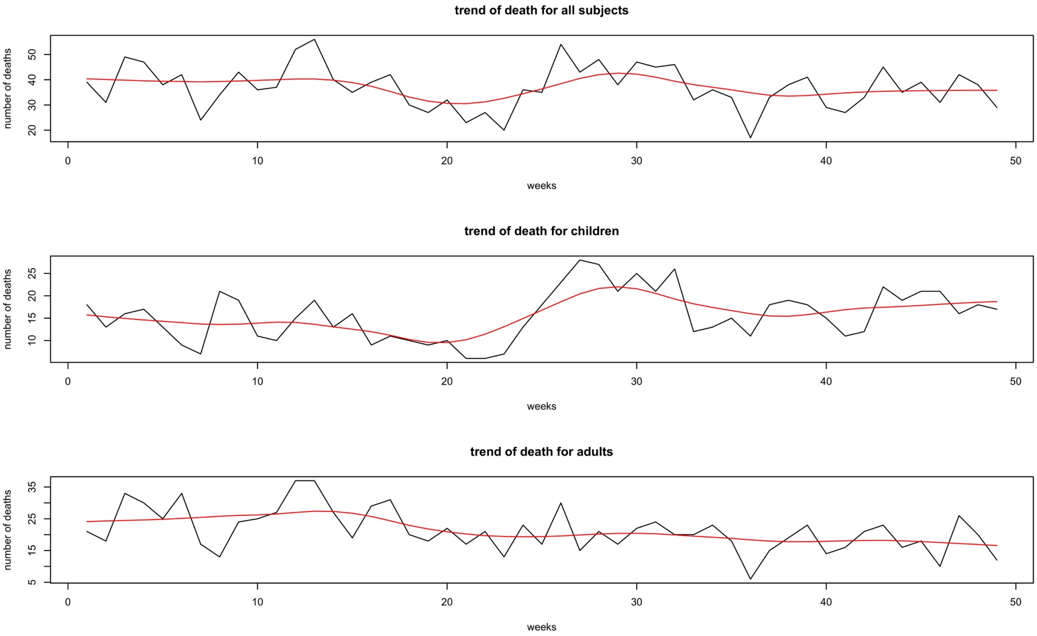


Figure 12: trends of the number of deaths for all subjects, children and adults.

Mortality in foundlings

The abandonment of children could be viewed as a chance to offer them a better probability of survival, considering that the high number of early deaths indicated a precarious condition. Hospitals and foundling homes had the aim to give a future to the foundlings thanks to the adoption and the fosterage (a sort of temporary adoption) by families (sometimes also the family of the "abandoned" children that seeks help to raise the child), a single man or women, or by religious. People who decided to welcome a child had the task of giving moral and material support (it is possible that the dowry promoted the hospitality of males, also if most abandoned children were females) (Albini, 2012).

The number of deaths among foundlings was likely very high, even if it is impossible to estimate mortality without precise data about the number of foundlings (Gavitt, 1990) (D'Andrea & Terpstra, 2017).

The register reports thirty deaths at the San Celso hospital, the foundlings' hospice. These deaths occurred in children between 2 and 9 years (median 3 years) whose foundling condition is confirmed by the surname *da Milano* common to all of them.

Unlike what was observed for deaths in children occurring at home, for these foundlings, the cause of death was reported in all but three cases, for whom the name of the certification author is missing. On the other hand, 25 certifications were given by the same doctor, Maestro Giovanni Casetti, probably the doctor in charge of hospital care. The most frequent cause of death reported is hectic fever, with 11 cases without other association and another 5 deaths associated with other diseases. Hectic fever was a chronic fever 'generated by bones' and often associated with phthisis or long-lasting inflammatory illnesses. The second most frequent cause of death appeared in 10 subjects: diarrheal diseases with or without fever. The distribution during the year of such cases does not suggest the occurrence of an epidemic cluster. Finally, four deaths were attributed to generalized edema, which might be attributed to protein deficiency due to malnutrition.

Deaths in childbirth

Among all deaths registered in Milan in the considered period, 30 concerned women died of childbirth and/or related causes (1,7% of the total cause of death and 3.3% of causes of death of females). The deaths in women of childbearing age (15-45 years) were 229 (42,4% of the total deaths in women). Considering the data in the registers, the mean age of women was 29.83 years (sd=9,21); the median was 30 years. Among the women who died of childbirth, the median age was 28.5 years (sd= 5,81, min=20 years, max=40 years); 12 of them (40%) resided within, and 18 (60%) outside the medieval

walls. Considering all women who died during childbearing age (15-45 years), 120 (52,4%) died within and 109 (47,6%) outside the medieval walls.
In S. Lorenzo *foris*, six childbirth deaths have occurred, one-fifth of the total, further suggesting that parish's probable populousness and possible poverty were responsible.

Deaths in old age and causes of death in the elderly.

Interestingly, according to Galen's hypothesis (the reference for the medicine of the period), aging is related to the gradual extinction of natural heat and humidity (Marinozzi, 2010).
The register reported 21 deaths of old age (9 males and 9 females, median age 90 years, range 80-100) without writing causes other than this. The age in these cases has been presumptively attributed, with an approximation to ten, in almost all of them. Only one of them died in a hospital, and five were supposed to have at death 100 years or more. The centenarians whose death from any cause is recorded in the register of 1480 were seven in all, three males and four females. Overall, deaths aged 70 and over were 205 (11,3%).

Starvation

Six cases between the ages of six and 60, five females, were reported in the register by Mastro Dionigi from Nuremberg (*Dionigi da Norimberga*) with the formula "found undernourished and consumed". In one case, the visit had occurred *post-mortem*, as probably happened in all these cases. It is, therefore, probable that these deaths occurred from starvation in people who had not had access to means of subsistence and had not received assistance before death. Consequently, it can be supposed that starvation was the cause of at least 0,3% of the deaths that occurred in Milan in 1480.

Not natural/Violent causes of death

There were 33 non-natural/violent deaths in the dataset, 1,8% of the total, a percentage almost identical to that detected by Dante Zanetti in the Milano Sforza Registers of 1503 (55 cases out of 3156 deaths, 1,74%) (Zanetti, 1976). Interestingly, in 2012, the percentage of non-natural deaths in the total recorded deaths in Italy was slightly higher (2,27%), but becomes quite similar if the deaths that occurred in transport accidents are subtracted from the total (ISTAT, 2014). A marked difference, however, was seen in the deaths from murder. In Italy, in 2012, homicide deaths were 0,07% of the total deaths recorded, with three-quarters of them were female victims. On the contrary, in the Milano of 1480, the homicide deaths represented 0,66% of the total deaths, and all of them were males with a median age of 27 years (range 17- 50). Although it cannot be excluded that violent deaths of women at any age may have been hidden under other apparently natural causes, it seems possible to conclude that femicide was not frequent in Milan at the end of the fifteenth century. In the study we are completing with the 1629 register, out of 59 murder cases, only one involved a woman. All the men murdered in 1480 were killed by sidearms, probably as a result of stabbings or duels with swords[4].

Due to the exclusion of suicides from burial in consecrated land, some of the deaths from suicide might have been registered as accidental, as in the case of an 18-year-old girl, recorded on January 30, who died from a 'fall from above'. An attempt to justify a suicide with a mental illness is that of *Johannina famula* (servant of) *Magdalene de Grulis,* an 18-year-old girl, registered on June 24[5]. The health officers were not so polite with another 21-year-old girl written on February 12[6].

Among the deaths due to accidental causes, two deserve to be mentioned as an example of the accuracy of the registrations in the MiSfoRe. The first is the case of *Vincenzo di* (son of) *Ambrogio Griffi*, 23 years old *qui balando cecidit super tavolam cum capite et facta forti contusione est cerebrum in 7ª* (dancing he fell on a table with his head and got a severe contusion to the brain and died in the 7th day). The second is the accident that occurred to *Giacomo Pozzobonelli*, a 96-year-old man who died *ex casu ab alto cum fractura cosse sinistre et contusione totius corporis* (for a fall from above with a fracture of the left thigh and bruising of the whole body)[7].

Deaths in hospitals

In addition to the 30 foundlings who died in San Celso, another 122 deaths occurred in 11 other hospitals. Seven of them (San Dionigi, 9 deaths; San Vincenzo, 5; San Lazzaro and Santa Caterina, 4 each; San Giacomo and San Simpliciano, 3 each; *Ospedale Nuovo*, 2) probably functioned as parish hospitals or hospices, with the only exception of San Lazzaro, which was dedicated to the treatment of wounds and skin diseases. Of the other four, the two most extensive (the *Ospedale Maggiore* and the Brolo hospital) recorded 32 and 31 deaths, respectively, while the Sant'Ambrogio hospital and

---

[4] For example , the unfortunate case of Bertolino del Signore, killed at 26 a letali vulneri in radice sinistre partis colli (by a fatal wound at the base of the left side of the neck) or by the case of Francesco da Desio, who died at 24 after "hit with a sword on the skull and also near the nostrils, that show the deep signs, with intervened fever, after the 20th day".

[5] *Semifatua nec ex toto sui iuris, se ipsa et de mente propria arsenicho venenavit et yta proprio presbitero confessa est penitentiam agens et in presentia Johannis Petri de Affori et Bernardini speciarii regis et sui famuli et Josep de Ferrariis quibus fides adhibenda est, presentibus etiam duabus matronis fidedignis* (Semidiot and not totally in the full capacity of her faculties, she poisoned herself with arsenic and on her initiative and therefore she confessed to her presbyter repenting and in the presence of Gio.Pietro da Affori and Bernardino, the sovereign's apothecary, and his servant and Giuseppe Ferrari, to whom credit must be given, also present two trustworthy married women).

[6] *Studiose animi, incitata passione, assumpto veneno argento sublimato die martis proxime passata colirica et mortali succedente passione* (With premeditation, driven by passion of love, taken for poison mercury last Tuesday, with an ensuing fatal colic attack).

[7] The case of this elderly gentleman, who belonged to an aristocratic family included in the *Matricula nobilium familiarum Mediolani* with the act dated 20 April 1377, describes a typical case of death in an elderly person following a fracture of the femur.

the *Ospedale della Pietà*, 18 and 11. These data confirm that in 1480, Milan had an extensive network of active hospitals. Among them was the cornerstone of the *Ospedale Maggiore*, strongly desired by Francesco Sforza, was laid on April 12, 1456. It began to function in 1472 while it was still under construction, and had not yet replaced the small hospitals in assuming the predominant role in health assistance to that the duke intended. Not considering the deaths in the foundling hospitals, the deaths in hospitals represent 0,07% of the deaths recorded in Milan in 1480.

3.4 Conclusion and discussion

Public health surveillance is "the ongoing, systematic collection, analysis, and interpretation of health-related data essential to planning, implementing, and evaluating public health practice" (Introduction to Public Health Surveillance|Public Health 101 Series|CDC, 2018).. In Milan, Duke Gian Galeazzo Visconti (1351-1402) established new public health surveillance practices. He made a notification system mandatory for all illnesses and deaths, established the figure of the health commissioner, which deals with compliance with health regulations, and finally decreed the separation of the healthy from the sick, also thanks to particular spaces outside the city. This collection of information was useful later, in the plague of 1468, to assess that most of the cases were present in the *Cinque vie* (Five Points or Five Streets) area and that some links in the chain of contagion in other parts of the city can be reported to this area (Carmichael, 1991).

The territorial information linked to the causes of death can be crucial to arresting an epidemic outbreak, such as what happened in a cholera epidemic in London during the nineteenth century. The study made by John Snow about the zones where the event took place allowed to establish the causes of the phenomenon (contaminated water) and to put in place measures to stop it (Tulchinsky, 2018).

Temporal analysis of the death event is essential in public health surveillance: comparing all causes of mortality in the same periods for different years allows estimation of the excess mortality, which can translate into public health actions. The UK Health Security Agency (UK Health Security Agency, 2023) put in place weekly all-cause mortality surveillance Fare clic o toccare qui per immettere il testo.More generally, the study can identify peaks due to the seasonality of the causes of death: a historical registry to study the pattern was the Florentine Dowry Fund, established in February 1425, and five periods (1436-1439, 1448-1451, 1456-1459, 1477-1480, 1526-1531) were taken in consideration. They showed relatively low mortality during the winter and high mortality in the summer (Morrison et al., 1985).

In addition to being able to carry out measurements in the presence of epidemics, the registration of deaths also allows the socio-economic conditions of a specific area to be assessed. Edwin Chadwick (1800-90) declared that poverty and disease were closely related, and Lemuel Shattuck (1793-1859) reported that living conditions have an important impact on infant and maternal mortality and morbidity (Declich & Carter, 1994).

We have thus chosen to perform spatial and temporal analyses of 1480, a period without war and epidemics, of the *Liber Mortuorum* of Milan to study the distribution of the death events in the city during this period. This source is extraordinary because it is, to our knowledge, the most ancient European registry that contains socio-demographical data at the individual level.

From the spatial and temporal analysis, it is clear that child deaths rule the pattern of the first and the trend of the second.

In the 1480 registry, there is a scarcity of information about the causes of death in children. Statutes only required identification of cause of death for individuals 3 years or older.

As it was written in the introduction, children's death is an important indicator to assess the health condition of an area. Early childhood was a very precarious condition: lack of hygiene, contaminated water, malnutrition, and other states led, for example, to worms, diarrhea, and smallpox.

The children of elite Milanese families were destined for a future of marriage, religious life, or conduits; the destiny of non-noble children depended strictly on the family's finances, which were connected to the physical environment, harvesting productivity, and the job market. The nutrition of infants has a crucial role in the growth and protection from illness: breast milk contains antibodies and many nutrients, but poor and malnourished mothers often did not have enough milk to feed the baby, and it is possible that a mother died in childbirth or the puerperium (Fass, 2013). From the distribution of the age of children, it can be seen that at least half of them were less than a year old.

The fact that the spatial clusters for all subjects and for children were the same possibly reflects that in the *Contrade*, there were poor neighborhoods (or wealthy neighborhoods for the low-low clusters), or it is possible that the areas were characterized either by a high or a low population density. Temporal autocorrelation can occur with illnesses affected by seasonality. For example, many respiratory diseases arise in spring and winter, and diarrheal disorders occur in summer and autumn. It is also possible that unknown conditions are connected with the increase of the death events over time.

Regarding the heatmaps, there is a significant peak of deaths in *Contrada* 21 for April for all subjects and adults, so the mortality of the adults may rule this peak. Still, in October, there was a peak for children and all subjects, so it is possible that a particular condition affected children in this area in this month. *Contrada* 11 had a high peak in August for all subjects and children. The study of death events in a period without epidemics can be critical in assessing the impact of the plague on the city. For example, it is possible to analyze excess mortality by age, by comparing the distribution of the age of death in plague or non-plague periods and sex and socioeconomic conditions. Studies generally show that there is not a relevant difference in the sex distribution of plague victims. Only some differences were found for the fifteenth century in Milan. However, it is possible that it was due to the poor condition of migrant women and widows. Furthermore, the event of dying concerning socioeconomic factors changed over time: during the Black Death, the plague was recognized as a universal killer. During the fifteenth century, poor people were primarily involved; finally, during the seventeenth century, the rich people (Alfani & Murphy, 2017).

PUBLISHED PAPER ON THIS PART OF THE PH.D. RESEARCH PROJECT:

- Luconi E., Boracchi P., Nodari R., Comandatore F., Marano G., Vaglienti F., Galli M., Biganzoli E. Spatial and Temporal Analyses of the Event of Death for 1480 in Milan Using the Data Contained in the Sforza's Registers of the Dead. Int J Environ Res Public Health. 2023 Feb 4;20(4):2783. doi: 10.3390/ijerph20042783. PMID: 36833481; PMCID: PMC9956338.

-Biganzoli E., Luconi E., Boracchi P., Nodari R., Comandatore F., Ferrara A., Castaldi S., Vaglienti F., Panella C., Galli M., The Milano Sforza Registers of the Dead: Health Policies in Italian Renaissance,
medRxiv 2021.02.10.20249093; doi:https://doi.org/10.1101/2021.02.10.20249093

4 THE PLAGUE OF 1630 IN MILAN: DID THE PROCESSION WITH THE BODY OF SAN CARLO PLAY A ROLE IN SPREADING THE EPIDEMIC?

"The plague arrived at its height towards the end of August. It is impossible to describe the wretchedness which reigned throughout Milan at this period. The weather was close, the air heavy, everywhere loaded with an equally dispersed fog, which appeared to obscure the sun without promising rain. The country around was uncultivated and arid; not even a drop of dew refreshed the faded verdure and the withered and languishing leaves. A dead silence reigned through the streets, unbroken, except by the melancholy sound of the bell carried by the *apparitori*, announcing the approach of the dead cart, heaped full of corpses" (Donovan, 1849)

4.1 Introduction

The plague that hit northern Italy in 1630 left a profound mark on the country's demography, economy, and culture and was a significant factor leading to the country's economic decline in the XVII century (Alfani & Percoco, 2019). It has been estimated that there were about two million victims, indicating a mortality rate of 30-35% (Alfani, 2013), with an estimated median rate of as high as 40% in the 26 largest cities of northern Italy (Alfani & Percoco, 2019). At that time, Milan was the capital of a vast duchy under Spanish rule and was greatly affected (Ripamonti, 1640).

During 1628 in France and Northern Europe, there were some plague outbreaks, but the movement of the troops in October and December for the war of Monferrato led to an epidemic. In addition, in this period, famine caused both an increase of beggars in the cities and led to famous bread riots. Scenes of death by starvation characterized the first six months of 1629, and the sending of the poor to the *Lazaretto* of the town condemned them to an ever more complicated situation. From July to September, foreign troops crossed the countryside of Milan. In October, a fact also contained in Manzoni's novel (the *Promessi Sposi* (The Betrothed)), Pietro Antonio Lovato, who lived in Porta Orientale, San Babila parish, returned to the city from Lecco or Chiavenna, carrying clothes from soldiers. He felt ill and died a few days later. He was considered the "patient zero" of the plague in Milan. On 31 October, two others died of plague in *Porta Orientale* district, and others in *Ospedale Maggiore*. After these deaths, there were no new cases until 22 November. Festivals characterized this month to celebrate the birth of the new heir to the Spanish throne: Carlo Domenico Baldassare. During December, there were new cases of plague but not new outbreaks.

In January 1630, the usual carnival celebrations took place. Fifty people were in the *Lazzaretto*, but only three deaths were recorded, reassuring the population that the city was "in health". Only on 30 January, the first "suspected case" of plague was recorded in the *Liber Mortuorum*, a young girl of 13 years old. In February 1630, celebrations about the carnival and the new heir continued. Only in March did the situation get complicated; by 11 April, there were four or five cases daily. The situation worsened in May, with the highest number of deaths, especially in Porta Orientale. In June, in only one night, more than 72 deaths were reported (Farinelli, 1988). At the end of March 1630, Milan's *Lazzaretto* held fewer than 300 people (La Croce, 1730). However, by June, that number reached thousands [8],and the health authority had also closed 300-500 'infected' private houses (Della Somaglia, 1653). On 11 June, a solemn procession took place (the San Carlo procession), proceded by two other processions: one on 9 June *(processione delle Torce)* and one on 10 June (*processione delle Arti*) (Farinelli, 1988).

Religious processions during times of plague took place all over Europe because illness was seen as a punishment from God for sins, so these rites were occasions to ask God for forgiveness and mercy (Chiu, 2018). The procession of the Body of San Carlo involved so many people that it is not so hard to believe that all the population of Milan took part in it, except people who were in closed houses (thus, houses in which plague cases occurred) or in the *Lazzaretto*. The citizens of Milan walked through the streets in the city center to participate in the solemn religious procession accompanying the body of San Carlo Borromeo (Farinelli, 1988). There were strict rules, often posted at the gates or nailed at the door of closed houses, about the entrance to the city to decrease the contagion risk (Donovan, 1849).

Two of the most important chroniclers of the time, Ripamonti and Tadino, affirmed that this event played a role in the marked increase in the number of plague cases: "*...eadem ab motis D. Caroli Reliquijs, & solicitato veluti Coelo terribilior multò violentiorque facta fuerit....*" (once the relics of San Carlo were moved and Heaven was solicited, the [plague] became much more terrible and violent), *Ripamonti* (*page 66* [9]); "*Finite queste solennità molto più si accese il fuoco della peste in*

---

[8] Archive of the State of Milan, Sanità 'ancient part', 286: letter from the grand chancellor Antonio Ferrer to the governor Ambrogio Spinola dated 2 June 1630.

[9] G. Ripamonti, *Iosephi Ripamontii cononici scalensis chronistae urbis mediolani de peste quae fuit anno MDCXXX: Libri V : desumpti ex annalibus urbis quos LX. decurionum autoritate scribebat* (Apud Malatestas, Regios ac Ducales Typoraphos, 1640).

tutte le parti della Città" (At the end of these solemnities, the fire of the plague flared even more throughout the city.), *Tadino page 108* (Tadino, 1648).

In July 1630, there were at least 500 daily deaths; in the meantime, some problematic episodes hit the city of Milan: an actual fire on the night of 22 July was followed, on the subsequent day, by a hefty downpour that caused many deaths in the *Lazzaretto*. Another fire broke out on 24 July. In August, there were thousands of deaths every day: on 31 August, this number reached two thousand. From September, the number of deaths began to decrease. Finally, the city was proclaimed free from the plague in February 1632 (Farinelli, 1988). No epidemiological reconstruction of this epidemic has been based on accurate epidemiological records. However, Milan's *Libri Mortuorum* offers the opportunity to study its spatiotemporal dynamics. This chapter of the thesis describes the first study's results based on accurate data concerning the impact of the San Carlo Borromeo procession on the epidemic.

4.2 Materials and methods

The three registers of 1630 of the *Mortuorum Libri*, with the related socio-demographic, geographical, and causes of death information, were consulted and digitized. In the State Archive of Milan, where the registers are kept, it was impossible to find the one with the data from 4 August 1630 to 31 August 1630, so the data still needed to be included. The registers were translated from Latin by an expert paleographer, Dott. Luca Fois; moreover, Dott. Riccardo Nodari, and Dott. Francesco Comandatore organized the data in a spreadsheet to conduct the analysis.

The recorded deaths were subdivided into those attributed to plague and those reported to other causes, according to the judgment of the expert on infectious disease, Prof. Massimo Galli, who analyzed the causes of death and attributed them into one of the two categories based on the description: in case of plague deaths, signs and symptoms are often reported. The subdivisions were used to reconstruct the epidemiological histogram of the number of daily deaths (August, the period without the data, was masked) in the graphical representation. For this analysis, the dataset was organized as follows. Each row contains individually reported information with the date of the event and information about whether the death was caused by plague or not. Information about the parish where the event happened was detailed in each case.

Clusters were made based on temporal dynamics according to the epidemic's progression, parish by parish, then expressed as the cumulative distribution of the number of plague deaths in that location. Data were organized as a matrix, with each column being a parish and each row being a day in which plague death happened. Days without events were also added to the matrix. For each parish, the data have been sorted by date, and the cumulative frequency distribution was calculated as presented:

number of death events in each day (in a parish) divided by the total number of deaths in the parish from January to December 1630. The analysis considered only parishes with at least one death every two weeks. The threshold (21 deaths) was calculated as the weeks of epidemics / 2. To determine the duration of the epidemic period, it is not appropriate to consider the date of the first victim of the plague and the last one because these may be isolated cases. So, the duration of the epidemic was determined according to the epidemic curves: the epidemic began around the middle of March and continued throughout 1630. Graphical representations of the death cumulative frequency curves ("plague cumulative curves") were performed, excluding parishes that did not reach the threshold.

The cumulative frequency curves were used to compute the distance matrix using Euclidean distance. The obtained matrix was used to perform the Principal Coordinates Analysis (PCA) (Gower, 1966): a method to create mappings of items based on distance. The PCA is a subtype of Multidimensional Scaling that deals with numerical distance. Like the Principal Components Analysis, the method reduces the dimensionality of the data, but the components are based on distance (not on variance). The Torgerson method was applied, and the matrix was double centered with those steps (Korstanje, 2021):

1) Distances were squared

2) Calculation of a centering matrix:

$C=I- (1/n) J_n$

I is the identity matrix (or unit matrix) that contains 1 in the diagonal and zero for other positions, n is the number of observations, and J is the original matrix distance. After that, we can compute the final matrix as:

3) $B= -1/2\, C D^{(2)} C$

4) Singular Value Decomposition is performed on matrix B.

The best number of axes for the PCA analysis was chosen according to the scree plot that is based on the eigenvalues.

The silhouette analysis was performed to select the appropriate number of clusters: this analysis is based on the separation distance between clusters. For each point in each cluster, S(i) is calculated as follows:

$S(i)= (b(i)- a(i)) /\max (a(i), b(i))$.

b (i) is the mean distance between point I and all other data points in the same clusters, a (I) is the smallest mean distance of the point I to all point in any other cluster.

The mean of S(i) for points in each cluster is considered to choose to select the appropriate number of clusters. If the value is near 1, there is a high separation between clusters, 0 is a sign of overlapping, and -1 value denote the worst decision.

After selecting the number of clusters, according to silhouette analysis, parishes were grouped using the unsupervised clustering algorithm K-means (Hartigan & Wong, 1979). In short, the algorithm is: a) random picking a point (a row in the data) as centers (centroid) of clusters (those centers cannot coincide) b) the distances of each point in the dataset with respect to the centroids is computed. c) the points are aggregated to the cluster based on the smallest distance d) the centroid of each cluster is reattributed based on the average distance of all the points belonging to the cluster e) the algorithm starts over point b and stops when there are no changes in the cluster compositions.

Permanova (Permutation Multivariate Analysis of Variance) test (Anderson, 2001) was used to assess the clusters' significance; the null hypothesis of this test is that the centroids and the dispersion of the clusters are equivalent for all groups. This test has no assumptions on normality or homoscedasticity, but the only assumption is that the observations are interchangeable under the null hypothesis. The matrix of Euclidean distances is required and, from this matrix, the distance of points between groups and within groups is considered, the Pseudo F statistic is calculated, and the distribution of the F statistic is calculated thanks to permutation.

So, the parishes, and therefore our cases, have been clustered based on the temporal progression of the epidemic. Now we can analyze the differences and similarities between the clusters.

The Mann-Whitney U test was applied on the dates that correspond to 25% and 50% of total plague deaths, the dates of the first plague deaths, and the date of the epidemic peak (highlighted by the changes of the concavity of the cumulative curve, the inflection point). This procedure was made to compare the trend of plague deaths in different clusters. Moreover, to make it easier to compare visually the differences in plague deaths over time of each cluster, a histogram of the weekly plague deaths was produced. Considering also the nonplague deaths in each cluster, the Mann-Whitney U test was used to compare all the deaths in them. This test is nonparametric and assesses if the two sampled groups are likely to derive from the same population in terms of the shape of the distribution of their data.

The longitude and latitude of each parish were determined using Quantum GIS software (Qgis Developer Team, 2021) and the position of the related church as shown in the digitized historical map of Milan dated 1629 (Blaeu, 1704). The positions of Parishes were also estimated according to a historical text (Rotta, 1891). Parishes were plotted on the historical map and colored according to the membership cluster to visualize the geographical distribution of the clusters. Then, the distance of each parish from the city center (defined as the centroid of the area delimited by the city's medieval walls) was computed using QGis.

4.3 Results

The digitization of the registers during the period of interest made it possible to get 8152 records, of which 5261 (64,5%) were reported as having been caused by plague. According to the diffusion of the epidemic in the city, the number of plague deaths in each parish varied widely, with San Stefano in Brolo having the highest number (592).

It can seem that the number of deaths due to plague in the register is low, considering that Milan's population in that period was around 130.000. However, it must be considered that the *Lazzaretto* register was not included in the analysis, and also the lost records in August (a volume covers the time of a month, then an exceptionally high number of people died in that period).

Figure 13 shows the daily incidence of recorded deaths. The absence of records between 4 and 30 August is due to the lack of the register reporting the deaths in that period.

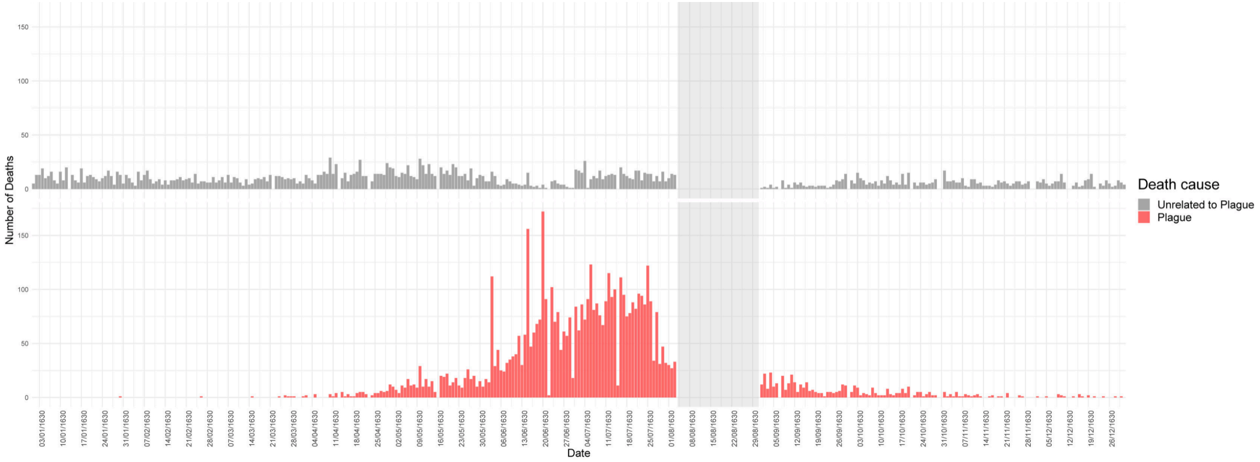


Figure 13: Barplots of the number of deaths per day in the city of Milan in 1630
The daily number of total plague deaths is shown in the red histogram at the bottom, while the daily number of those unrelated to plague is in the gray histogram at the top. The shaded area (4–30 August) indicates the period for which the death register is unavailable.

Figure 14 shows the cumulative number of parishes with one plague death between 1 January 1630 and 4 August 1630 with the indication of the time of the procession of San Carlo (11 June).

Interestingly, the epidemiological situation worsened dramatically after this event: for 63 out of 94 parishes (67%), the first plague event happened after the procession.

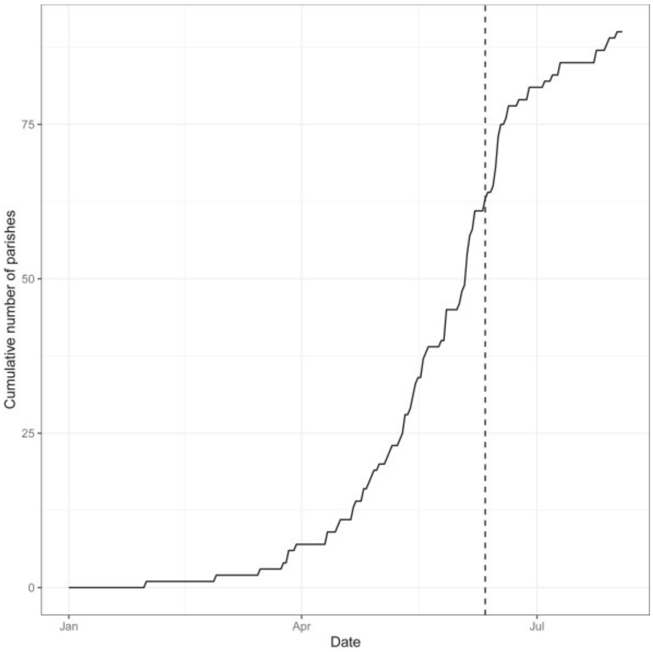


Figure 14: Cumulative number of parishes with one plague death over time.
For each day between 1 January 1630 and 4 August 1630, the number of parishes with at least one plague death is plotted. The vertical dashed line refers to 11 June 1630, the date of the San Carlo procession.

There were 43 parishes with more than one death over two weeks (with more than 21 recorded deaths), and they were grouped according to their cumulative epidemiological curves (Figure 15A). The silhouette analyses allowed us to choose the optimal number of clusters: 2. The K-means subdivided 14 parishes in cluster 1 and 29 in cluster 2. Permanova test confirmed the significance of the obtained clusters (p-value <0,001) (Figure 15B).

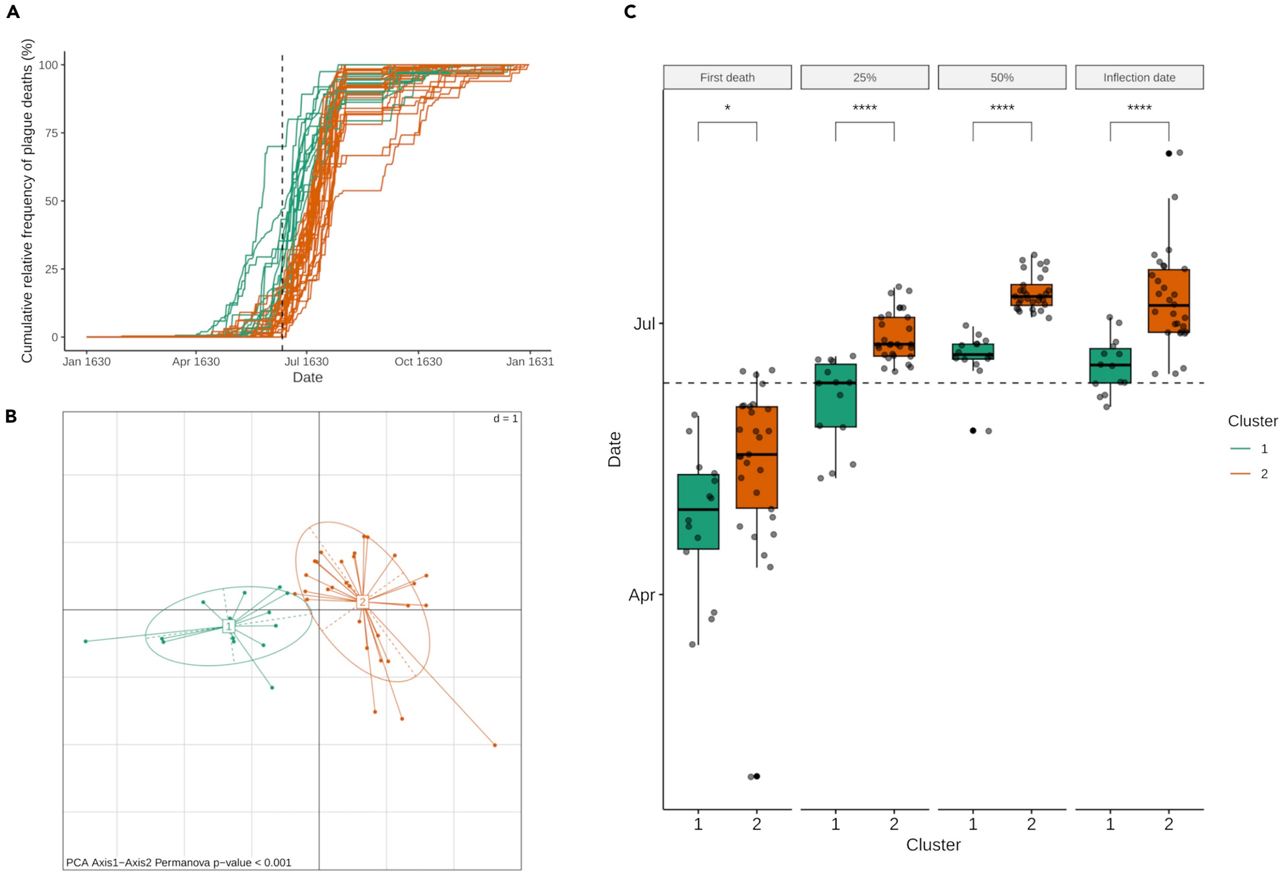


Figure 15: Clustering of the parishes on plague cumulative relative frequency curves. Analysis of the cumulative relative frequency curves of the 43 parishes with, on average, at least one plague death every two weeks of the epidemic (i.e., 21 total plague deaths). The curves were subjected to clustering analysis and principal coordinates analysis (PCA) which identified two clusters (labeled "Cluster 1" and "Cluster 2"). In all the plots, cluster 1 is colored in green, and cluster 2 in orange. (A) Plague cumulative relative frequency curves of the 43 parishes are colored based on the cluster: on the x-axis, the date, and, on the y axes, the cumulative number of plague deaths. The dotted gray line indicates the day of the San Carlo procession (11 June 1630). (B) PCA using the Euclidean distance on the plague cumulative relative frequency curves. (C) From left to right, boxplots of the dates for which the parishes of the two clusters reached the first plague death, 25% of total plague deaths, 50% of total plague deaths, and the inflection point of the curves (i.e., the date at which the cumulative curve changes concavity, corresponding to the epidemic peak). In each boxplot, the dates for the two clusters were compared using a Mann-Whitney U test ('*' p-value <0.05; '****' p-value <0.0001). The boundaries of the boxplot whiskers are based on the 1.5 Interquartile range (IQR) value. The dotted line indicates the date of the San Carlo procession. Figure 3C shows the comparisons of parameters relative to the epidemiological curves of the parishes of the two clusters. The parishes of cluster 1 experienced the first plague death significantly earlier than those of cluster 2 (Mann-Whitney U test; median cluster 1: 1630-04-29, median cluster 2: 1630-05-18; p value <0.05). Similarly, the parishes of cluster 1 reached 25% (median cluster 1: 1630-06-11, median cluster 2: 1630-06-24) and 50% (median cluster 1: 1630-06-20, median cluster 2: 1630-07-10) of their total plague deaths earlier than the parishes of cluster 2 (Mann-Whitney U test; p value <0.0001). Lastly, the parishes of the two clusters significantly differ also for the date of the inflection point (the date at which the curve changes concavity) of their epidemiological curves (Mann-Whitney U test; median cluster 1: 1630-06-17, median cluster 2: 1630-07-07; p value <0.0001). Moreover, the two clusters showed no significant difference in both global deaths (Mann-Whitney U test; cluster 1: 2,640, cluster 2: 4,362; p value = 0.39) and total plague deaths (Mann-Whitney U test; cluster 1: 1,625, cluster 2: 2,979; p value = 0.99).

Seventy-nine of the 94 parishes recorded in the registers were successfully geolocalized and accounted for 93% of all the reported plague deaths; the remaining 15 parishes accounted for 0.99%. The parish information was absent in the registers for 6% of all plague deaths.

The position of the 79 georeferenced parishes is shown on Milan's historical map (Figure 16A). The weekly incidences of plague deaths for cluster 1 and cluster 2 parishes are reported in Figures 16B and 16C, respectively. Lastly, the median distance from the city center of the parishes of cluster 1 is not significantly different from those of cluster 2 (one-tailed Mann-Whitney U test; median cluster 1: 790 m, cluster 2: 515 m; p-value=0,12).

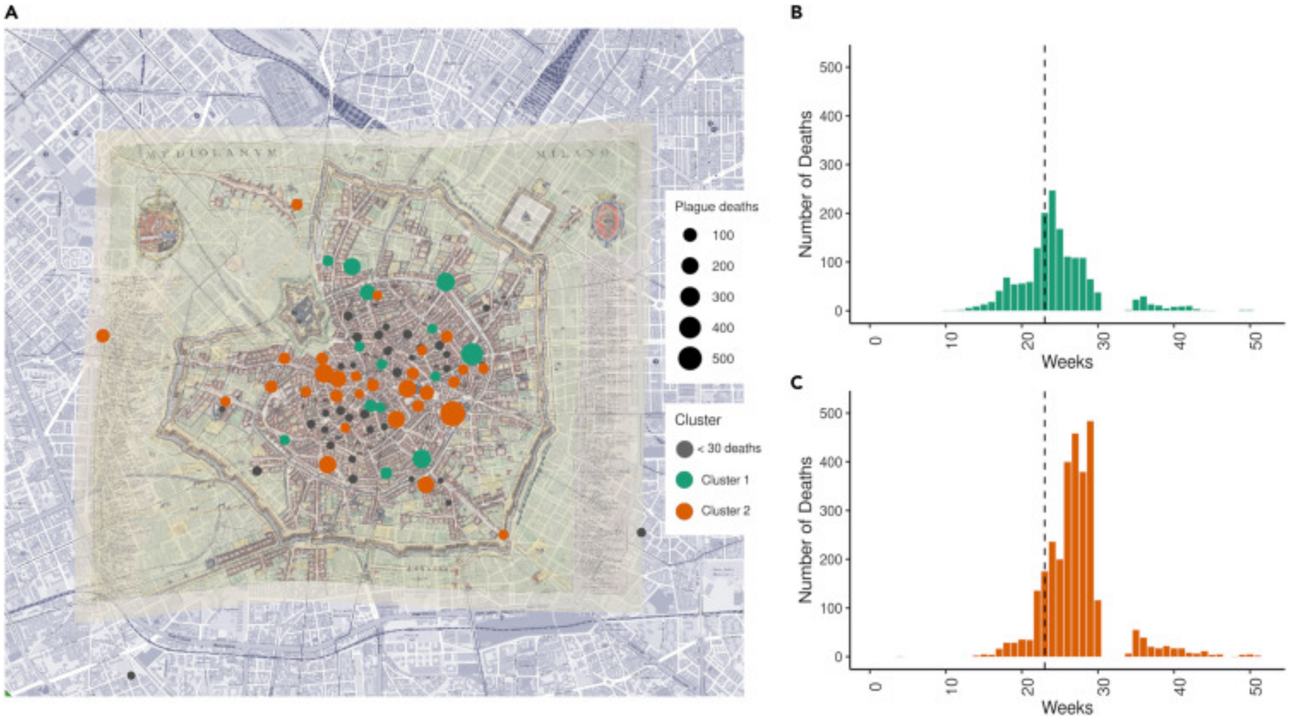


Figure 16: Clusters spatiotemporal distribution. (A) Parish localization on a historical 1629 Milan map. In green are parishes of cluster 1; in orange are those of cluster 2; in gray are the parishes with less than 21 total plague deaths (not used in the clustering analysis). (B) Total weekly plague deaths of cluster 1 parishes; the dashed line shows the procession date (11 June 1630). (C) Total weekly plague deaths of cluster 2 parishes; the dashed line shows the procession date (11 June 1630).

4.4 Discussion

In 1630, northern Italy was the scene of the War of Succession of Mantua, an episode in the Thirty Years' War. The plague was brought to the duchy of Milan by the army that crossed the Alps to besiege Mantua (Farinelli, 1988). However, contemporary chroniclers reported that the spread of the infection remained limited at least until the spring and that the epidemic accelerated at the beginning of summer, during the weeks after the San Carlo procession of 11 June 1630 (Ripamonti, n.d.) (Tadino, 1648). The procession took about 12 hours and involved most of the city inhabitants. It started from Cathedral Square at 7 a.m. (Ripamonti, n.d.), covered all the main streets in the city center, and returned to the cathedral at 7 p.m. For eight days after the procession, the body of San Carlo was placed in public view, attracting the masses who came to pray to the saint (Ripamonti, n.d.). According to the chroniclers, after the procession, the number of plague deaths rose from about 100 to about 1.700 per day (Della Somaglia, 1653).

In this study, we digitized the death registers of Milan for the year 1630 for the first time. These registers contain individual records of more than 8000 individuals who died in Milan in 1630. The dataset obtained allowed us to perform the first quantitative reconstruction of one of the most important plague epidemics in Italian history, also narrated in the masterpiece Italian novel "The Betrothed" of Alessandro Manzoni.

Yersinia pestis is mainly transmitted to humans by the bite of an infected flea through the rat-flea-human transmission chain. The results obtained from our analyses and contemporary chronicles open a biological question: Can a mass gathering event affect the evolution of an epidemic of a vector-borne disease like the plague?

Considering only the rat-flea-human transmission chain, the answer to this question should be "no"; indeed, there is no apparent relationship between the gathering of people and the movement of rats. However, this is not the only possible plague transmission chain. *Y. pestis* can also be transmitted from human to human through the air (i.e., pneumonic plague) or human ectoparasites (Pollitzer, 1954) (Barbieri et al., 2021). The human-to-human transmission chain could explain why a mass gathering event could have affected the evolution of the epidemic.

Pneumonic plague is a rare and highly debilitating plague that affected no more than 2-3% of people during a 19th-century epidemic. It is characterized by a short incubation time, a high mortality rate (about 100% in untreated individuals), and a time from first symptoms to death of 48-72 hours (Lathem et al., 2005). Due to its high contagiousness, a resurgence of the epidemic due to a procession involving many people could be explained by airborne transmission. However, droplet transmission occurs in the final stage of the disease, when patients are so severely ill that their mobility and ability to take part in a long procession are greatly reduced. Furthermore, the debilitating nature of the disease and the noticeable symptoms displayed by infected people (cough, breathing difficulties, bloody sputum (Perry & Fetherston, 1997)) suggest that it was unlikely that people with the pneumonic form of the disease were able or allowed to take part to the gathering, also considering the recommendations made to prevent the participation of infected people to the procession.

As stated above, *Y. pestis* can be transmitted from human to human through human ectoparasites (i.e., lice and fleas). The role of human ectoparasites in plague transmission is currently debated (Barbieri et al., 2021). Still, the role of human ectoparasites in plague transmission has been hypothesized since the beginning of the 20th century (Swellengrebel & Otten, 1914) (Blanc & Baltazard, 1942) (Pollitzer, 1954). This hypothesis received particular interest in the last twenty years after the demonstrations of the ability of fleas and lice to carry and transmit *Y. pestis* from one mammal to another (Ayyadurai et al., 2010) (Houhamdi et al., 2006) (Zhao & Yin, 2016) and the presence of infected human ectoparasites in areas with plague cases (Blanc & Baltazard, 1942) (Drali et al., 2015) (Ratovonjato et al., 2014). The hypothesis of human plague transmission through ectoparasite is also supported by epidemiological (Dean et al., 2018) (Dean et al., 2019), historical (Barbieri et al., 2022; Davis, 1986), and archaeological studies (Hufthammer & Walløe, 2013). Moreover, the bite of an infected ectoparasite gives rise to the bubonic form of the disease, which is more coherent with the temporal dynamic seen in this work. The bubonic form of plague has a time from infection to death/recovery compatible with the split in the dynamics of plague deaths between the two parish clusters about one to two weeks after the San Carlo procession.

A more recent work (Miarinjara et al., 2021) about *Pulex irritants,* a vector that is involved in human plague epidemiology, concerned four experimental conditions with this insect:

1) Fleas infected using human blood and fed sterile human blood daily from day 1-2 after infection (conditions connected with human plague epidemiology).
2) Fleas infected using human blood and subsequently fed sterile human blood every two days.
3) Fleas infected using rat blood and fed sterile rat blood daily from day two after infection.
4) Fleas infected using rat blood and subsequently fed sterile rat blood every 2 to 3 days.

It was found that *Pulex irritants* is an inefficient vector for *Yersinia pestis*, but the transmission event after the first infective blood meal is essential. Therefore, human transmission from those parasites is effective only in the presence of many of them for a host.

The work involved only a strain of *Pulex irritants* from American wild animals: it is possible to find differences in animals from different geographical zones and a human environment.

A study regarding other forms of transmission not directly mediated by lice or other ectoparasites (Jullien et al., 2021) concluded that droplets from the body or manipulated carcasses could cause pneumonic plague. Still, the contact had to be prolonged and very close. In addition, bubonic plague can be transmitted by skin-fluid contact (in particular blood; there is a lack of literature regarding contact from other fluids) and septicemic plague from cuts.

It is highly improbable that that way of transmission is connected with a procession.

It is difficult to identify a social stratification in Milan: 70% of the five Parishes were composed of nobles in the first years of the seventeenth century. However, a more defined social stratification was possible considering streets or narrow city blocks.

Green areas characterized the area between the Spanish and medieval walls with a scattered, high-density populated area. In the sixteenth century, there were almost 25.000 people, and the area beyond the Spanish Walls was less inhabited (10.000 or 20.000 people). The area into the medieval walls was more densely populated with 100.000 people.

Wealthy artisans commercialized in the *Broletto*, *Duomo*, and *Palazzo Ducale* areas, but the zone was also the market area for the poor people. The zone between the *Duomo* and the *Cordusio* was the *Contrada delle Arti*, with specific activities in different streets (for example, goldsmiths, perfumers, or gunsmiths). Textile activities characterized the zone east of the *Duomo*, and were also present in the northern one. In the second, there were also printers and booksellers.

The external zone into the Medieval walls was a residential area: the aristocracy and the wealthy traders preferred to live far from the shops in a calmer place.

There were soldiers with families and poor people in the *Castello* area. The presence of criminals and poor people also characterized it.

The area nearest the medieval walls was highly economically productive (D'Amico, 2012).

Due to the extreme rarity of detailed historical records on past events, the exact number of plague victims for the 1630 epidemic in Milan is still not known. The main limitation of the work presented here is that the registers used for the analysis only reported deaths at home or on city streets. Unfortunately, to our knowledge, the data concerning the deaths in the *Lazzaretto* (the hospital reserved for people suspected to be infected with a different area for infected people), hospitals, and convents were not preserved or are unavailable for the analyzed period. Nevertheless, the data relating to deaths in the city and streets are probably the most suitable for investigating the spatiotemporal dynamics of the epidemic.

In conclusion, this is the first spatiotemporal reconstruction of the plague epidemic in Milan in 1630, one of the most important in Italian history. The results, coherently with the chroniclers of the time, suggest that a significant mass gathering event could have played a role in amplifying the plague epidemic, opening the biological question of whether human-to-human transmission could have had a pivotal role during this epidemic. Further, multidisciplinary investigations are necessary to study this exciting hypothesis.

In this work, we just scratched the surface of the knowledge contained in the death registers of Milan, which cover the period from the first half of the 15th century to the beginning of the 19th century. Studying historical records opens a window into our past, allowing us to understand important information about pre-modern human societies and their pathogens.

PUBLISHED PAPER ON THIS PART OF THE PH.D. RESEARCH PROJECT:

Galli M., Nodari R., Perini M., Luconi E., Fois L., Vaglienti F., Bandi C., Biganzoli E., Comandatore F., “A spatiotemporal reconstruction of the 1630 plague epidemic in Milan." *Iscience* 26.5 (2023), 106704, ISSN 2589-0042, https://doi.org/10.1016/j.isci.2023.106704.

Nodari R., Perini M., Fois L., Sterzi L., Luconi E., Vaglienti F., Bandi C., Biganzoli E., Galli M., Comandatore F., Computational protocol to perform a spatiotemporal reconstruction of an epidemic, STAR Protocols, Volume 4, Issue 4,2023,102548, ISSN 2666-1667,https://doi.org/10.1016/j.xpro.2023.102548.(https://www.sciencedirect.com/science/article/pii/S2666166723005154)

5 THE CLINICAL FEATURES AND THE RISK FACTORS ASSOCIATED WITH DEATHS IN THE PLAGUE OF 1485 IN MILAN

5.1 Introduction

Plague is caused by the gram-negative bacteria *Yersinia pestis*, discovered in 1894 by Alexandre Yersin; it is essentially a rodent illness, and it passes from one rodent to another by fleas (at least 30 species of fleas can be vectors). It is necessary to distinguish two types of rodents: resistant to disease and susceptible. The first does not die due to the disease; thanks to them, the disease can be endemic, but the second dies due to infection. If the rodent host dies, the fleas search for, preferably, another rodent. Still, in case of the absence of rodents, the animal can be established in another warm-blooded species, for example, humans (Duncan & Scott, 2005) (McEvedy, 1988). Animal-to-human transmission can also occur with direct contact: infected animal carcasses, animal bites, and the consumption of infected meat (Barbieri et al., 2020)

Chronicles described symptoms during the Black Death. According to this documentation, it is possible that the Black Death can be ascribed to the bubonic plague. A famous chronicle was Giovanni Boccaccio, an Italian writer and poet, who wrote in his *Decameron* (a collection of short stories), dated between 1349 and 1353, that the plague ''began in both men and women with certain swellings either in the groin or under the armpits, some of which grew to the size of a normal apple and others to the size of an egg (more or less), and the people called them *gavoccioli*'' (Carmichael, 2008).

A Venetian inscription in the cloister of the Church of *Santa Maria della Carità* of 1348 reported that "Alguni spudava sangue p(er) la boca e alguni vegniva glanduxe soto li scaii e a le lençene e alguni vignia lo mal del carbo(n) p(er) le carne e pareva che q(ue)sti mali se piase l'un da l'oltro çoè li sani da li,'nfermi (et) era la çe(n)te i(n) tanto spave(n)to che 'l pare no voleva andar dal fio né 'l fio dal pare e durà q(ue)sta mortalidate cerca mexi VI" (Vaccaro, 2021).

To search the causes of the Black Death, some research on the skeletal remains of the period was carried out in Italy.

Four individuals buried during the Black Death epidemic were studied in the cemetery of the 'ospitale' of San Nicolao (Genova), one of the places of rest for people in the route of Via Francigena.

As a result, the *Yersinia pestis* F1 antigen in three individuals was found in the analysis of bone samples thanks to the Rapid Diagnostic Test (Cesana et al., 2017).

Other research in Venice involved burial locations in the zone of the *Lazzaretto Vecchio* dating to the 14th–16th centuries. This site was used in the Black Death period and subsequent plague epidemics. One hundred seventy-three dental pulp specimens collected from 46 graves were investigated using suicide PCR. The *Bartonella quintana* DNA was identified in five samples: two from the 15th century and three from the 16th century. *Yersinia pestis* DNA was identified in three samples: two from the 14th century and one from the 16th century (Tran et al., 2011).

Outside Italy, researchers study graves in Barcelona (Spain) dated back in 1300-1420, Bolgar City (Russia) dated back 1298-1388, and Ellwangen (Germany) dated back in 1487-1627. One hundred seventy-eight individuals were considered, and 223 DNA extracts from teeth were obtained to study the presence of the *Y. pestis* genomic material through a species-specific quantitative PCR. In addition, there were 53 potentially positive DNA extracts from 32 individuals (Spyrou et al., 2016).

Archival registers can be looked at along with the possibility of studying skeletal remains and retrieving evidence of *Yersinia pestis*,. The parish register of Givry (France), numbered CC74, contains data about the period of the Black Death of 1348. The records contain data about marriages and deaths. At the end of July, the plague begins to claim victims. From August, in the locality, deaths occurred every day: 28 people on 11 August, 32 from 12 August to 21 August, 50 from 22 to 31 August, and from 1 to 11 September ninety-four people die, from 11 to 20 one hundred people die, one hundred and seven from 21 to 30 September. After this month, the number of deaths began to decrease (Gras, 1939).

Although much later, Barcelona death registers, dating from 1457 to 1590, were compiled based on the work of the letter carrier (*correu*), who had to report the number of burials and baptisms in each parish daily. These data made it possible to determine the peak of deaths in plague periods: in 1476, 1483, 1489, 1501, 1520, and 1558, the peak was in June, and in 1457, 1494, and 1515, the peak was in July. For two periods, the peaks were not in summer: in 1530, it was in April, and in 1589, it was in September (Smith, 1936).

Other death registers were available for a subsequent period, for example, the Bills of Mortality of London.

In the last twenty years, the security of the Black Death's etiology has been discussed with both medievalists and scientists (Pobst, 2013).

An exciting work that involved data from the registers is about seasonality: in particular, the time of the peak during the Black Death was between April and October, and before this period (before 1347), there were two peaks in distribution: one in January-February and one in October-November, but recent results, shows that the peak, generally, was between November and April. The authors suggest that the inverse seasonal mortality was due to a variant of the modern *Yersinia pestis* and that, in the past, the movement of people and goods in the warm season allowed a lot of human contact and disease transmission (Welford & Bossak, 2009).

Some facts would make it doubtful that the Black Death was due to the plague (Duncan & Scott, 2005):

- The absence of resistant rodents to plague in Europe.
- The plague spread very rapidly: in less than three years, it covered vast areas of Europe, but the black rat, which is considered the principal host of the epidemic, has a home range of 100 meters.
- The flea's life requires a temperature between 18 and 27 degrees Celsius and humidity of at least 70%. In northern Europe, those conditions are improbable in the territories of the diffusion.
- The plague was recognized to be transmitted from human to human (in fact, the lockdowns and the separation between ill and healthy as preventive measures could prove this), but one person affected by bubonic plague, to our knowledge, cannot transmit the illness directly; only pulmonary plague can be transmitted directly, but this form of plague requires the presence of bubonic plague.
- A study based on data from registers with the application of the susceptible-infectious-recovered (SIR) model of plague transmission with a human ectoparasite vector was compared to the model for pneumonic and rat-flea transmission, and the results suggest that plague transmission occurred through ectoparasites (Dean et al., 2018).

Those arguments can be applied to subsequent plague waves, but an interesting source of information on the etiological agents can be searched in symptoms.

During the Renaissance, ascertainment that a death had occurred after the fourth day of illness was considered sufficient proof to exclude the suspicion of plague.[10] This is because plague signs and symptoms such as an abscess (*Ascesso*), were not enough proof to make a diagnosis of plague. Medicine in the past was on Hippocratic and Galenic bases, according to which this sign is a symptom of the putrefaction of a humor, which might be due to the absorption of poison (in this case it is the manifestation of plague). In this era, different views of causation could be held, one based on a victim's clinical history and the other on postmortem bodily inspection (Carmichael, 2019).

---

[10] Doubtful cases differ from plague. The first are those in which there is the presence of buboes, fever, vomiting, but death does not happen quickly. A case is labeled a plague if death occurs within a few days (two or three on average). Even a death that occurred quickly, but without symptoms, is considered a doubtful case (Samuel K. Cohn and Guido Alfani, 'Households and Plague in Early Modern Italy', *Journal of Interdisciplinary History*, 2007, 38 (2): 177–205).

Interestingly, chronicles of the time rarely report physical signs and the course of the disease, instead describing the territorial diffusion of the epidemic, the number of deaths and the speed of the diffusion, the brevity of the “incubation period” and the limited amount of time between contagion and death (S. Cohn, 2002).

With today’s clinical knowledge, we can tentatively interpret the 15th century evidence of plague cases differently. Some symptoms are characteristic of bubonic plague, a sudden onset in most cases. There is also the presence of a *Carbone*, an abscess with the center black or yellow that becomes an ulcer in two days with the possible presence of cysts around it. The primary *Carbone* occurred at the site of the flea's bite, and the secondary *Carboni* were due to the hematogenous dissemination. The presence of more than one primary *Carbone* is a sign of more than one bite from the infected insect. The *Bubbone* appeared to be an agglomeration of lymph nodes in which the bacteria accumulate and can be found more frequently in the inguinal, axillary, and neck (a more ominous outcome associated with the latter position). It could be the presence of one or more secondary *Bubboni*, smaller than the primary one. In addition, there could be the formation of an *Edema* near the swelling. The *Ascesso* could be a product of the changing of the *Bubbone* over time.

Secondary pneumonic plague could be found in 8 to 10% of patients with the bubonic plague, and a cough with phlegm and blood and fever characterizes it. Primary pneumonic plague is very rare (Nikiforov et al., 2016).

The *Mortuorum Liber* of Milano of 1485, a year characterized by plague, contains very detailed clinical information about the victims and many sick individuals, so it can be used to establish the signs and symptoms and the interesting part of the body. In addition, because these registers also contain data about the individuals who survived the plague (and other causes of death), the symptoms can be related to the risk of death. In the registry, for some subjects, it is recorded on the day the symptoms started. Using this information with the date of the recording in the registry, it is possible to study the time distribution between the onset of the signs and the death.

5.2 Methods

The register of 1485 was translated from Latin by an expert paleographer (Dott. Luca Fois), and the information was organized and transferred to a spreadsheet to make analyses.

The records previously identified as repeated in the data entry were checked and excluded from the analysis. Those records could be because there were duplicated records in the source. However, additional control of other repeated records was performed based on all the information.

The age of the subjects can be expressed in years, months, or days (the last two used for children); it was rounded (were necessary) and reported with the same unit (years).

Gender, as previously illustrated, is not included in the registry, but it can be inferred from the subject's name (if it is present and legible).

The textual content was used to derive if the subject survived from January to December or not, as well as the duration of the illness (to death), the presence of plague symptoms, and the interested part of the body.

Table 13 contains the terms searched in the causes of death (this is the field's name on the dataset, but, as presented before, it can also be ascribed as "description of the illness"). They are synonyms of the same conditions (death) or symptoms (such as *Bubbone*) (reported as "Condition or symptom" in the table).

The searched terms are reported in Italian because it is the language of the data transcribed from the registers. Terms were searched in upper case and lower case. It can be noted that the terms were truncated to consider singular, plural, and little variances of them.

Condition or symptom	**Related words searched in the "causes of death"**
Death	"mort" OR "deced"
Ascesso	"ascess", "apostem"
Carbone	"carbon", "pustol", "antrac"
Fever (*Febbre*)	"febbr"
Morbilli	"morbill"
Bubboni	"bubbon", "ghiandol"," tumor", "bubo","tumefazion"

Table 13: Terms searched in the causes of death for constructing the variables of signs and symptoms

Women of childbearing age were from 15 to 49 years old (Peller, 1947).For those subjects, additional information was retrieved (Table 14).

Condition or symptom	**Relate words searched in the “causes of death”**
Aborto (Abortion)	“abort”
Flusso mestruale (menstrual flow)	“mestr”
Incinta (pregnant)	“incint”
Parto (childbirth)	“parto”
Puerperio (puerperium)	“puer”

Table 14: Terms searched in the causes of death for constructing the conditions of women in childbearing age.

This first analysis was performed using the Knime Analytic Platform (Berthold et al., 2008), and in particular, using the String Manipulation Node. This node searches for a term in a string and gives the position of it, if it is present. The procedure was performed for each term (Related words searched in the “causes of death” provided in Tables 13 and 14), and the output was the presence of signs and symptoms of plague for each subject, and if the subject died during 1485 according to the *Liber Mortuorum*. Columns were added for each condition or symptom in the database. For each subject, the category of those variables was codified as “Yes” or “No” (presence or absence).

Subjects without signs and symptoms might have died due to the plague, but for some people, the causes could be very different (for example, childbirth). In a situation where it is impossible to evaluate the case, it was decided that, in the absence of symptoms, the death or the illness was probably not due to plague. It was challenging to evaluate the signs and symptoms for the subjects with the wording “similar symptoms” or “same symptoms” (it is not clear with respect to which subject in the registry this indication is referred to). Those subjects were considered to have no symptoms, so they probably were no plague cases.

An age pyramid for death and survived people according to age and gender is provided. Trends of the deaths for all subjects and the percentages of plague deaths referred to the number of registered subjects are provided.

A multiple component analysis (MCA) on plague signs and symptoms was performed to analyze the association between them.

Logistic regressions were performed to investigate the association between the death (response variable) and gender, age class, plague signs and symptoms, and conditions in women of childbearing age (explanatory variables). Results were expressed as odds ratios (accompanied by 95% confidence intervals and p-values). Only univariate analysis was performed on this topic. Only the presence and the absence of each condition were considered (two categories). Plague cases recorded in the

*Mortuorum Lib*er are part of all infected subjects in the city because they do not include the dead inside the *Lazzaretto* (probably, in the considered period, it was the area in *Domus Montanee* (now the zone of the University of Milan, in Via Festa del Perdono (Zanoboni, 2020) and the hospitals.

Moreover, for each symptom of plague (excluding fever), a study of the part of the body where it was manifested was performed. The parts of the body searched in this analysis are reported in Table 15.

It can be noted that the terms were truncated to consider singular, plural, and little variances of terms.

Part of the body	**Relate words searched in the "causes of death"**
Cervello (top of the head)	"cervello"
Coscia (thigh)	"cosc"
Ascella (armpit)	"ascell"
Inguine (groin)	"inguin"
Collo (neck)	"collo"
Orecchio (ear)	"orecchi"
Bocca (mouth)	"bocca"
Mento (chin)	"mento"
Spalla (shouldes)	"spall"
Gola (throat)	"gola"
Petto (chest)	"petto"
Mascella (maxilla)	"mascell"
Mandibola (mandible)	"mandibol"
Omero (humerus)	"omero"
Braccia (arm)	"bracci"
Natica (buttock)	"natic"
Ventre (belly)	"ventre"
Ginocchio (knee)	"ginocch"
Tibia (shinbone)	"tibi"
Mammella (breast)	"mammell"
Gomito (elbow)	"gomit"

Table 15: position of the symptoms searched in the field "Causes of death" for subsequent analysis.

The order of the words was considered to associate the symptoms with the part of the body. For example: " *deceduto per ascesso pestifero dietro l'orecchio sinistro con morbilli pestiferi*" (deceased with a pestiferous abscess behind the left ear, with pestilential measles). The order of the words of interest was: "*ascesso*", "*orecchio*" and "*morbilli*". The *Ascesso* is in the ear area; it is unclear whether the *morbilli* are in the same area.

We analyzed the length of time between symptoms and death. The date of the event was considered the date indicated in the register (the day of the registration): the day of the week (expressed as a number) of it was obtained thanks to the Knime node "Extract Date&Time". The String Manipulation Node, as described before, was used to extract the information about the day of the manifestation of the symptoms in the string of the cause of death. In particular, the names of the day of the week and the wording "*ieri*" (yesterday) were searched. The days of the first observation, occasionally, were also reported in a numerical form: for example, the "*Prima*" referred to the first day of the week; this search requested caution because it can be referred to the hour of the day or other topic, so a check was performed.

It was possible that there was more than one indication about the day in the causes of death, and those records were evaluated individually. Only the records of dead subjects were considered for this analysis.

The indication about the day in the field about the causes of death was translated into a numerical form: in particular, Mondays were considered as 1 and so on until Sundays (considered as 7). The term "Yesterday" was considered as 1. The length of time between the event and the starting of the symptoms was obtained as the difference between the day retrieved from the date of the registration and the one in the string about the causes of death. If it was negative, a correction was made by adding seven to the result. The same was made if the difference was equal to zero. The median of the distribution was considered.

5.3 Results

The spreadsheet consists of 4334 records. After eliminating duplicated records, 4329 subjects were considered in the analysis, and their information was in register 77/1 (Fondo: Atti di Governo di Polazione parte Antica). Of the 4310 whose gender was specified, 2391 were females (55,5%), and 1919 males (44.5%). The median age of the 3997 cases in which this data was available was 22 years (limits 0 - 100 years, IQR 13-40); the mean was 26.63 (s.d. 17,26).

Of the cases reported, 1676 were found to have died at the first observation, or after some time from it, while of the other 2653, whose deaths were not reported during the year in the *Mortuorum Liber*, it is presumable that they survived the epidemic and, as such, were considered for the analysis.

Figure 17 shows the distribution of deaths and the number of subjects recorded in the registry according to age and gender.

Figure 18 shows the number of subjects recorded in the registry per week at their entry into the register. They are divided between alive and dead.

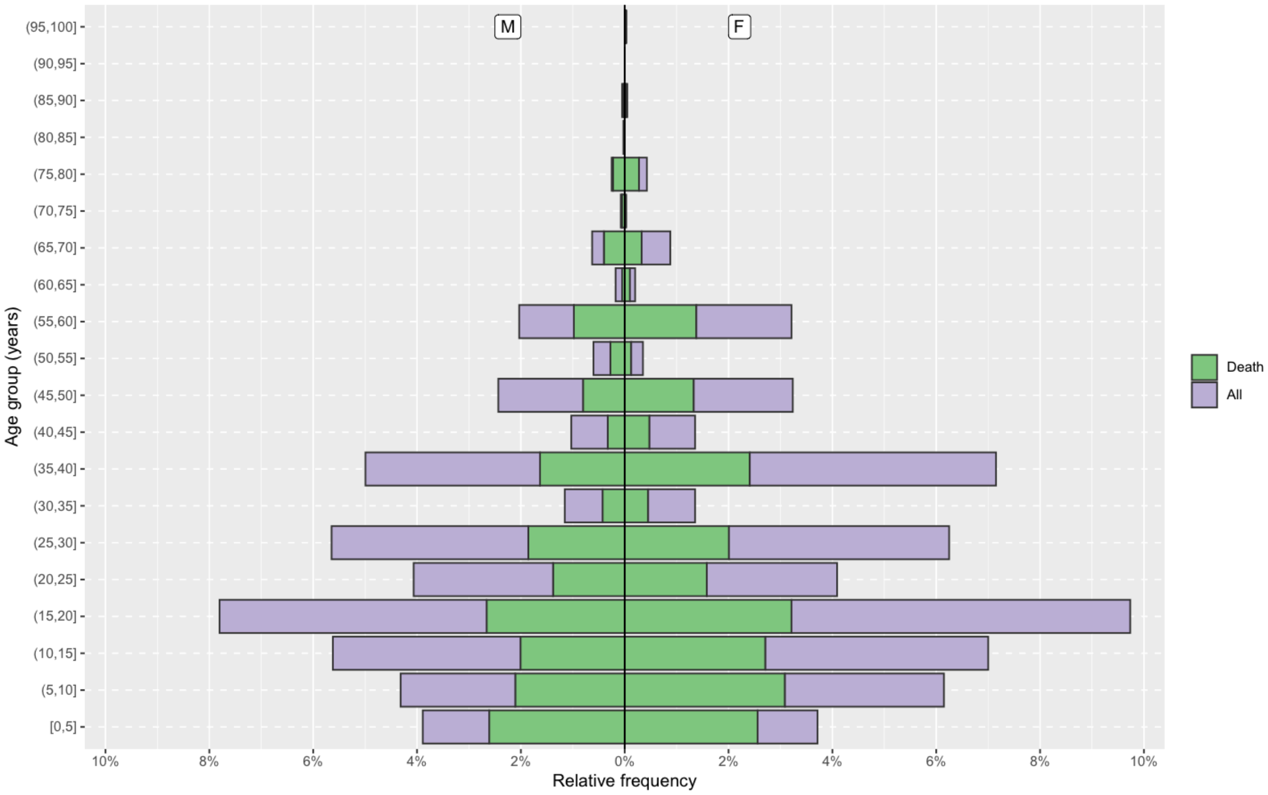


Figure 17: Distribution of deaths and the number of subjects recorded in the registry by age and gender. The violet bar represents the number of subjects. The green bar represented those who died for each category. Only subjects for which age and gender are reported were considered for the analysis (3985).

It can be observed that the highest frequency of age class recorded in the registry for males and females was the (15,20] one. The highest frequency of deaths for males was in the class (15,20] and [0,5]; for females was in the class (15,20], and (5,10].

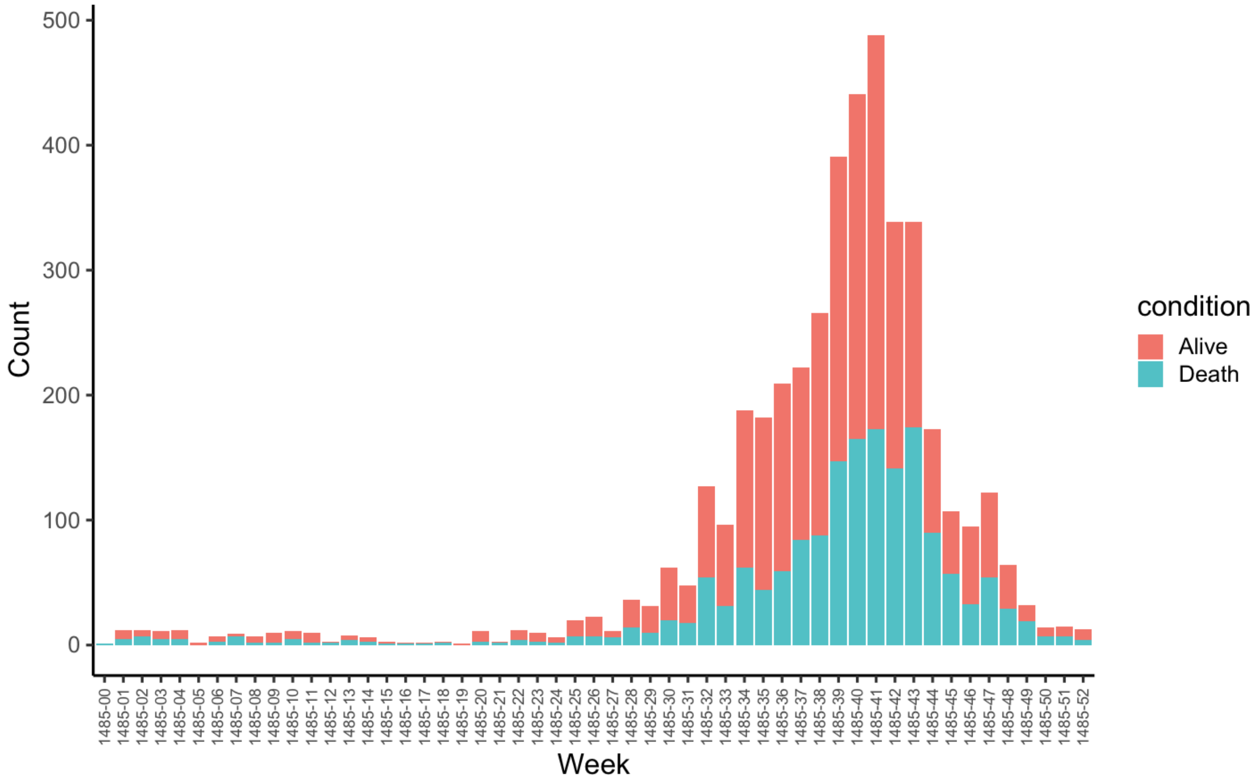


Figure 18: Number of subjects recorded per week at their entry in the register. They are divided between alive and dead in the considered year. For one subject, the entry date in the registry was missing.

Figure 19 reports the percentage of deaths among registered subjects each week.

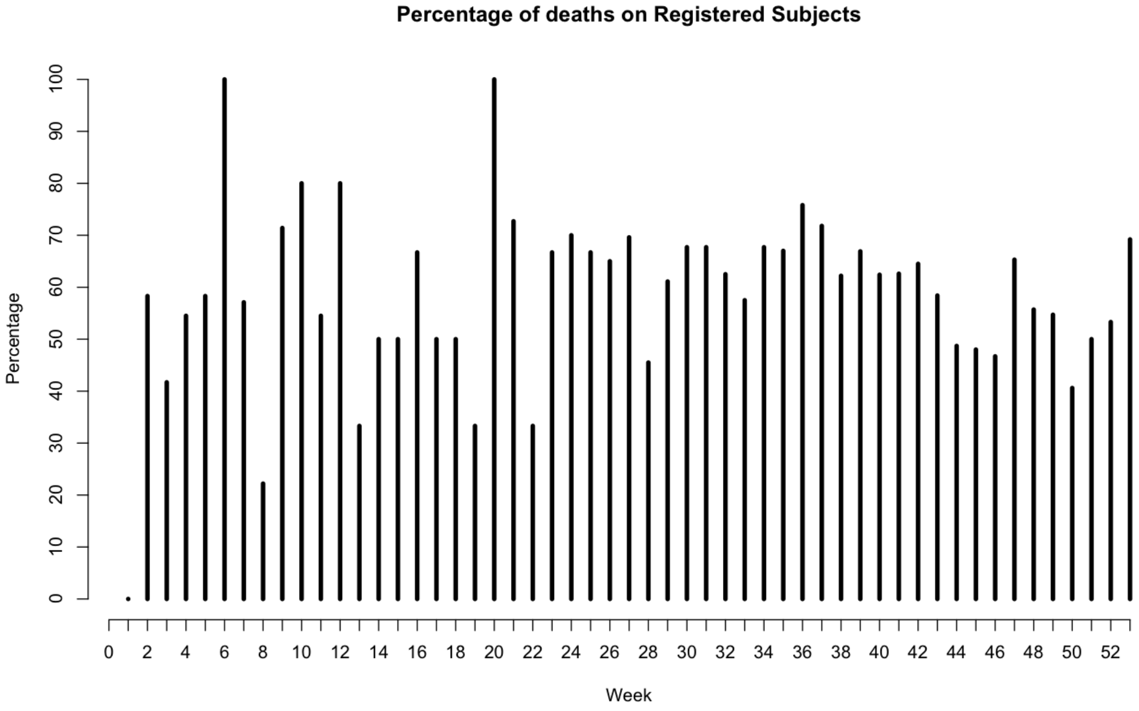


Figure 19: percentage of deaths on the Registered subjects in each week

In weeks number six and 20, all the entries in the register were about death subjects.

Table 16 presents the number of recorded subjects in 1485 each month (the first register date was 2/1/1485, and the last was 31/12/1485) and the number of deaths. For one subject, the date of the death was missing.

Month	Number of subjects	Number of deaths
1	48	23
2	25	12
3	39	13
4	16	8
5	18	7
6	58	19
7	148	52
8	519	183
9	1055	340
10	1767	719
11	524	243
12	111	57

Table 16: Number of recorded subjects and deaths in each month of 1485.

The highest number of recorded subjects was in October (1767). In August, there were 519; in September, 1055; and in November, 524. The highest number of deaths was in October (719). In August, there were 183 deaths. In September, there were 340, and in November, there were 243 deaths. The most recorded subjects in the registry were around September, October, and November. The highest number of deaths were also in the same months.

Table 17 reported the characteristics of all the subjects, subjects that did not die in the year, and those who died.

	All subjects		Alive		Death	
Variable	**N**	**%**	**N**	**%**	**N**	**%**
SESSO (GENDER)	4310		2639		1671	
... F	2391	55.5%	1457	55.2%	934	55.9%
... M	1919	44.5%	1182	44.8%	737	44.1%
ASCESSO	4329		2653		1676	
... No	3691	85.3%	2263	85.3%	1428	85.2%
... Yes	638	14.7%	390	14.7%	248	14.8%
BUBBONE	4329		2653		1676	
... No	2161	49.9%	1098	41.4%	1063	63.4%
... Yes	2168	50.1%	1555	58.6%	613	36.6%
CARBONE	4329		2653		1676	
... No	3664	84.6%	2151	81.1%	1513	90.3%
... Yes	665	15.4%	502	18.9%	163	9.7%
FEBBRE	4329		2653		1676	
... No	3828	88.4%	2230	84.1%	1598	95.3%
... Yes	501	11.6%	423	15.9%	78	4.7%
MORBILLI	4329		2653		1676	
... No	3276	75.7%	2513	94.7%	763	45.5%
... Yes	1053	24.3%	140	5.3%	913	54.5%

Table 17: characteristics of all subjects, those who alive and not in 1485.

Subjects who did not die in 1485 had a median age of 24 years, and those who died had a median age of 20. Women of childbearing age in the registry were 1268 (34.1% died) with a median age of 26 (min: 15 years, max: 48 years). Women of childbearing age who did not died in 1485 had a median age of 26 years and who died had a median age of 25 years.

Table 18 reports the characteristics of all women of childbearing age, those who did not die during 1485, and those who died during the period.

	Women in childbearing age (total)		Alive		Death	
Variable	**N**	**%**	**N**	**%**	**N**	**%**
ASCESSO	1268		835		433	
... No	1084	85.5%	705	84.4%	379	87.5%
... Yes	184	14.5%	130	15.6%	54	12.5%
BUBBONE	1268		835		433	
... No	579	45.7%	315	37.7%	264	61%
... Yes	689	54.3%	520	62.3%	169	39%
CARBONE	1268		835		433	
... No	1100	86.8%	706	84.6%	394	91%
... Yes	168	13.2%	129	15.4%	39	9%
FEBBRE	1268		835		433	
... No	1106	87.2%	691	82.8%	415	95.8%
... Yes	162	12.8%	144	17.2%	18	4.2%
MORBILLI	1268		835		433	
... No	974	76.8%	791	94.7%	183	42.3%
... Yes	294	23.2%	44	5.3%	250	57.7%
ABORTO	1268		835		433	
... No	1238	97.6%	833	99.8%	405	93.5%
... Yes	30	2.4%	2	0.2%	28	6.5%
INCINTA	1268		835		433	
... No	1253	98.8%	830	99.4%	423	97.7%
... Yes	15	1.2%	5	0.6%	10	2.3%
MESTRUAZIONI_	1268		835		433	
... No	1261	99.4%	831	99.5%	430	99.3%
... Yes	7	0.6%	4	0.5%	3	0.7%
PARTO	1268		835		433	
... No	1260	99.4%	835	100%	425	98.2%
... Yes	8	0.6%	0	0%	8	1.8%
PUERPERIO	1268		835		433	
... No	1266	99.8%	834	99.9%	432	99.8%
... Yes	2	0.2%	1	0.1%	1	0.2%

Table 18: characteristics of all women of childbearing age, those who did not die during 1485, and those who died during the period.

For the low frequency of each condition in those women, it was chosen to group all the conditions in one variable that assumes the value “Yes” if almost one condition is present and “No” otherwise (Table 19)

	DEATH		ALIVE	
Variable	**N**	**%**	**N**	**%**
variabile_unica	835		433	
... No	823	98.6%	384	88.7%
... Yes	12	1.4%	49	11.3%

Table 19: variable that groping of the condition in childbearing women (*variabile_unica*)

In the Multiple Component Analysis (MCA), the active variables were the symptoms (five variables), and the passive variables were age classes, gender, and outcome (alive or death).

Figure 20 reports the decomposition of the total inertia in the MCA. The total inertia explains the percentage of variance of the data explained by the dimensions.

	eigenvalue	percentage of variance	cumulative percentage of variance
dim 1	0.29	29.01	29.01
dim 2	0.22	22.12	51.13
dim 3	0.21	21.45	72.57
dim 4	0.21	20.59	93.17
dim 5	0.07	6.83	100.00

Figure 20: decomposition of the total inertia in the MCA.

Based on the decomposition of total inertia and on the number of active variables, it was chosen to consider two dimensions: the explained variability was 51.13% (29.01% in the first dimension and 22.12% in the second dimension).

The percentage of explained variability was relatively low, so it is possible that the symptoms were not highly correlated.

Figure 21 shows the correlation of the variables and the two axes (dimensions) of the MCA:

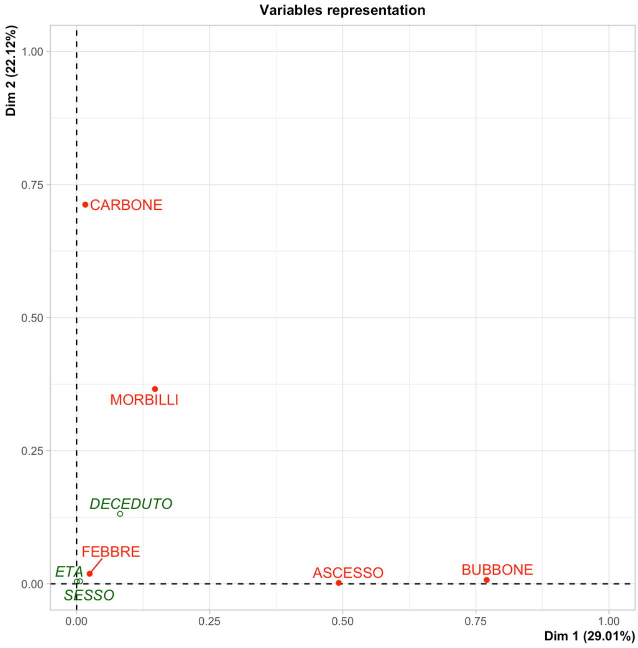


Figure 21: correlation of the variables and the two axes (dimensions) of the MCA.

Ascessi and *Bubboni* are correlated with the first axes: the seconds represent the consequence of the first, so this dimension may represent the post-transmission symptoms by the vector. *Carboni* and *Morbilli* are associated with the second axes, so it is possible that this dimension indicated the signs of the puncture by the carrier.

Febbre (Fever) is near the origin, so it is not associated with a dimension.

Passive variables do not contribute to the axis construction. Still, they are not associated with the axis because they are near the origin: the variable "death" is a little more represented on the second axis.

The cos2 is used to estimate the quality of the representation of the active variables in the two axes; this measure expresses the correlation of the respective points of the variables in the multidimensional space, defined by the dimensions themselves. Table 20 shows the cos2 of each variable for the first and the second dimension:

Variable	First dimension	Second dimension
Ascesso	0.492	0.002
Bubbone	0.770	0.007
Carbone	0.016	0.712
Febbre	0.024	0.019
Morbilli	0.147	0.366

Table 20: cos2 of each active variable for the first and the second dimension.

Table 21 reports the association of the two axes, the contribution of the construction of the two axes (ctr) and the quality of the representation ($\cos^2$) for the categories of the active variables:

Categories	**Dim.1 (first dimension)**	**Ctr (first dimension)**	**cos2 (first dimension)**	**Dim.2 (second dimension)**	**Ctr (second dimension)**	**cos2 (second dimension)**
ASCESSO_No	-0.292	5.003	0.492	-0.017	0.022	0.002
ASCESSO_Yes	1.688	28.941	0.492	0.098	0.127	0.002
BUBBONE_No	0.879	26.596	0.770	0.085	0.326	0.007
BUBBONE_Yes	-0.876	26.511	0.770	-0.085	0.325	0.007
CARBONE_No	-0.054	0.172	0.016	-0.360	9.893	0.712
CARBONE_Yes	0.299	0.948	0.016	1.981	54.508	0.712
FEBBRE_No	0.057	0.195	0.024	0.050	0.198	0.019
FEBBRE_Yes	-0.432	1.487	0.024	-0.380	1.511	0.019
MORBILLI_No	-0.217	2.468	0.147	0.343	8.049	0.366
MORBILLI_Yes	0.677	7.679	0.147	-1.067	25.040	0.366

Table 21: association of the category of the active variables with the two dimensions.

To obtain the overall quality of the representation for each variable category, the sum of the squared cosine of the two axes of one of the categories (because of the different categories of the same variable) was considered. If the result of this sum was at least 0.5, the categories of the variables were well represented. Table 22 represents the sum of the cosine of each category of each active variable.

Variable	Sum of the quality of the representation in the two axes
Ascesso	0.494
Bubbone	0.777
Carbone	0.728
Febbre	0.043
Morbilli	0.513

Table 22: sum of the cosine of the categories of each active variable.

Ascesso and *Fever* are not well represented, concerning the other active variables. It can be seen also in the MCA factor map in Figure 22.

The presence of a *Carbone* is not associated with another category. So, people with this sign of plague may not have other signs and symptoms. This is the same for the presence of *Morbilli*.

The absence of a *Carbone* has a weak association with the presence of fever (*Febbre*). The presence of a *Bubbone* does not seem to be associated with the presence of other symptoms.

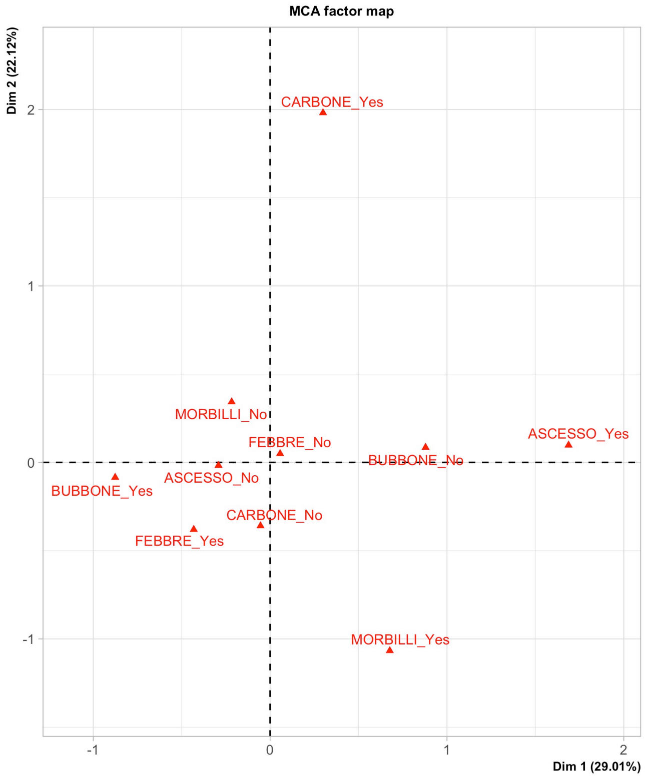


Figure 22: MCA factor map of categories of active variables.

Table 23 presents the quality of the representation of the passive variables:

Passive variables	First dimension	Second dimension
Sex	0.000	0.004
Age	0.006	0.005
Deceased	0.082	0.131

Table 23: quality of the representation of the passive variables

The quality of the representation for all the passive variables is scarce. Moreover, from Figure 23, it can be seen that all the categories of the passive variables are near the origin of the axes.

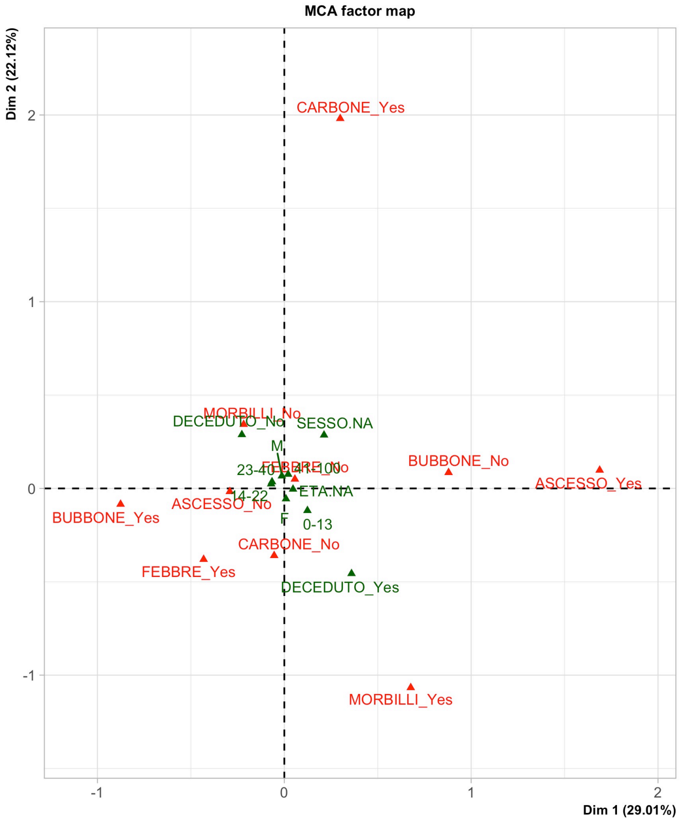


Figure 23: quality of the representation in the axes of MCA for active and passive variables.

Subsequent analyses were performed to understand better the association between plague symptoms and the connection between plague symptoms and the death outcome.

Table 24 reports the combination of plague symptoms and the number of subjects with the different combinations.

Ascesso	*Bubbone*	*Carbone*	Fever	*Morbilli*	Number of subject
No	Yes	No	No	No	1521
No	No	No	No	Yes	554
Yes	No	No	No	No	477
No	No	Yes	No	No	358
No	No	No	No	No	290
No	Yes	No	No	Yes	251
No	No	No	Yes	No	194
No	Yes	No	Yes	No	185
No	Yes	Yes	No	No	157
Yes	No	No	No	Yes	95
No	No	Yes	No	Yes	65
No	No	No	Yes	Yes	51
Yes	No	Yes	No	No	30
No	No	Yes	Yes	No	21
Yes	No	No	Yes	No	16
Yes	Yes	No	No	No	15
No	Yes	No	Yes	Yes	14
No	Yes	Yes	No	Yes	11
No	Yes	Yes	Yes	No	10
No	No	Yes	Yes	Yes	7
No	Yes	Yes	Yes	Yes	2
Yes	No	Yes	No	Yes	2
Yes	Yes	Yes	No	No	2
Yes	No	No	Yes	Yes	1

Table 24: combination of plague symptoms and the number of subjects with the different combinations.

It can be seen that the majority of subjects had only one plague symptom, and, in particular, the presence of only the *Bubbone* (1521) is the most common, followed by the presence of only the *Morbilli* (554) and the *Ascessi* (477). Subjects without the considered plague symptoms were 290.

Univariate models with GLM were performed to quantify the association between the subject status (able to survive to 1485 or death during the year), sex, age (three classes), and the plague symptoms. The presence of women's condition was also considered for the association of the outcome (alive or death), but it was estimated the association was only for the presence of almost one condition and the absence of them because of the low frequency of the conditions, as seen before. The association was expressed as odds ratios (and the relative 95% C.I.). In those analyses, the reference category was the absence of the considered symptoms (i.e., for the *Bubbone*, the reference was the absence of this sign).

Table 25 shows the result of the univariate and multivariate analysis with the odds ratio (95% C.I. and p value):

Univariate analysis with symptoms and condition with two categories		
VARIABLE	**ODD RATIO (95% CONFIDENCE INTERVAL)**	**P VALUE**
GENDER (males) (Reference category: females)	0.97 (0.86, 1.1)	0.66
AGE (Reference category:0-13)		
14-22	0.47 (0.39, 0.56)	<0.001
23-40	0.48 (0.40, 0.56)	<0.001
41-100	0.68 (0.56, 0.83)	0.00013
Ascesso (Yes) (Reference category: No)	1.01 (0.85, 1.20)	0.93
Bubbone (Yes) (Reference category: No)	0.41 (0.36, 0.46)	<0.001
Carbone (Yes) (Reference category: No)	0.46 (0.38, 0.56)	<0.001
Febbre (Yes) (Reference category: No)	0.26 (0.20, 0.33)	<0.001
Morbilli (Yes) (Reference category: No)	21.48 (17.67, 26.12)	<0.001
WOMEN IN CHILDBEARING AGE		
Female_condition (Yes) (Reference category: No)	8.75 (4.60, 16.64)	<0.001

Table 25: result of the univariate analysis with the odds ratio (95% C.I. and p value).

Regarding sex, males have an odds of 0.97* odds of females: this value is near 1, and the 95% confidence interval includes 1, so the two odds were similar, and the gender seems not to associate with the outcome (death). Moreover, the test is not significant, so there is no evidence of a significant association.

Regarding age, the reference category is “0-13” years. Taking into consideration the age class “14-22,” the odds of “14-22” is 0.47 times the odds of “0-13”, so higher age classes seem protective of death events. In addition, the 95% confidence interval does not include 1, and the test is significant, so there is evidence of the association. It is the same for other contrasts about the age classes ("23-40" vs. "0-13" and "41-100" vs. "0-13").

The presence of female conditions, such as pregnancy and menstrual flux, are associated with the outcome (death): in almost one of them, the odds are 8.75 times higher than the odds of no condition.

The presence of *Ascesso* does not seem to be significantly associated with the outcome (death) concerning the condition of an absence of the *Ascesso*.

The presence of *Bubbone* seems to be a protective condition against death (concerning the absence of *Bubbone*). It is the same for *Carboni* and *Febbre*.

Morbilli were highly associated with death: the odds of having a *Morbilli* was 21.48 times the odds of not having *Morbilli*.

Bodily location of plague signs

Only subjects with at least one symptom of plague (fever was excluded because it has no connection with a specific part of the body) referred to at least one part of the body were included in the analysis. Three thousand forty-one subjects were included. Two thousand three hundred four subjects have only one symptom and one part of the body in the description of the cause of death.

In Table 26, for each sign, is presented the absolute frequency of each part of the body:

Position of the body	**Number of times the sign is in the position of the body**
ASCESSO	
Inguin	200
Cosc	146
Orecchi	112
Ascell	85
Cervello	21
Collo	17
Gola	12
Mento	7
Bracci	2
Mammell	2

Bocca	1
Ginocch	1
Mandib	1
Mascell	1
Natic	1
Omero	1
Spall	1
Tibi	1
CARBONE	
Tibi	74
Cosc	68
Bracci	51
Collo	33
Omero	32
Petto	30
Spall	30
Ginocch	16
Natic	15
Orecchi	15
Inguin	14
Mento	14
Ascell	13
Gola	12
Mascell	11
Ventre	10
Mammell	8
Mandib	5
Cervello	3
Bocca	2
Gomit	1
MORBILL	
Cosc	7
Inguin	4
Petto	3
Ascell	2
Natic	2

Bocca	1
Gola	1
Mento	1
Omero	1
Ventre	1
BUBBON	
Inguin	716
Cosc	642
Ascell	596
Orecchi	37
Cervello	30
Collo	17
Gola	17
Bracci	6
Bracci	5
Ventre	4
Mandib	3
Mento	3
Natic	2
Spall	2
Bocca	1
Ginocch	1
Gomit	1
Mammell	1
Mascell	1

Table 26: for each sign, is presented the absolute frequency of each part of the body.

Time between symptoms and death

For 414 subjects, there was a single indication about a day of the week in the cause of death.

For seven subjects, there were two indications of a day of the week in the cause of death, and for 2, there were four. Four subjects with multiple days cited were excluded because the day the symptoms offset was not clear.

Table 27 presents the distribution of the time from symptoms to death.

Time (days)	Number of subjects
1	34
2	76
3	116
4	147
5	5
6	2
7	34
9	3
10	10

Table 27: distribution of the time from symptoms to death for subjects.

The median time from symptoms to death was 3 days (IQR: 2-4).

5.4 Discussion and conclusion

Thanks to the Milan death register of 1485, a plague year, we analyzed the presence and the combination of plague symptoms and their association with death events. This analysis was possible because the registry also contains people who survived the period (beyond the initial reporting event, when confirmation of a plague case was made). Studying the time between the symptoms offset and the death was also possible.

The principal signs and symptoms described in the registry and referred as plague were *Febbre* (fever), *Bubboni, Carboni, Morbilli,* and *Ascessi*. The most frequent symptom in the analyzed cases was the *Bubbone.* The definition of those signs probably was not different from one physician to another, but the diagnosis for the same swelling could be different (S. K. Cohn, 2010).

The definitions for those terms used in the Renaissance are different from those used in modern medicine, in which the terminology used for plague consists of buboes, ulcers, and glandular swellings. Some physicians explained some terminology used in the past. A Udinese physician, in 1576, wrote that there were three categories of symptoms: *Petecchie* are stained or bumps that could appear in many parts of the body, and they are similar to flea bites. *Carboni* or *Antraci* can have different shapes and colors, appear in all body parts, and are associated with fever. The *Bubboni* appear in proximity of the lymph nodes. A Genovese surgeon in 1575-78 wrote that the *Bubboni* have an elongated form at first appearance, but then they become rounded. They can have different colors, but black is mainly associated with death; the *Carboni* were tiny and circular, and the subject has a higher probability of death if they are associated with fever and other symptoms. Paolo Bellintani, a Capuchin who paid service in the *Lazzaretti* of Milan, Marseille, and Brescia, also observed the association between the black color and the increased likelihood of death. Ancient treaties on this topic in Milan did not survive in time. Still, the recognition of plague symptoms was of primary importance for all people, not only those who were enrolled in the health magistrate (Cohn, 2010).

The symptoms were scarcely associated with each other. In the univariate analysis, which considered only the presence and absence of symptoms, it appeared that *Bubboni,* or *Carboni* seemed protective against death, but that was not true for the presence of *Morbilli*. The *Ascesso* was not associated with an outcome.

The presence of at least one particular condition in women of fertile age, such as childbirth, abortion or puerperium, was associated with an outcome.

The peak of the plague epidemic in 1485 in Milan was in autumn. A study about the plague in the Swiss territory from 1628 to 1630 with data mainly from the parish registers was conducted on the seasonality of plague peaks. Thirty-eight communities scattered through the territory were considered. Forty-three plague peaks, in different towns and over the course of a century, were considered: 39 were from September to January, one was in spring, and three were in summer. This pattern is also present in the period of the fourteenth and fifteenth centuries in continental Europe. Still, there were summer peaks in Great Britain (1349-1665), with the diminution of mortality in the cooler months. It is possible that in the summer there were conditions incompatible with the vector: the temperature was too high; in Great Britain, this condition was mitigated by the nearness of the sea. (Eckert, 1980).

We analyzed the age and gender distribution of subjects written in the registry and whose deaths were attributed to plague, we found similar distribution of females and males, with a slightly excess of females deaths in the class [15-20]. The analysis shows that gender is not associated with the odds of death. Age is significantly associated with the event of death. Many studies were about individual susceptibility to plague: the study of the skeletons in the Black Death period concluded that there were no significant differences in death events in males and females, and it was the same in England and North Italy for the fifteenth century period. Studies about Milan from the period 1452-1523 showed a prevalence of female deaths from plague, and could be related to the poverty of widows and migrant women (Alfani & Murphy, 2017).

The 1485 data shows a lot of registration for the age classes [15-20] [10-15] [25-30], and there were a major percentage of death in the classes [0-5], [15-20]. To our knowledge, there is no document about the composition of the population in the considered period.

Studies that considered skeletal evidence in the period of the Black Death suggested there is a relationship between frailty, poor health, age, and the possibility of dying from plague. From the fifteenth to the sixteenth century, the plague hit mainly subjects less than 12 years old, particularly considering the excess of mortality, with this order of age classes: 11-20, 21-30, <5.

It was possible to take into account the demographic structure of the population in a study about the plague of Marseille in 1720-1722: from this work emerges that an increase in age has a positive association with the possibility of death for plague; also considering Milan in the period 1629-1630, it appeared that children (<10 years old) were less affected and the age classes more hit was 40-60 years (Alfani & Murphy, 2017). The mortality risk factors in a modern population are linked to poor hygienic conditions, overcrowding, and poor nutrition. Those conditions are connected. In a non-plague situation in a contemporary period, the risk of death increases with age, but there is a peak in the perinatal age. In an epidemic plague, this trend completely transforms with the addition of many peaks that reflect the categories most at risk of dying from the plague. Children from 5 to 9 years old are at higher risk because they play or sleep on the ground, so they are more prone to flea bites. For young adults, risk is highly determined by the work activities. In particular farmers, manufacturers, and trades have a higher risk of death (Rubini et al., 2016).

Women were at high risk of contracting the plague probably because they spent much time at home, where rats and insects live.

The median time from symptoms to death was three days. Considering the Milanese epidemics of 1452, 1468, 1483, 1502, and 1523, Cohn reported that the median time to death was less than three, and the modal was two days (S. K. Cohn & Alfani, 2007). A similar time-to-death estimation for bubonic plague was obtained for an urban plague epidemic in Madagascar (August-November 2017): 2 days from symptoms to death (Randremanana et al., 2019).

6 TIME SERIES ANALYSIS ON HISTORICAL DEATH DATA OF MILAN FROM THE MILAN REGISTERS (1452-1845)

6.1 Introduction

Time series based on historical data carry precious information about past events (above all, health and disease).

For example, the London Bills of Mortality, a historical registry of burials and christenings in the Anglican parishes, captured data from the last part of the sixteenth century until 1858. But only from 1664 general causes of death were recorded (thus, not only plague deaths). John Graunt's observation on the Bills of Mortality, published in 1662 (before the 1665 plague), performed one of the first numerical analyses on demographical data and the first life tables. An example of his work is about the plague in 1592, 1603, 1625, and 1636: he considered seasonality, comorbidity, and time dependence. Many call him the "father of demography."

His conclusions are fascinating:

"It may be further observed, that the time of the Plagues continuance at the height was of several durations, for Anno 1592 it continued from the first week in July to the second of September, without increasing or decreasing above 100 in 1600; whereas in 1603 it remain'd but three weeks at the state, decreasing near 1⁄4 the next week after the height; Anno 1625 it remain'd not three weeks at a stay, increasing 1⁄16 part the next week before the height, and decreasing as much the next week after. 1636 it stood five weeks without increasing or decreasing above 1⁄16 part afore mentioned" (Graunt, 1676). And "The last thing I shall observe is, that in all the four great years of mortality above-mentioned, I do not find that any week the Plague increased to the double of the precedent week above five times" (Graunt, 1676).

All these aspects were also considered in the modern analysis of public health. Another study of time series based on historical death registries in England was based on The Bills of Mortality of London and the Parish Registry of Penrith. Penrith was a small, isolated rural town. The analysis of the series has been made to study the periodicity of the smallpox epidemics in the two locations and to assess the impact of the external variables, such as the shortage of food. Spectral analysis was employed (Krylova & Earn, 2020).

Giuseppe Ferrario was born in Milan in 1802. In 1825, he graduated in medicine and surgery at the University of Pavia. He worked at the *Ospedale Maggiore* in Milan, the *Istituto S. Corona*, and the *Accademia dei Filodrammatici*. But he could have stood out more as an author in the clinical field. He was an acculturate man and he wrote some historical essays about the Milanese healthcare, such as "*Cenni storici sugli ospizi ed antichi spedali de' vecchi in Milano e statuto medico-economico del Pio Albergo Trivulzio nell'anno 1831"*. He founded a medical mutual aid association, the first in Italy.

He made essential studies in Lombardy's health history since 1833, with his work about sudden deaths (Ferrario, 1834).

In 1845 he founded the Physio-medical-statistical Academy of Milan (Ferrario, 1846).

He died in Milan on 2 November 1870 (Crespi, 1996)

The second volume of his work, entitled "*Statistica medica di Milano dal secolo 15. fino ai nostri giorni*" (Ferrario, 1840), reports the history of Milan, the topography of the city, observations about the climate of Milan (and Lombardy), the censuses of Milan and in the provinces in different periods, the mean annual price in Milan for food, liquids (such as wine and oil), wood and coal, silk and clothes, a study about coin values, some information about the population of Milan (and Lombardy) in different periods, and a chapter about hospices and particular Milanese hospitals.

The chapter about the population in Milan and Lombardy contains tables about the number of deaths in parishes and the *Corpi Santi* (outside the city walls) from 1452 to 1845. Different tables were created for the deaths in the hospitals, but they were not considered in the analysis because the numbers were missing until 1756. Moreover, until 1756, the deaths of religious people were excluded because their deaths were not reported to the *Ufficio di Sanità* (Health Office). The table reports the number of deaths in each month. Those data, as well as all historical time series, reflect the events and their consequences (Griffin & Isaac, 1992).

Many time series show structural differences in trends, frequency, and probability distribution. These changes can be due to external factors and/or internal systematic changes. Here, we identified changepoints, one or more time points that locates this change and subdivided a time series into homogeneous segments.

There is a lot of changepoints typology (Schroth et al., 2021):

-1-Changing in the mean, that is, a level change.

-2-Changing in the variance: the value of the mean is constant along the sequence, but there are differences in variances. It is observable thanks to changes in the peaks of the oscillations.

-3-Changing in the periodicity: it is studied thanks to the wavelet

-4- Changing the pattern: it is studied thanks to the Wasserstein method.

Outliers in a time series analysis are observations that are different from the pattern of other values: they can be generated from measurement errors, but they can also be derived from something that impacted the series in a way that determined the different behaviors of the monitoring variable. Identifying the outliers (with an automatic method to avoid subjectivity) and interpreting them is essential. Recognizing the presence of outliers is also crucial to the subsequent series analysis (i.e., ARIMA models are impacted by the presence of outliers) (Tsay, 1988).

There are five types of outliers, classified according to the effects on the time series (Peña, 2000):

-Additive outliers: can be due to a particular event happening in a moment of time of the time series (in our case, it can be an epidemic event). Additive outliers affect the residual and the parameter estimations but do not influence the data from the subsequent period. Additive outliers can be either exceptionally high values or exceptionally low values.

-Level shift: an observation corresponding to a change of the mean that persists until the last observation of the time series. From a stationary series, a level shift changes it to a non-stationary process.

-Innovational outliers: a level shift with a temporary effect. It affects the autocorrelation (so in the estimation of the parameters). An example of innovational outliers in historical time series can be a demographic change in the population under study.

- Temporary change is similar to level shift, but the effect decreases exponentially on the subsequent observations. The time series may return to the original level.

-Seasonal level shift: an additive outlier verified at regular intervals.

Outliers will be explained according to the work of Alfonso Corradi (Corradi, 1972), and details will be added according to epidemiological research. Outliers that are not linked to epidemic events will be searched according to other historical references. Unexplained outliers can be studied by referring to the causes of death in the *Libri Mortuorum*. Still, this source could be employed to better understand an epidemic phenomenon, for example, according to socio-demographical data.

Changepoints could be interpreted according to demographical events.

In this study, for the first time, a method to evaluate the presence of outliers and changepoints is applied in Ferrario's time series. In particular, it employed a method for studying environmental time series.

The study of Ferrario's time series can provide new information on the impact of known events, such as the plague of 1630. It can be a starting point for research, for example, using the causes of death in the *Mortuorum Libri*, to always have more excellent knowledge on the epidemics and other events that affected Milan.

6.2 Methods

Data contained in Ferrario's tables were included in a spreadsheet. We began with the period from 1581, because the data were complete from January of that year. The large amount of missing data from the previous period did not allow us to estimate the missing values. In addition, any estimation would have been highly impacted by external variables, that we could not determine, such as good demographic information.

The time series was subdivided into periods of about fifty years each, as recommended by prior literature for long time series (Munoz-Tuduri et al., 2006). Separating the data offered from the *Mortuorum Libri* (1452-1801) to the later sources of data from Trivulziana Library (from 1802 to 1845) was preferred.

The number of days for each month was corrected because it was possible that the differences in the number of days in each month could impact the data. So a "calendar adjustment" was performed: Wt=Yt *((365.25/12)/(number of days in the considered month)). Wt is the corrected number of deaths for the considered month, and Yt is the number of deaths in the Ferrario (Wheelwright et al., 1998).

A time series analysis was performed to evaluate the presence of changepoints and outliers.

Changepoints indicate a change in pattern concerning the previous time period. Changepoints identify intervals with similar characteristics, which make an analysis of seasonality or longer cycles easier.

Outliers are data in which the number of deaths was either very low or very high with respect to previous and subsequent months. They can be associated with known mortality crises, such as epidemics, famine, or significant population fluctuations.

To this aim, the method proposed by Čampulová (Čampulová et al., 2018) was used because it was developed using environmental data that are affected by several external variables. This package seemed interesting because Ferrario's time series of deaths in Milan was perturbed by plague and other epidemics and war. When the deaths occurred outside of the city, they were not recorded, but the events could have negative repercussions for the urban population, for example, in the provisioning of the city.

The method is divided into three steps:

1)Kernel **regression** is applied to the time series. It was performed using a local method and a global method for smoothing.

2) Change points **are identified on the base of kernel regression residuals**. Thanks to changepoints, the series was divided into segments.

3)Outliers were identified in each segment.

Kernel regression

Was the non-parametric method used to estimate the relationship between time points (independent or predictor variable) and the number of deaths (dependent variable) on the n sample points (t_i, D_i).

$$D_i = m(t_i) + \sigma(t_i)\varepsilon_i$$

ε_i are symmetric random variables with zero mean and unit variance, $\sigma(t_i)$ is a function for the standard deviation of D_i depending on t_i.

$m(t_i)$ is estimated by

$$\widehat{m}(t_i) = \sum_{i=1}^{n} D_i \int_{s_{i-1}}^{s_i} \frac{1}{h_{t_i}} K\left(\frac{t_i-u}{h_{t_i}}\right) du$$

For each t_i, a weighted sum of t_i (i=1…,n) is calculated; the weights are given by $\int_{s_{i-1}}^{s_i} \frac{1}{h_{t_i}} K\left(\frac{t_i-u}{h_{t_i}}\right) du$ where $K\left(\frac{t_i-u}{h_{t_i}}\right)$ are the kernel values correspond to t_i. h_{t_i} =h(t_i) is a bandwidth in point t_i. The integration limits, centered at t_i, are $s_i=0.5(t_i+t_{i+1})$ for i=1,…, n-1.

The degree of smoothing of the function is given by the bandwidth h which can depend on t_i (local) or can be a common value for all t_i values (global). The smoothness of the function increases with increase of h.

Epanechnikov kernel was used: $K(z) = 0.75(1 - z^2)$ where z= $\left(\frac{t-u}{h_t}\right), z \in [-1,1]$

After choosing the kernel function, to obtain $\widehat{m}(t_i)$, the estimate of bandwidth is needed.

For local bandwidth local error (MSE= $E(\widehat{m}(t, h_t) - m(t))^2$) and for global bandwidth the error mean integrated square error (MISE=$E\left(\int_a^b w(t)[\widehat{m}\,(t, h_x) - m(t))^2\, dt\right)$ where the w(t) is a weight function) are minimized.

A local plug-in estimator is used where h_t can be found by an iterative procedure.

The global and local methods presented differences in the plug-in to choose the bandwidth: the local method is the one proposed by Herrmann (Herrmann, 1997), and the global method is the one by Gasser et al. (Gasser et al., 1991).

The series was subsequently subdivided into J segments s_J by J-1 changepoints τ_j. These are defined as the time points in which the distribution of kernel regression residuals changes. Within each segment, the distribution of residuals is supposed to be homogeneous, and it is supposed to be different among residuals pertaining to different segments. The number J-1 and the position τ_j of the change points were iteratively identified by the non-parametric E-Divisive method. The algorithm starts with a single change point τ_1; a new change point is added at each iteration, and its contribution is tested. The test statistics is based on the Euclidean distance between sample observations pertaining to the different segments defined within change points, and the permutation test was used to obtain the exact p-value for the null hypothesis that no additional change point is needed. The procedure is stopped after the first non-statistically significant result (James & Matteson, 2013; Matteson & James, 2014).

The above method for change point identification is based on the assumption of independence among residuals. This was investigated by the autocorrelation function. For each time point, the corresponding regression residual is considered; the autocorrelation is the correlation between a residual r_{t+h}(corresponding to a time t+h), with the residual r_t corresponding to a previous time (t). The difference between the two times (h) is the lag.

For each lag (h) firstly the autocovariance was estimated by

$$v(h) = \frac{1}{n}\sum_{i=\max(1,-h)}^{\min(n-h,n)}(r_{i+h} - \overline{r})\,(r_i - \overline{r})$$

then the autocorrelation was estimated by $a(h) = \frac{v(h)}{v(0)}$

The autocorrelation function was plotted together with the approximate 95% limits of the expected correlation coefficients under the null hypothesis of no correlation.

The autocorrelation of residuals is expected to depend on the kernel regression bandwidth; thus, a significant autocorrelation could be observed for low lag times (e.g., lower than 4). A non significant correlation for lag times greater than 4 could be considered empirically reasonable to assume the independence between residuals “enough far one from another” (Capasso, 2021).

Moreover, sparse significant autocorrelation, which is of small value, could be considered an “artifact” without a clear indication of a “systematic” dependence among residuals.

Finally, the outliers are identified in each segment using the Chebychev inequality. Outliers are defined as the observations lying outside of the interval $\pm L s_j$ where s_j is the standard deviation of the distribution of residuals in the segment j. L is chosen as the value which maximize the largest GAP among the ordered residuals in the set of residuals $\leq L s_j$ (see Appendix).

Outliers were tentatively explained by some historical source, such as the work of Alfonso Corradi (Corradi, 1972).

6.3 Results

Ferrario's time series covers the period between 1452 and 1845. It reports the monthly number of deaths; only data beginning in 1581 was considered because of the high frequency of missing values for the previous period.

Figure 24 presents the complete series with the "calendar adjustment" (see method).

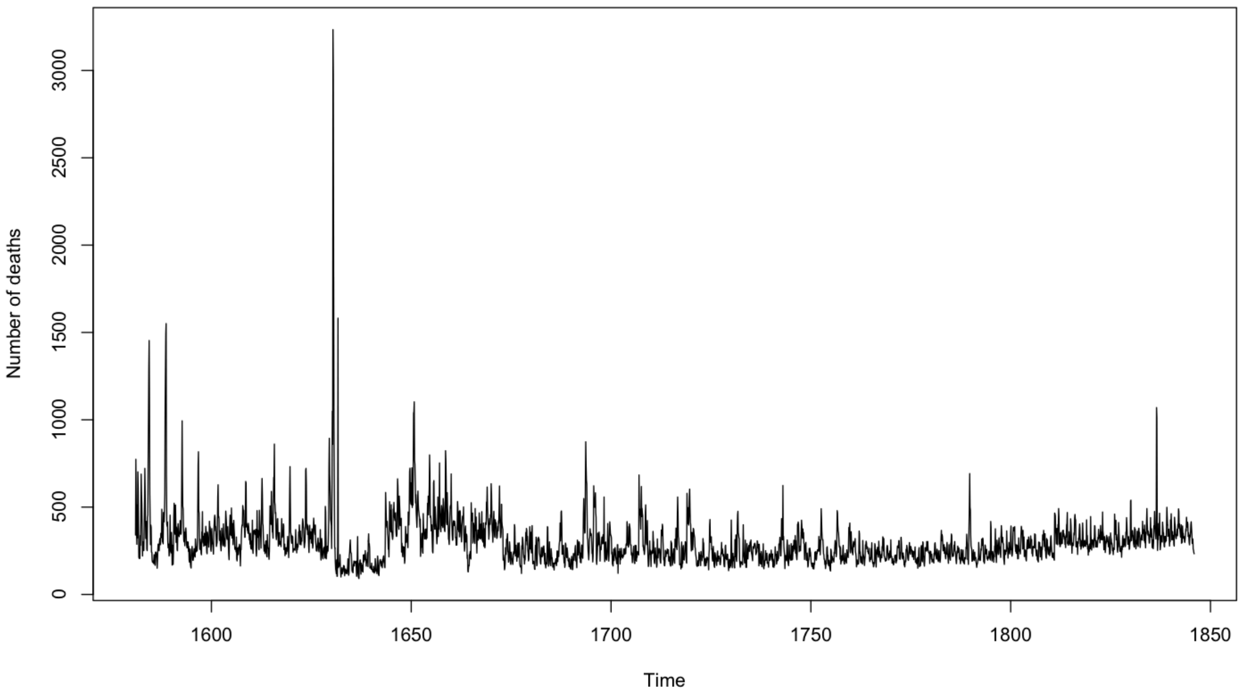


Figure 24: Monthly number of deaths in the Ferrario's time series from 1581 to 1845.

The higher peak due to the 1630 plague is apparent. Other relatively high peaks are identified mainly during the first part of the series (from 1581 to 1650). The series appears to be more "regular" from about 1675. This decline in mortality variations may be due to the disappearance of certain diseases in Milan, for example, plague. It was also likely to improve socio-sanitary conditions and medical discoveries (Tasca, 2013).

During the eighteenth century, there were many doctors, new hospitals, and all social classes had access to cures. From the seventeenth to the eighteenth century, public welfare was translated from a concept of happiness, justice, and tranquility to a more concrete way:

- Distribution of goods
- Development of the city and communication routes

- Improving the living conditions
- Implementation of health

The politics of health during the seventeenth century notably led to the development of these three aspects: families were involved in the well-being of the children; hygiene and medicine were viewed as a form of social control of the population (the eradication of epidemics, the decrease in the mortality rate and an improved life expectancy become goals of primary importance and this translated into an increase of power for the doctor). Hospitals become not "only" care facilities for the poor population but disease treatment facilities. Notably, physicians' education became more standardized in this period (Foucault, 2014).

Moreover, from the second half of the seventeenth century, in different parts of Europe, physicians focused on the disease as a product of disorder between men and the environment. With the study on the factors that can cause a disease, it became possible not only to act on the symptoms to mitigate them (as has always been done), but also prevent the disease itself. In the middle of the eighteenth century, the "medicine of avoidance" arosed. The physicians had to monitor the environment and, eventually, remove all the elements that can promote disease with the drainage, lavation, ventilation, and reinterment. This approach led to a decrease in mortality from the seventeenth century (Riley, 1987).

It is important to note that the population of Milan increased from 1500 to 1800 (in 1500, it was estimated to have 100.000 inhabitants. In 1650, 120.000, and, in 1800, 135.000) (Bairoch et al., 1988). Hence, the differences in the number of deaths in different periods have to be interpreted according to this fact.

The series was also divided into five periods of about 50 years: 1581-1654, 1655-1705, 1706- 1754, 1755- 1805, and 1806- 1845. Each period was considered separately.

Figure 25 presents the local and global Kernel smoothing method for the first piece of the time series (1581-1654).

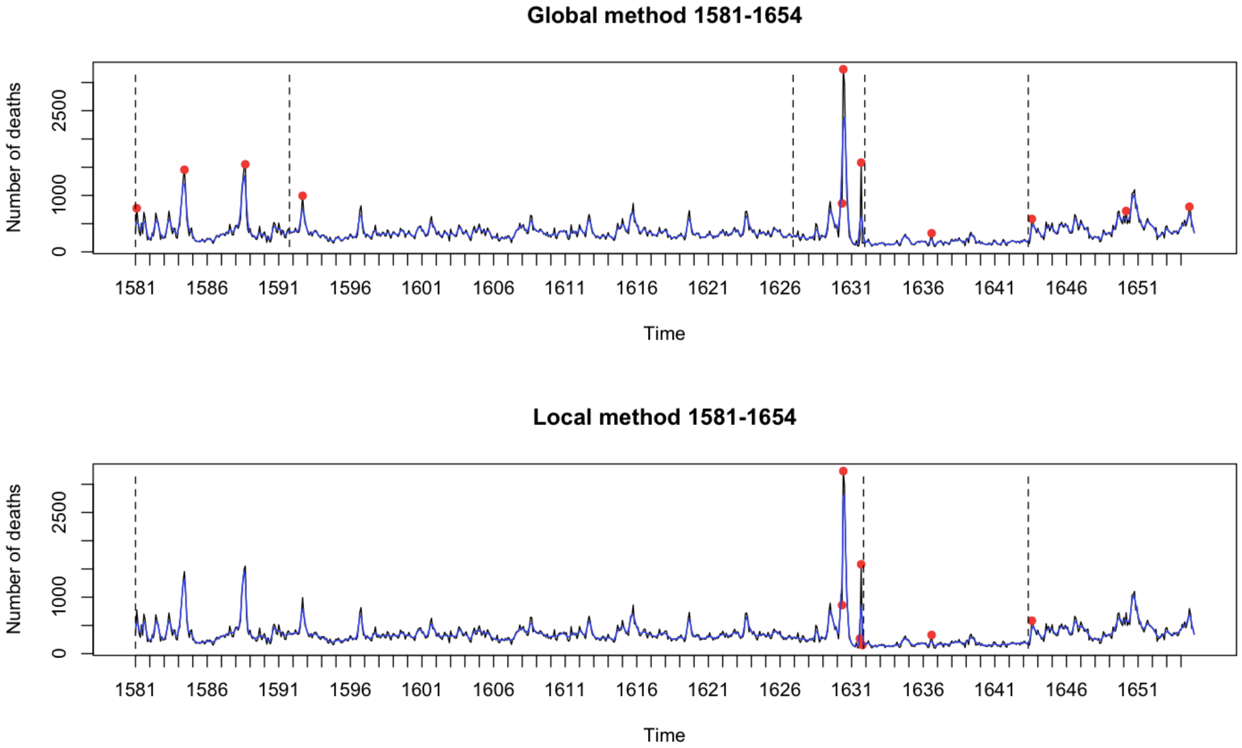
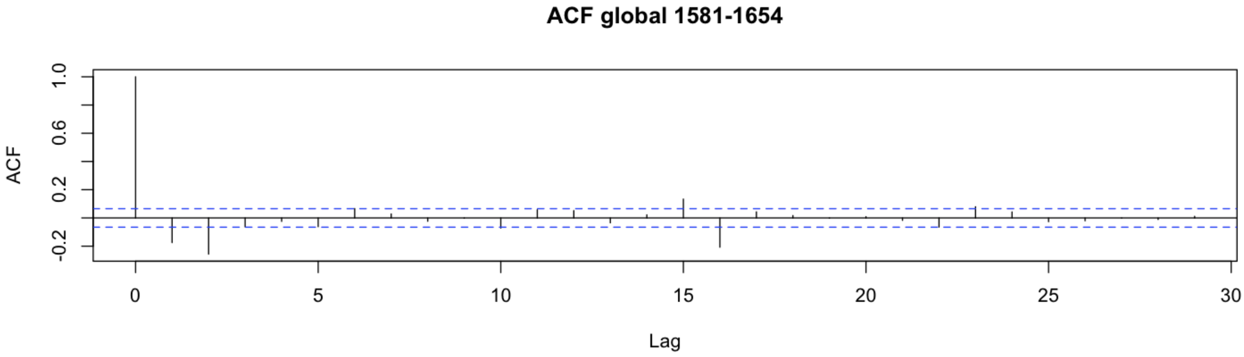


Figure 25: local and global Kernel smoothing for the time series from 1581 to 1654. The points in red are the outliers found by the two methods. The black line represents the original series, the blue line represents the Kernel smoothing series, and the vertical bars represent the changepoints.

Figure 26 presents the autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1581-1654).



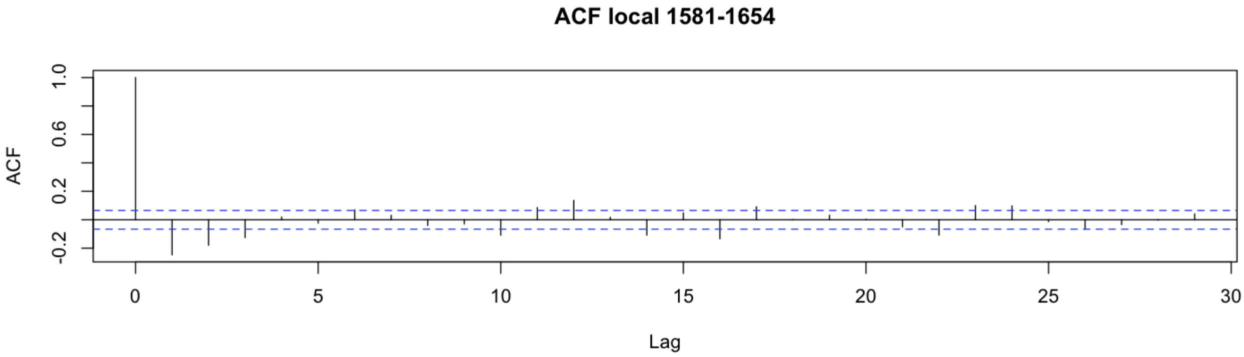


Figure 26: autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1581-1654.

Table 28 reports the outliers. The events in the outliers' date (or years), according to Corradi (Corradi, 1972), were reported. Table 29 reports the bandwidths identified by the two methods.

LOCAL METHOD		**GLOBAL METHOD**		**EVENTS**
OUTLIERS	NUMBER OF DEATHS (Original series)	OUTLIERS	NUMBER OF DEATHS (Original series)	
		1581-02-01	774	
		1584-06-01	1454	
		1588-09-01	1551	
		1592-09-01	994	
1630-05-01	859	1630-05-01	859	Plague
1630-06-01	3234	1630-06-01	3234	Plague
1631-08-01	266			
1631-09-01	1582	1631-09-01	1582	
1631-10-01	153			
1636-08-01	331	1636-08-01	331	
1643-08-01	584	1643-08-01	584	
		1650-03-01	725	
		1654-08-01	799	

Table 28: outliers in the time series 1581-1654

LOCAL METHOD		**GLOBAL METHOD**	
BANDWIDTH (Start date)	BANDWIDTH (End date)	BANDWIDTH (Start date)	BANDWIDTH (End date)
1581-01-01	1631-10-01	1581-01-01	1591-09-01
1631-11-01	1643-04-01	1591-10-01	1626-11-01
1643-05-01	1654-12-01	1626-12-01	1631-11-01
		1631-12-01	1643-04-01
		1643-05-01	1654-12-01

Table 29: bandwidths (delimited by changepoints) in the time series 1581-1654

It can be noted that the local method was put into the same segment, delimited by the changepoints, the plague period (from around 1629 to 1632), and the previous period. This approach was judged unreliable because changepoints usually occur where the characteristics change abruptly, and it is reasonable to assume that plague greatly impacted this series with its appearance. In the global method, the period of plague and the immediate subsequent period are “separated” from the rest of the series. The global method seems to produce better results, according to the fact that plague led to an impressive number of deaths and was preceded by famine, war, and another epidemic (Galli, Nodari, et al., 2023; Galli, Oreni, et al., 2023) and the epidemic was followed by a period of a drastic decline in deaths, produced more reliable results. In the subsequent period (about from 1632 to 1643), it can be seen that there was a downward shift in the number of recorded deaths. Also there was a vertical shift, after this period, and the series level became comparable to the pre-plague one.

In 1500, Milan was one of Europe's most important centers of artisan production, commerce, and agriculture. In 1535, it was annexed to the kingdom of Spain by Charles V and experienced an unprecedented increase of inhabitants: in 1541, it reached 60.000 people, and in 1574, 120.000. In this period, about half of the heads of families were artisans or merchants. By 1570, it was one of the most populated cities in Europe. With the plague of 1576, about 15% of the inhabitants died, but by 1581-1582, as the city reached again 100,000 people and social structures were restored before 1600. The plague of 1576 made a significant impact. Although was not devastating for Milan in terms of the number of deaths, but economically it was punishing. Demand for Milanese goods from other cities and the competition from non-Italian regions had a tremendous negative impact. The situation worsened with the poor agricultural harvests in the Po River valley, which led to inflation and poverty. From 1581, many bankers in Milan faced bankruptcy. These disruptions contributed to the outliers detected in the first part of the series. The situation improved only from 1594 when Milan had 125,000 inhabitants: this data remained stable until the plague of 1630 (and this can also be seen in the time series in which, according to the employed method, there were no outliers and the series is mainly stable) (D’Amico, 2000).

Milan recovered faster from plague events from an economic and demographical point of view. At the end of the sixteenth century, Milan recovered as a production center in the textile and metallurgical sectors. From 1616 to 1659, we see another demographic decline due to wars and plague (D’Amico, 2001), in which it can be seen that the number of deaths, even if there was a population increase, was low. Could those data reflect the advantage that poor people took from the plague, for example, having greater accessibility to work and higher pay (as many workers have died)? A study about the time series of the mason wages in Milan (compared to Genoa, which was spared from the 1630 plague, and Florence, less impacted by it) does not confirm this hypothesis. It is true that with respect to the smaller urban centers, Milan was less affected economically, but the plague nonetheless had a long-term effect on urban growth. Moreover, Lombardy was affected by the war of Mantua (1629-1631) (Alfani & Percoco, 2019).

The population of Milan passed from 75.000 inhabitants in 1633 to 100.000 in 1647 because of migratory flux from rural territories and merchants and artisans. Thanks to the investments Milan re-established its value as a commercial hub by land (D’Amico, 2001).

In the last part of the series, the increase in the number of deaths could reflect the rise in the population. The exact motivation for the presence of outliers could be investigated thanks to the *Liber Mortuorum.*

Figure 27 presents the local and global Kernel smoothing method for the second time series (1655-1705).

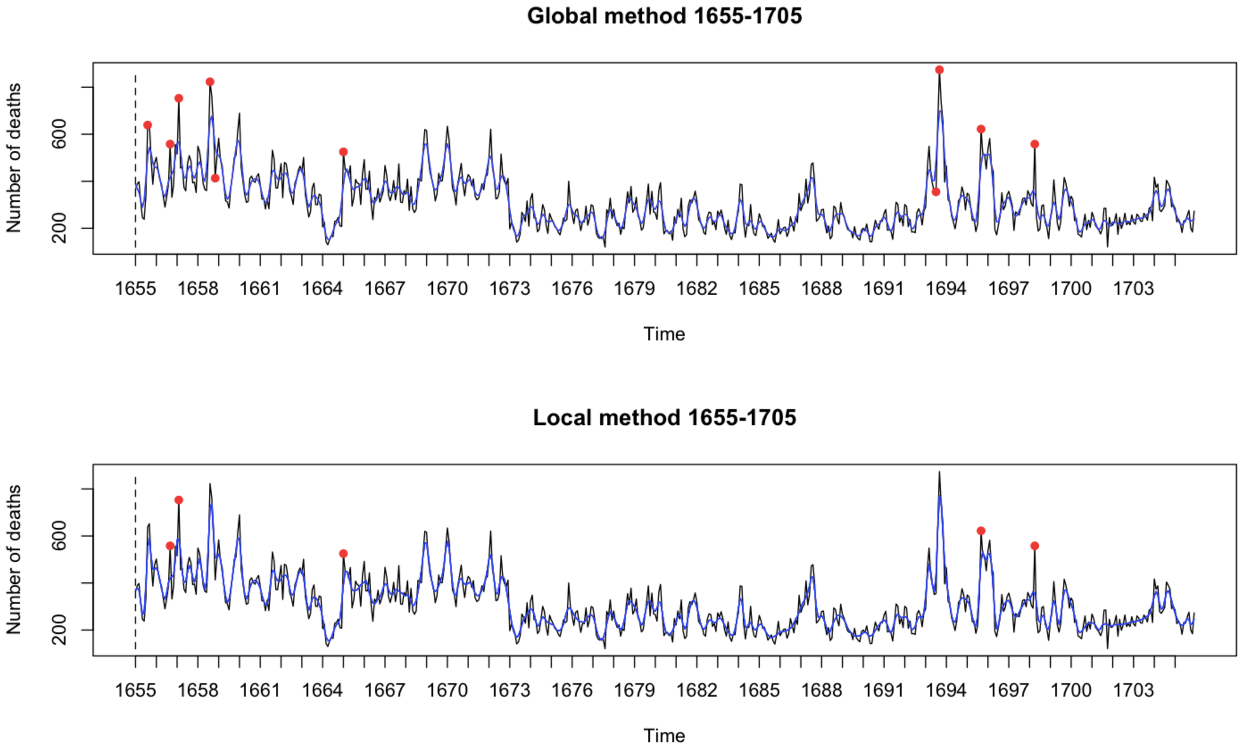


Figure 27: local and global Kernel smoothing for the time series from 1655 to 1705. The points in red are the outliers found by the two methods. The black line represents the original series, the blue line represents the Kernel smoothing series, and the vertical bars represent the changepoints.

Figure 28 presents the autocorrelation plot for the residuals of the global and local Kernel smoothing methods for the time series (1655-1705).

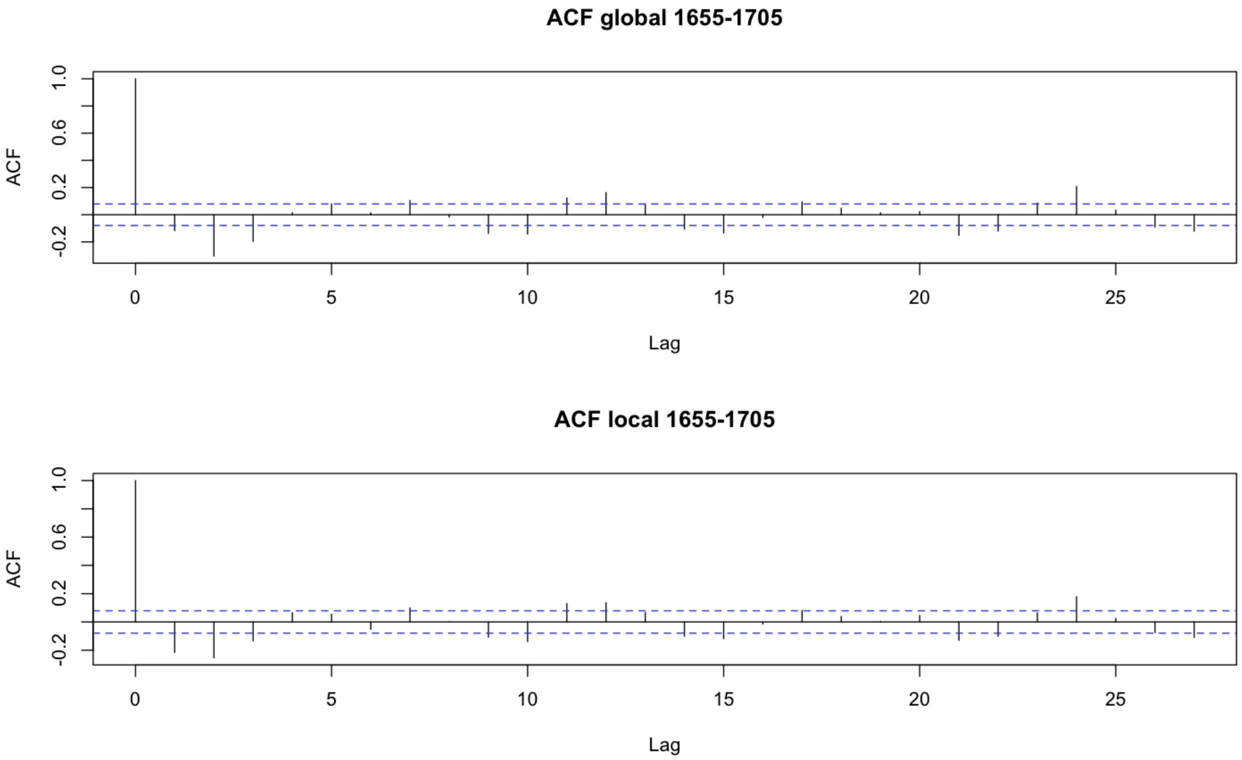


Figure 28: autocorrelation plot for the residuals of global and local Kernel smoothing methods for the time series 1655-1705.

Table 30 reports the outliers. The events in the outliers' date (or years), according to Corradi (Corradi, 1972) were reported.

LOCAL METHOD		**GLOBAL METHOD**		**EVENTS**
OUTLIERS	NUMBER OF DEATHS (Original series)	OUTLIERS	NUMBER OF DEATHS (Original series)	
		1655-08-01	639	
1656-09-01	558	1656-09-01	558	
1657-02-01	753	1657-02-01	753	
		1658-08-01	823	
		1658-11-01	413	
1665-01-01	525	1665-01-01	525	
		1693-07-01	355	Smallpox (which mainly affected children)
		1693-09-01	874	Smallpox (which mainly affected children)
1695-09-01	622	1695-09-01	622	
1698-04-01	558	1698-04-01	558	

Table 30: outliers in the time series 1655-1705

Considering that the Local method did not recognized as outliers any months of 1693 (in which many children died because of smallpox), it could lead us to consider the global method more reliable for the analysis of this time series.

The smallpox virus belongs to the family of Poxviridae. The first site of the infection is the respiratory tract. The transmission is due to the long-lasting face contact at close range (<2m). In the respiratory tracts, there is the duplication of the virus; after that, thanks to the macrophages, the virus spreads to different parts of the body (in the skin, the infection is manifested with macules that become papules, vesicles, pustules, and scabs). The death happened due to toxemia (coagulopathy, hypotension, and multiorgan failure). There are two variants of the smallpox virus: the *Variola major*, with a 30% case fatality rate, and the *Variola minor*, with 1% (Moore et al., 2006).

Interestingly, smallpox was considered an overall benign condition from the ninth to the sixteenth century. For the Islamic physician Rhazes, this infantile condition allows it to pass from the wet blood of childhood to the dry blood of an adult (this follows the humor theory). For Avicenna, smallpox allowed the clearness from the contamination of the mother's menstrual blood in the womb. For Aemilius Campolongus, this state permitted the expulsion of the bad humor through the skin (otherwise, it can corrupt blood and organs).

From 1300 to 1500, it was not considered a dangerous condition, and epidemics were not common; considering London, Italy, Spain, and Paris, there was only one that hit an important part of the population: 1444 in Paris, a city with 35.000 inhabitants, 6.000 inhabitants were infected. Italy was infected during the fourteenth century, a limited number of times: Florence in 1335 (many children), 1363 Siena, 1336 Naples, 1386 Vicenza, and 1393 Bologna.

It is challenging to have some data about the fourteenth century. Still, some numerical information is retrievable from the subsequent century. In the registers of the death of Florence, which contains the detailed causes of death only for the plague period, it can be understood that from 1424 to 1458, only 84 subjects died because of smallpox. Also, in London, the situation for subsequent periods was similar, according to the parish registry of All Hallows London Wall. From 1574 to 1598, 12 people died from smallpox, and 10 were under seven years of age.

There was a change in the number of epidemics in the sixteenth century, as testified by the Italian chronicles: in 1300-1400, there were 7. In 1400-1500, there were 4. In 1500-1550 there were 7, but in 1550-1600 there were 38. According to the data, it is possible that the epidemics until 1550 were caused by *Variola minor* and the subsequent ones by *Variola major* (Carmichael & Silverstein, 1987).

So, it is unsurprising that in Milan, an important and populous center, there were smallpox epidemics that were outliers only in this period according to the number of deaths. Interestingly, Corradi notes that epidemics affected mainly children.

Figure 29 presents the local and global Kernel smoothing method for time series 1706-1754.

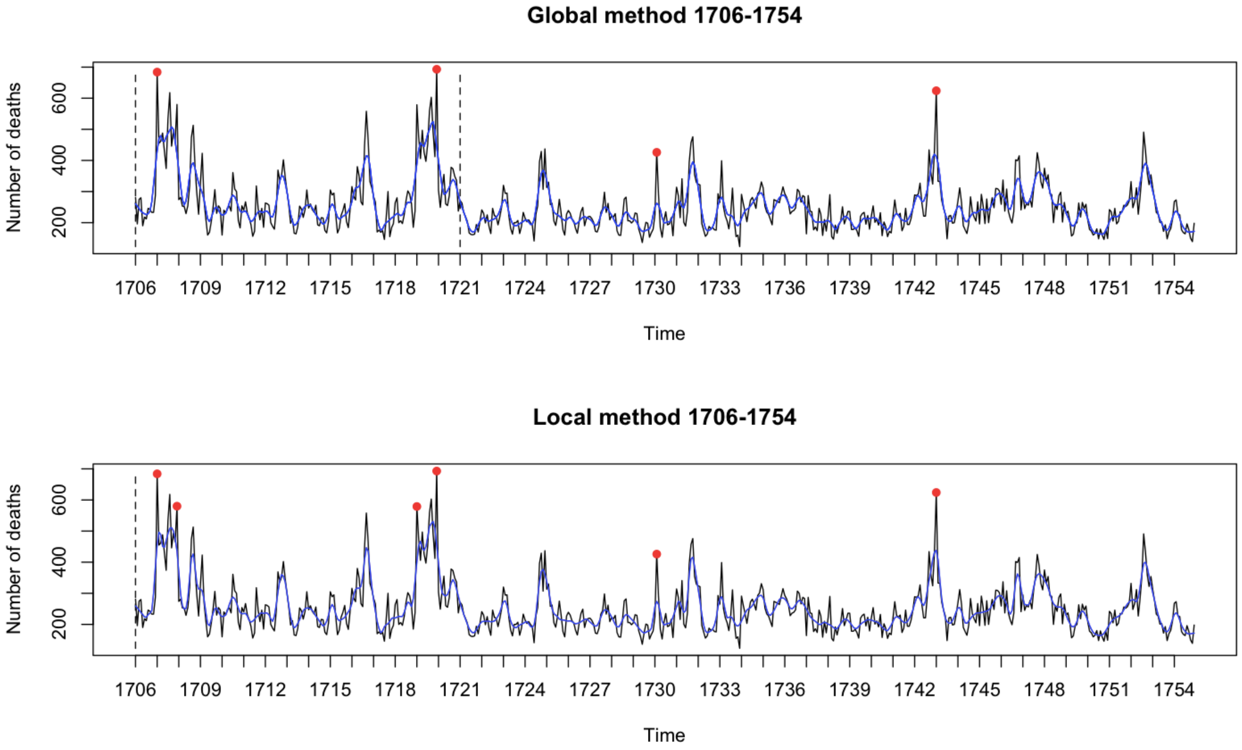


Figure 29: local and global Kernel smoothing for the time series from 1706 to 1754. The points in red are the outliers found by the two methods. The black line represents the original series, the blue line represents the Kernel smoothing series, and the vertical bars represent the changepoints.

Figure 30 presents the autocorrelation plot for the residuals of the global and local Kernel smoothing methods for the time series 1706-1754.

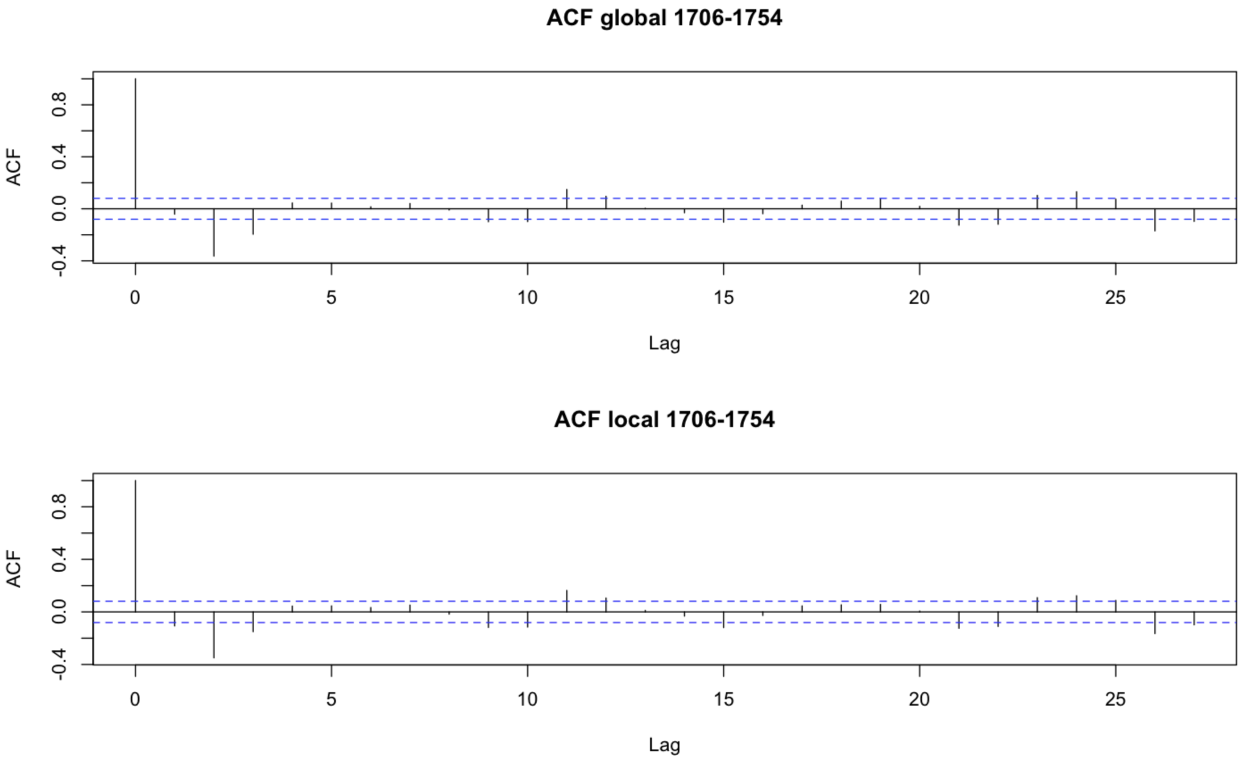


Figure 30: autocorrelation plot for the residuals of global and local Kernel smoothing method for the time series 1706-1754.

Table 31 reports the outliers. The events in the outliers' date (or years), according to Corradi (Corradi, 1972), were reported. Table 32 reports the bandwidths identified by the two methods.

LOCAL METHOD		**GLOBAL METHOD**		**EVENTS**
OUTLIERS	NUMBER OF DEATHS (Original series)	OUTLIERS	NUMBER OF DEATHS (Original series)	
1707-01-01	684	1707-01-01	684	
1707-12-01	580			
1719-01-01	579			Dysentery, *Febbre biliosa* (Bilious Fever), Smallpox
1719-12-01	693	1719-12-01	693	Dysentery, *Febbre biliosa* (Bilious Fever), Smallpox
1730-02-01	426	1730-02-01	426	Catarrhal epidemic (in Lombardy)
1743-01-01	624	1743-01-01	624	*Influenza* (flu)

Table 31: outliers in the time series 1706-1754.

LOCAL METHOD		**GLOBAL METHOD**	
BANDWIDTH (Start date)	BANDWIDTH (End date)	BANDWIDTH (Start date)	BANDWIDTH (End date)
		1706-01-01	1720-12-01
		1721-01-01	1754-12-01

Table 32: bandwidths (delimited by changepoints) in the time series 1706-1754

Bacillary dysentery was common in rural Europe and became a killer due to poor sanitary conditions and poor nutrition (Black, 1999). In a modern definition, dysentery, provided by the Oxford English Dictionary, is "A disease characterized by inflammation of the mucous membrane and glands of the large intestine, accompanied with griping pains, and mucous and bloody evacuations". Bloody evacuations are not always present in the case of dysentery.

There are two types of dysentery, bacillary and amebic, with similar symptoms. The bacillary one has a mortality rate of 15% in modern times in Africa and Latin America; death is due to dehydration and poisoning by the bacterium's toxins. The amebic one is chronic, with an intermittent and long-lasting effect; it causes ulcers, liver abscesses, and secondary infection to the lungs, brain, and spleen, thus not a likely cause of dysentery deaths in Milan. Dysentery is typical of tropical areas but is also present in temperate territories, where the bacillary form is more diffuse (Haycock, 2002).

Gram-negative bacteria of the Enterobacter class cause the bacillary form. It was discovered by Kiyoshi Shiga (1870-1957) during an epidemic in Japan in 1897: 90.000 people were infected, and 20% died. The reservoir is the human digestive tract, so the feces make dispersion possible. Transmission is possible through dirty hands and contaminated water (Nicolas et al., 2007).

In the eighteenth century, it was believed that dysentery was caused by air. The corrupted air is caused by marshland reclamation (putrefaction of the corpses of plants and animals), human feces in the presence of bad air, and straw rotting in tents and hospitals (Haycock, 2002).

The *Febbre biliosa* (bilious fever), as reported by a treater of the nineteenth century, is a disease that interests the abdominal venous system (in particular, the hepatic portal vein). The first symptoms were tiredness, lack of appetite, sickness, and headache. But for some subjects, the first phase of the illness was utterly asymptomatic. Symptoms in the second time include cold and shivering, especially during the afternoon, and then feeling hot accompanied by thirst, headache, difficulty breathing, increased heart rate, and pins and needles. Symptoms in subsequent phrases include:

- Sparkling eyes with dilated pupils.
- The presence of blood.
- Confusion.
- Strange colors of the tongue (black, yellow, or bright red with furrows).
- Pale or blackish lips.
- Ringing in the ears.

It is reported that generally healthy subjects were hit: more frequently, men, followed by women, and finally children (Meli, 1837).

A tract of the eighteenth century about the *Febbri biliose* that describes the epidemic disease in Lausanne (Switzerland) in 1755 reports that this fever belongs to the class of *febbri putride* (putrid fever). Breathing problems were frequently attributed to *Febbri putride*, as were foods of animal origin (prone to putrefaction), and the corruption of the bile (one of the four humors) that is also prone to corruption (Tissot, 1772).

Interestingly, today the term *Febbre biliosa* is referred to "*Febbre biliare emoglobinurica*" a hemolytic complication of malaria. The symptoms are abdominal pain, bilious vomiting, and the presence of hemoglobin in the urine (*FEBBRE BILIARE EMOGLOBINURICA - Dizionario Medico*, n.d.), but the most common interpretation of this cause of death among historians of medicine is typhus fever. Another interpretation is possible: bilious fever is characterized by bile — thus, its victims were jaundiced. With jaundice as the principal symptom, the differential diagnosis would include malaria, yellow fever, and dengue fever. The former was indeed a problem for those involved in rice cultivation, or living near areas where mosquitos were abundant.

The *Febbre Catarrale* is probably another name for flu (Vicentini et al., 2015).

Historically, it is often difficult to be sure that an epidemic is due to influenza. Still, five signs are indicators of the nature of the illness: sudden onset in a place and all cases in one place occur in one or two months, rapid diffusion in an area, high morbidity and low mortality (but the overall mortality can be higher according to the period and the area (for example, during the pandemic of 1729-1730, in London the overall mortality rate was doubled)), hit all social classes and subjects of all age, but leads to death mostly elderly people and who have chronic illnesses, and the disease resolves itself in a few days (with death or recovery) (Beveridge, 1991).

It is challenging to have a high certainty that some epidemics were due to flu during the sixteenth century: in 1551, Caius, the English physician, reported an epidemic characterized by headache, fever, and myalgia. Was it flu or not? (Cunha, 2004).

In 1588, there was the first recognized influenza pandemic: from Asia, the illness interested Africa, and, after, Europe and America). Everywhere it lasted six weeks. During the eighteenth century, there were 13 epidemics, and in the nineteenth century, 12 (Ghendon, 1994).

Figure 31 presents the local and global Kernel smoothing method for the time series 1755-1801.

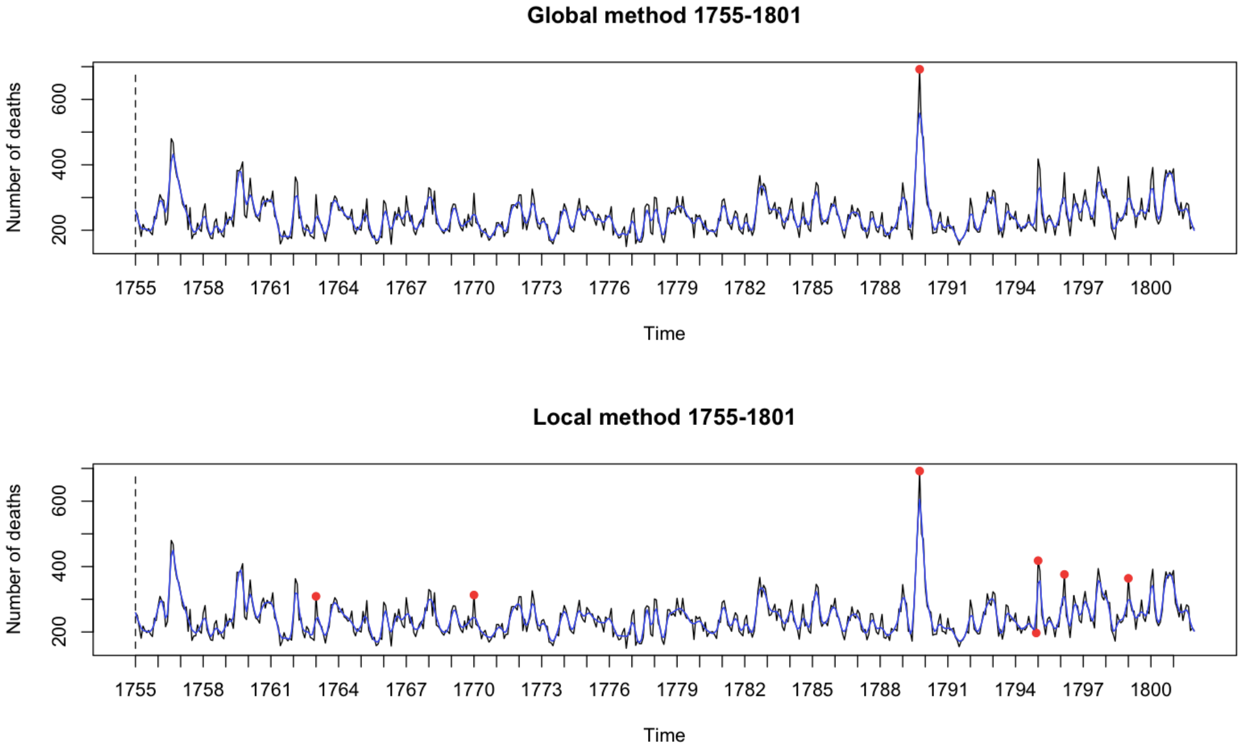


Figure 31: local and global Kernel smoothing for the time series from 1755 to 1801. The points in red are the outliers found by the two methods. The black line represents the original series, the blue line represents the Kernel smoothing series, and the vertical bars represent the changepoints.

Figure 32 presents the autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1755 -1801.

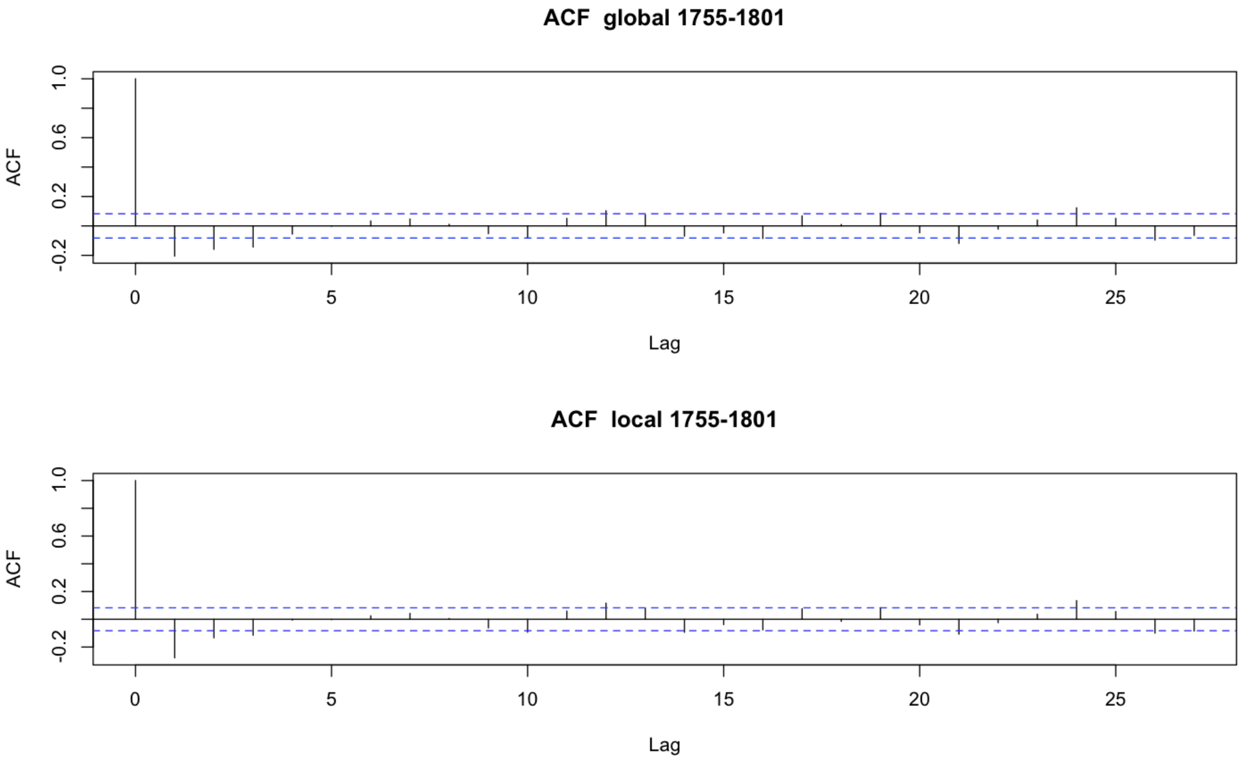


Figure 32: autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1755-1801.

Table 33 reports the outliers. The events in the outliers' date (or years), according to Corradi (Corradi, 1972), were reported.

LOCAL METHOD		**GLOBAL METHOD**		**EVENTS**
OUTLIERS	NUMBER OF DEATHS (Original series)	OUTLIERS	NUMBER OF DEATHS (Original series)	
1759-11-01	246			
1770-01-01	313			
1773-12-01	163			
1777-11-01	193			
1789-10-01	692	1789-10-01	692	Typhus
1796-03-01	376			Typhus

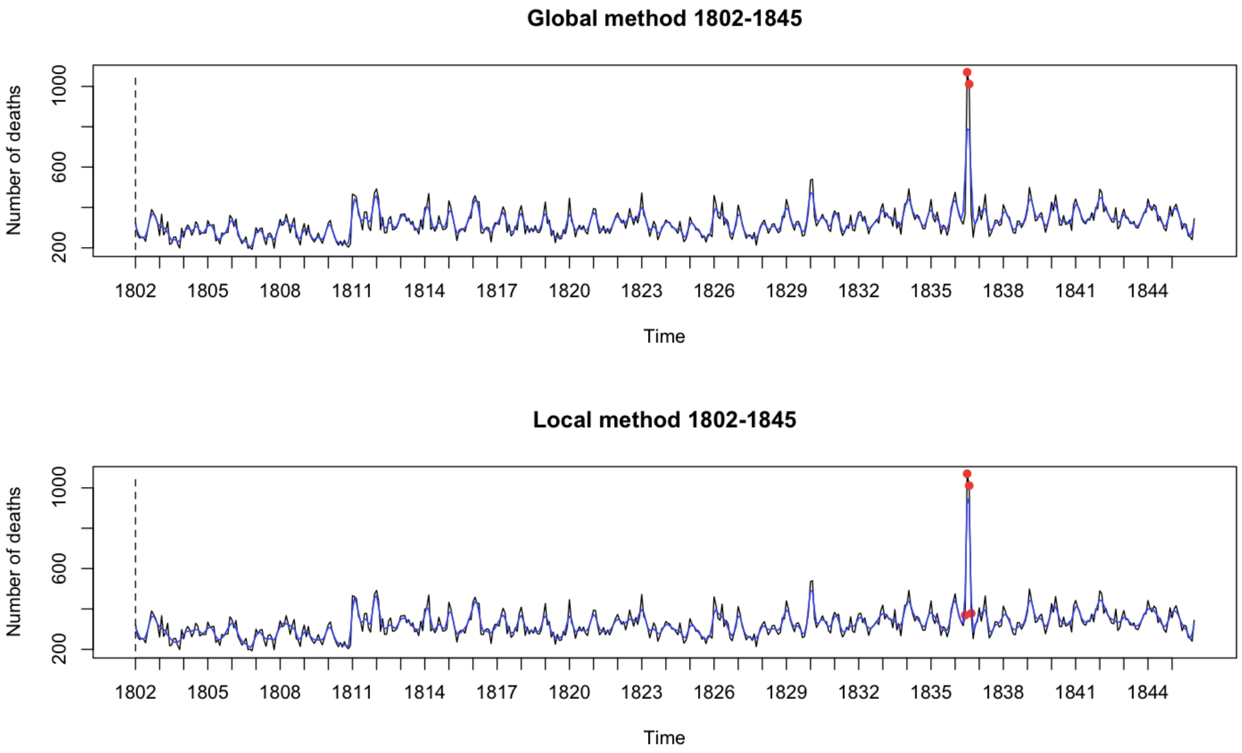
Table 33: outliers in the time series 1755-1801.

Typhus (likely called putrid fever) caused massive epidemics in Italy in 1590-3, 1648-9, 1671, 1764, and 1817-18. In general, epidemics happen in winter, when people wear heavy clothes and are in poor sanitary conditions. The events occurred after wars, climate events, and a bad harvest that cannot be repaired thanks to commercial activitie, and demographic upheavals.

Typhus is caused by *Rickettsia prowazekii*, a gram-negative bacterium transmitted by the flea *Pediculus humanus corporis*. The transmission of the bacterium does not occur directly due to the flea bite but through the contamination by flea feces. The itching caused by the animal's bite and the resulting injury from scratching causes the animal's excrement to enter the body. The incubation lasts 10-14 days, and two weeks after the incubation period, the patient's patient stabilizes or dies. The patient may never heal because of the Brill-Zinser disease: after the healing of the patient's symptoms, the bacterium can remain in the body in a latent form, even for life. With a lowering of the immune defenses, it can pass into an active state. Usually, the Brill-Zinser disease does not lead to death but can cause a new epidemic event in the population. Symptoms of typhus include rash and vasculitis on the skin, heart, central nervous system, skeletal muscles, and kidneys.

If the infection of *Rickettsia prowazekii* is not treated it causes death in 13%-30% of cases (Hanlon, 2000; Raoult et al., 2004).

Figure 33 presents local and global Kernel smoothing for the time series from 1802 to 1845



Global method 1802-1845

Local method 1802-1845

Figure 33: local and global Kernel smoothing for the time series from 1802 to 1845. The points in red are the outliers found by the two methods. The black line represents the original series, the blue line represents the Kernel smoothing series, and the vertical bars represent the changepoints.

Figure 34 presents the autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1802 to 1845.

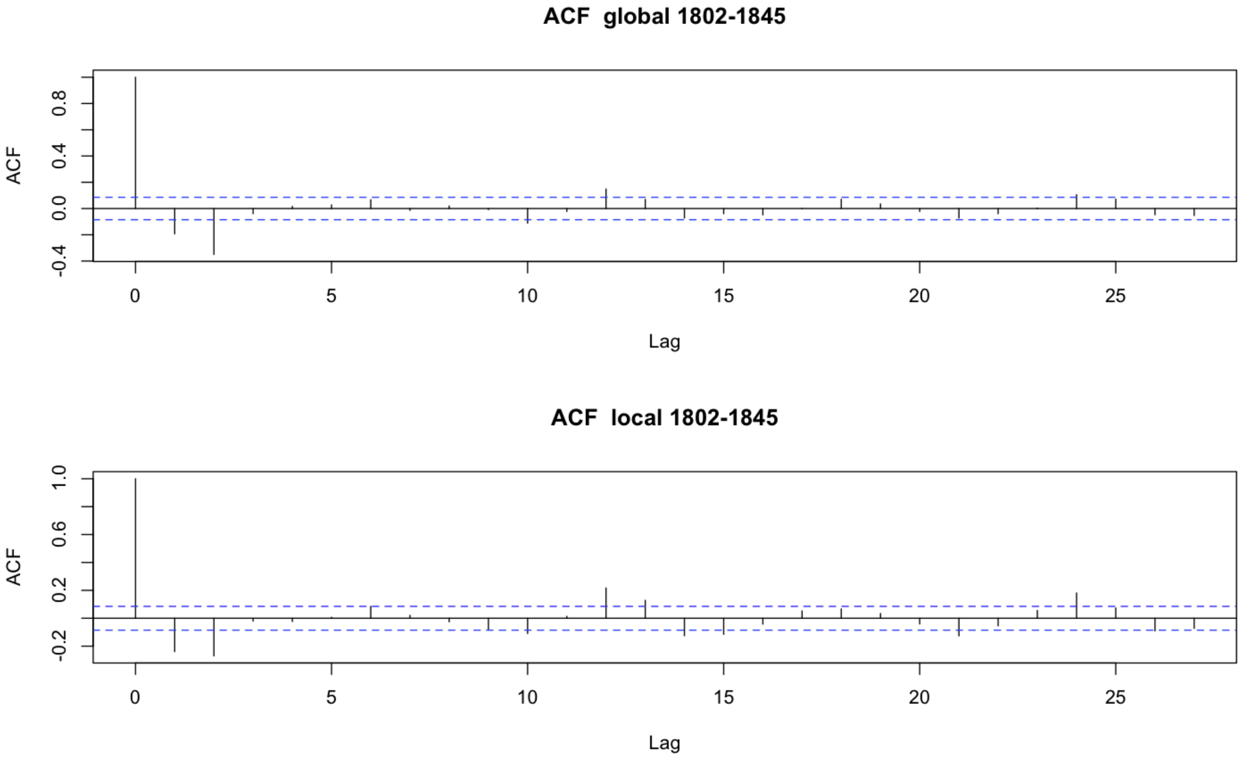


Figure 34: autocorrelation plot of residuals global and local Kernel smoothing method for the time series 1802 to 1845.

Table 34 reports the outliers. The events in the outliers' date (or years), according to Corradi (Corradi, 1972), were reported.

LOCAL METHOD		**GLOBAL METHOD**		**EVENTS**
OUTLIERS	NUMBER OF DEATHS (Original series)	OUTLIERS	NUMBER OF DEATHS (Original series)	
1836-06-01	369			Cholera
1836-07-01	1070	1836-07-01	1070	Cholera
1836-08-01	1011	1836-08-01	1011	Cholera
1836-09-01	378			Cholera

Table 34: outliers in the time series 1802-1845.

Cholera is caused by the bacterium *Vibrio cholerae*, particularly by the serogroups O1 and O139. It is a bacterium that lives in an aquatic environment. It can be transmitted directly from human to human with the fecal-oral route or indirectly with contaminated fluid. The pattern of infection is clustered due to direct contamination or contaminated local environment (i.e., the case of the epidemics of Broad Street in 1854 because of using a water pump that takes the liquid from a contaminated source) (Deen et al., 2020).

Symptoms include diarrhea with yellowish feces, nausea (without vomiting), weight loss, thirst, and a white tongue. Subsequent symptoms comprise vomiting, physical and mental debilitation, pallor, sweating, decreased urination, and the aggravated symptoms of the first phrase (Stillé, 1885)

An analysis of the epidemic in Milan in 1836, thanks to the *Liber Mortuorum*, established that the most dramatic period was from the second half of July to the first week of August. The cholera mainly affected men, the elderly, and the artisans. The epidemic mostly hit the territories outside the city walls, but also some zones inside the walls: they were characterized by social backwardness (Faron, 1997)

6.4 Discussion

We applied a method proposed Čampulová for studying environmental time series to Ferrario's time series. The latter is based on Milano Sforza's registers of death, which contained detailed information about half a million individuals in Milan. The registration covered the period from 1452 to 1801, and so this is the time span of the time series. Moreover, data from 1802 to 1845 were recovered from the Trivulziana Library.

Only data from 1581 was considered because the algorithm does not admit the presence of missing values.

The time series was divided into periods of approximately 50 years, and the “calendar adjustment” was performed. The local and global algorithms on each time series were performed. Changepoints and outliers were found and explained according to Corradi’s volumes and other historical sources that report socio-demographic information or non epidemic events.

Considering the first time series (1581-1564), the distribution of changepoints is more reliable, particularly considering the plague period: the plague period is divided by the rest of the series, and this also occurs for the period immediately following the epidemic. The behavior of the last parts of the series, delimited by changepoints, could be due to demographic increase and, therefore, an increase in the number of deaths. The selective effect of the plague was no longer decisive in this part. The first part of the series (delimited by changepoints) probably was highly impacted by the crop failure and economic problems affecting the city. The local method subdivided the series into three parts, but the plague period is not divided from the previous period of the series, and is this has yet to be judged as reliable.

The subdivision of this series according to global methods was:

1)From 01/1581 to 09/1591 2) From 10/1591 to 11/1626 3) From 12/1626 to 11/1631 4) From 12/1631 to 04/1643 5)From 05/1643 to 12/1654.

There were many more outliers identified by global methods than those determined by the local method: 11 vs. 7. Parts of outliers identified by the global method are in the first part of the series (02/1581, 06/1584, 09/1588, and 09/1592: according to the documents that Corradi compiled, there were no epidemic events in that period, but other historical source referred the presence of crop failure and economic problems that could be the cause of those outliers. Those outliers were not identified by using the local method. Local and global methods identified outliers in the plague period (5/1630 and 6/1630). For subsequent periods, local and global methods identified five outliers, but only three were in common: 09/1631, 08/1636, and 08/1643: it was not possible to find the events associated with those outliers.

Considering the second time series (1655-1705), the local and global method did not identify any changepoints, so the series is slightly homogeneous. The global method identified ten outliers, and the local method identified five. Some outliers identified by the global method (07/1693 and 09/1693), according to Corradi’s work, could be explained by smallpox epidemics. Those outliers were not identified by the local method, so it is possible that the global method is more reliable for this time series.

Considering the third time series (1706-1754), the global method subdivided the series into two parts according to changepoints: 1) from 01/1706 to 12/1720 2) from 01/1721 to 12/1754. The global method identified four outliers, and the local method identified six outliers. The outliers in the period 1719-1730 (of which three were identified by both the local and the global method: 12/1719, 02/1730, 01/1743, and one only by local method 01/1719) could be due to dysentery, smallpox, *Febbre biliosa*, and flu.

The global and local methods identified a low number of outliers: four according to the global method and six according to the local method.

Considering the fourth series (1755-1801), the local and global methods did not identify any changepoints. The global method identified one outlier (10/1789), and the local method identified six outliers. Outliers in 1789 and 1796 could be due to typhus.

The fifth time series (1802 to 1845) refers, probably, to another source on the number of deaths in the city of Milan: they come from the collections of the Civic Historical Archive (Trivulziana Library), where the registry sources of the Municipality of Milan are kept.

The local and global methods did not find any changepoints in the series. Both methods identify a low number of outliers: two for the global method (07/1836 and 08/1836) and four for the local method (06/1836, 07/1836, 08/1836, and 09/1836). All those outliers could be due to Cholera. Interestingly, all the outliers occurred in the same year.

6.5 Conclusions

Giuseppe Ferrario (1802-1870), in his work, "*Statistica medica di Milano dal secolo 15esimo fino ai giorni nostri*", reported very interesting information about the history of Milan. A table that reported the monthly numbers of deaths according to the *Mortuorum Liber* and other demographical sources from 1802- 1845 was considered for the analysis. The package envoutliers performed the analysis on this time series. Outliers were investigated mainly according to Corradi's work and other sources about socio-demographic phenomena.

The significance of the outliers could be investigated according to the causes of death listed in the *Mortuorum Liber*.

Monthly data characterize data of Ferrario's time series. Still, in Corradi's book, information about epidemics is organized in years, so the presence of an outlier during an epidemic year cannot be due to the epidemic phenomenon.

7 CONCLUSIONS

My Ph.D. project was about the *Mortuorum Libri* of Milan. This source contains detailed information about one and half a million subjects who lived in the city of Milan (or who, in any case, were in the city only temporarily, for example, for commercial reasons) and covers the period 1452-1801. Socio-demographical information included name, age, social class, and job (the last ones are not present in all the registers). From the name, it can be inferred the family the subject belongs to and the gender. Geographical information indicates where the subject lived or where the event happened. Clinical information was about the cause of death or the nature of the illness. In the registry, there is an indication of the *Necroscopo*, the crucial figures who inspected the corpses of people who did not die in hospital or who was not under the care of a community physician or surgeon.

The study of the registers involves people from different disciplines: historians, anthropologists, epidemiologists, and experts in the history of medicine. The information in them could also be important for the people who live in Milan today: they can discover the history of the city made by "normal people" and not the rulers. Citizens could become emotionally attached to their neighbors from 300 years ago. For all these reasons, making information accessible is mandatory. REDCap, a web application to construct databases and store data securely, was chosen because it is free for the universities and it is relatively simple to use. REDCap can store the data in a format that allows the analysis from software (such as R or GIS). In the future, we do not exclude to put them online and accessible to everyone.

A first analysis was made on 1480 data, a period free from plague and other epidemics and characterized by peace. Spatio-temporal analysis on the death density in different *Contrade* was performed. From this analysis, it was clear that children greatly impacted the distribution. During the Renaissance, the children mortality was high, and it had a strict connection with the socioeconomic situation in an area. We also give detailed information on:

Accidental and violent death (suicide or homicide)
Deaths in the hospital
Deaths in the elderly
Deaths in women
Deaths of orphans

The year 1630 was a major plague year: a spatiotemporal reconstruction of it was performed on the data of *Mortuorum Libri*, also considering an event that, according to the chronicles of the period, had a significant impact on the evolution of the epidemic situation: a religious profession featuring the remains of San Carlo Borromeo, which Church authorities organized hoping to mitigate the epidemic, took place in early July. The first plague event happened after the procession for 63 out of 94 parishes (67%). We found two different spatial patterns of deaths, based on the Principal Coordinates Analysis and the Silhouette method.

The epidemiological curves of the parishes of the clusters were considered. The parishes of Cluster 1 experienced the first plague death significantly earlier than those of Cluster 2. Similarly, the parishes of Cluster 1 reached 25% and 50% of their total plague deaths earlier than the parishes of Cluster 2. Lastly, the parishes of the two clusters significantly differ in the date of the inflection point (the date at which the curve changes concavity) of their epidemiological curves. The two clusters showed no significant difference in global deaths (plague and nonplague deaths).

The results raised some questions: How could a procession greatly impact the dynamics of an epidemic? The hypothesis of transmission to humans by vectors is still debated.

The *Mortuorum Liber* of 1485, another severe plague event, was considered to perform analyses on the association between gender, age, symptoms, and death events. This analysis was possible because in the registry there were also subjects with plague signs, but still alive at the time of the initial report. The only symptom with a positive association with the outcome is the *Morbilli* (chronicles confirm this), but there is, to our knowledge, no explanation for the reason the *Bubboni, Carboni*, and *Morbilli* seem to be protective for death. The association between different symptoms is scarce, and in fact the majority of subjects have only one symptoms. The positions of symptoms in different parts of the body were also evaluated.

The median time from symptoms to death was 3 days.

Further analysis will involve the spatio-temporal dynamics of this epidemic event and the association between the plague deaths and the professions of the subjects.

A time series analysis based on Giuseppe Ferrario's data (monthly number of deaths from 1450 to 1845) was performed. The study was performed with a method also used in ecological data (similarly to historical data, they had been impacted by external events). The series was divided into five periods of about 50 years. Global and local kernel smoothing was performed for each piece of the time series, and we obtained changepoints and outliers. Changepoints were investigated according to some demographical information, and outliers were investigated according to Corradi's work and other sources. There are some "unexplained" outliers, and the meaning can be given by querying the *Mortuorum Liber*: it can allow the discovery of other epidemiological events for which there is little study or mention. The *Mortuorum Liber* can be used to analyze the outliers (for example, the data on a year characterized by Cholera can be used to investigate if this illness hit some age classes or some social categories. Moreover, a spatio-temporal analysis can be performed to discover which city zones were hit and the order in which they were affected).

8 REFERENCES

Albini, G. (1982). *Guerra, fame, peste: crisi di mortalità e sistema sanitario nella Lombardia tardomedioevale*. Cappelli.

Albini, G. (2012). Dall'abbandono all'affido : storie di bambini nella Milano del tardo Quattrocento. *Mélanges de l'École Française de Rome. Italie et Méditerranée*, *124–1*. https://doi.org/10.4000/mefrim.243

Alfani, G. (2013). Plague in seventeenth-century Europe and the decline of Italy: an epidemiological hypothesis. *European Review of Economic History*, *17*(4), 408–430.

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

Alfani, G., & Percoco, M. (2019). Plague and long-term development: the lasting effects of the 1629–30 epidemic on the Italian cities. *The Economic History Review*, *72*(4), 1175–1201.

Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, *26*(1), 32–46.

Anselin, L. (1995). Local indicators of spatial association—LISA. *Geographical Analysis*, *27*(2), 93–115.

Antonielli, L. (2008). Una disputa corporativa nella Milano di fine settecento: anziani versus sepoltori. In *Con la ragione e col cuore: studi dedicati a Carlo Capra* (pp. 341–367). FrancoAngeli.

Appelhans, T., Detsch, F., Reudenbach, C., & Woellauer, S. (2021). *mapview: Interactive Viewing of Spatial Data in R* (R package version 2.10.0).

Ayyadurai, S., Sebbane, F., Raoult, D., & Drancourt, M. (2010). Body lice, Yersinia pestis orientalis, and black death. *Emerging Infectious Diseases*, *16*(5), 892.

Bairoch, P., Batou, J., & Pierre, C. (1988). *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)* (Vol. 4). Librairie Droz.

Bamji, A. (2019). Marginalia and mortality in early modern Venice. *Renaissance Studies*, *33*(5), 808–831. https://doi.org/10.1111/REST.12551

Barbieri, R., Drancourt, M., & Raoult, D. (2021). The role of louse-transmitted diseases in historical plague pandemics. *The Lancet Infectious Diseases*, *21*(2), e17–e25.

Barbieri, R., Nodari, R., Signoli, M., Epis, S., Raoult, D., & Drancourt, M. (2022). Differential word expression analyses highlight plague dynamics during the second pandemic. *Royal Society Open Science*, *9*(1), 210039.

Barbieri, R., Signoli, M., Chevé, D., Costedoat, C., Tzortzis, S., Aboudharam, G., Raoult, D., & Drancourt, M. (2020). Yersinia pestis: the natural history of plague. *Clinical Microbiology Reviews*, *34*(1), 10–1128.

Bertoglio, T. (2021). Indagini mediche e medici cittadini nella Milano tardomedievale: uno studio dei Libri Mortuorum del 1478. *Studi Di Storia Medioevale e Di Diplomatica-Nuova Serie*, 189–211.

Beveridge, W. I. B. (1991). The chronicle of influenza epidemics. *History and Philosophy of the Life Sciences*, 223–234.

Bivand, R. (2020). *classInt: Choose Univariate Class Intervals* (R package version 0.4-3).

Bivand, R., Keitt, T., & Rowlingson, B. (2022). *rgdal: Bindings for the "Geospatial" Data Abstraction Library* (R package version 1.5-29).

Bivand, R., & Lewin-Koh, N. (2022). *maptools: Tools for Handling Spatial Objects* (R package version 1.1-3).

Bivand, R. S., & Wong, D. W. S. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, *27*(3), 716–748. https://doi.org/10.1007/s11749-018-0599-x

Black, J. (1999). *Eighteenth Century Europe, 1700-1789*. Bloomsbury Publishing.

Blaeu, J. (1704). *Nouveau théâtre d'Italie, ou, description exacte de ses villes, palais, eglises, &c. et les cartes geographiques de toutes ses provinces*. P. Mortier Libraire.

Blanc, G., & Baltazard, M. (1942). Rôle des ectoparasites humains dans la transmission de la peste. *Bull Acad Natl Med*, *126*, 446.

Boucheron, P. (2006). Milano ei suoi sobborghi: identità urbana e pratiche socio-economiche ai confini di uno spazio incerto (1400ca.-1550ca.). *Società e Storia*, *112*, 235–252.

Bourgeois-Pichat, J. (1951). La mesure de la mortalité infantile. I. Principes et méthodes. *Population (French Edition)*, 233–248.

Boyce, N. (2020). Bills of Mortality: tracking disease in early modern London. *The Lancet*, *395*(10231), 1186–1187. https://doi.org/10.1016/S0140-6736(20)30725-X

Brunsdon, C., & Chen, H. (2014). *GISTools: Some further GIS capabilities for R* (R package version 0.7-4).

Būntgen, U., Ginzler, C., Esper, J., Tegel, W., & McMichael, A. J. (2012). Digitizing historical plague. *Clinical Infectious Diseases*, *55*(11), 1586–1588.

Byass, P. (2007). Who Needs Cause-of-Death Data? *PLOS Medicine*, *4*(11), e333. https://doi.org/10.1371/JOURNAL.PMED.0040333

Čampulová, M., Michálek, J., Mikuška, P., & Bokal, D. (2018). Nonparametric algorithm for identification of outliers in environmental data. *Journal of Chemometrics*, *32*(5), e2997.

Capasso, G. (2021). *Identificazione di valori anomali nelle serie storiche per dati ambientali*. Università degli Studi di Padova.

Carmichael, A. G. (1991). Contagion theory and contagion practice in fifteenth-century Milan. *Renaissance Quarterly*, *44*(2), 213–256.

Carmichael, A. G. (2008). 1 Universal and Particular: The Language of Plague, 1348–1500. *Medical History*, *52*(S27), 17–52.

Carmichael, A. G. (2016). Registering deaths and causes of death in late medieval Milan. In *Death in Medieval Europe* (pp. 223–250). Routledge.

Carmichael, A. G. (2019). Epidemics and state medicine in fifteenth-century Milan. In *Medicine from the Black Death to the French disease* (pp. 221–247). Routledge.

Carmichael, A. G. (2022). The Forensic Tradition in Milan's Civic Mortality Registers. *Journal of Early Modern Studies*, *Continuous*.

Carmichael, A. G., & Silverstein, A. M. (1987). Smallpox in Europe before the seventeenth century: virulent killer or benign disease? *Journal of the History of Medicine and Allied Sciences*, *42*(2), 147–168.

Cernik, L. (2016). *Survival of the Lucky: Foundling Children in the Ospedale degli Innocenti*. http://hdl.handle.net/20.500.12648/1177

Cesana, D., Benedictow, O. J., & Bianucci, R. (2017). The origin and early spread of the Black Death in Italy: first evidence of plague victims from 14th-century Liguria (northern Italy). *Anthropological Science*, *125*(1), 15–24.

Chandler, T. (1987). *Four thousand years of urban growth: An historical census*. Mellen.

Chiu, R. (2018). Singing on the Street and in the Home in Times of Pestilence: Lessons from the 1576–78 Plague of Milan. In *Domestic devotions in early modern Italy* (pp. 27–44). Brill.

Cipolla, C. (1976). *Public Health and the Medical Profession in the Renaissance* . CUP Archive.

Cipolla, C. M. (1978). The 'bills of mortality'of Florence. *Population Studies*, *32*(3), 543–548.

Cohn, S. (2002). *The Black Death Transformed: Disease and Culture in Early Renaissance Europe*. Arnold.

Cohn, S. K. (2010). *Cultures of plague: medical thinking at the end of the Renaissance*. Oxford University Press.

Cohn, S. K., & Alfani, G. (2007). Households and Plague in Early Modern Italy. *The Journal of Interdisciplinary History*, *38*(2), 177–205. http://www.jstor.org/stable/20143583

Corradi, A. (1972). Corradi Alfonso: Annali delle epidemie occorse in Italia dalle prime memorie fino al 1850, five volumes. *Società Medico-Chirurgica Di Bologna (Ed.), Forni, Bologna, Italy (in Italian)*.

Crespi, M. (1996). *FERRARIO, Giuseppe*. Dizionario Biografico Degli Italiani - Volume 46. https://www.treccani.it/enciclopedia/giuseppe-ferrario_%28Dizionario-Biografico%29/

Cummins, N., Kelly, M., & Ó Gráda, C. (2016). Living standards and plague in L ondon, 1560–1665. *The Economic History Review*, *69*(1), 3–34.

Cunha, B. A. (2004). Influenza: historical aspects of epidemics and pandemics. *Infectious Disease Clinics*, *18*(1), 141–155.

D'Amato, A., Zilberstein, G., Zilberstein, S., Compagnoni, B. L., & Righetti, P. G. (2018). Of mice and men: traces of life in the death registries of the 1630 plague in Milano. *Journal of Proteomics*, *180*, 128–137.

D'Amico, S. (1994). *Le contrade e la città. Sistema produttivo e spazio urbano a Milano fra Cinque e Seicento*. FrancoAngeli.

D'Amico, S. (2000). Crisis and transformation: economic organization and social structures in Milan, 1570-1610. *Social History*, *25*(1), 1–21.

D'Amico, S. (2001). Rebirth of a city: Immigration and trade in Milan, 1630-59. *Sixteenth Century Journal*, 697–721.

d'Amico, S. (2012). *Spanish Milan: A city within the Empire, 1535-1706*. Springer.

D'Andrea, D. M., & Terpstra, N. (2017). INFANT MORTALITY AND FOUNDLING HOMES. *The World of Renaissance Italy: A Daily Life Encyclopedia [2 Volumes]*, *366*, 263–274.

Davis, D. E. (1986). The scarcity of rats and the Black Death: an ecological history. *The Journal of Interdisciplinary History*, *16*(3), 455–470.

Dean, K. R., Krauer, F., & Schmid, B. V. (2019). Epidemiology of a bubonic plague outbreak in Glasgow, Scotland in 1900. *Royal Society Open Science*, *6*(1), 181695.

Dean, K. R., Krauer, F., Walløe, L., Lingjærde, O. C., Bramanti, B., Stenseth, N. C., & Schmid, B. V. (2018). Human ectoparasites and the spread of plague in Europe during the Second Pandemic. *Proceedings of the National Academy of Sciences*, *115*(6), 1304–1309.

Declich, S., & Carter, A. O. (1994). Public health surveillance: historical origins, methods and evaluation. *Bulletin of the World Health Organization*, *72*(2), 285.

Deen, J., Mengel, M. A., & Clemens, J. D. (2020). Epidemiology of cholera. *Vaccine*, *38*, A31–A40.

della Somaglia, C. G. C. (n.d.). *Alleggiamento dello Stato di Milano per le imposte, e loro ripartimenti*. nella Reg. Duc. Corte, per Gio, Battista, e Giulio Cesare fratelli Malatesta ….

DeWitte, S. N. (2014). Mortality risk and survival in the aftermath of the medieval Black Death. *PloS One*, *9*(5), e96513.

Di Bello, A. (2017). *CONTRIBUTO A UNA BIBLIOGRAFIA DELLE CHIESE DI MILANO*. https://milano.biblioteche.it/sites/milano/assets/www/SormaniStoriaLocale/MilanoStoriaLocaleChiese.pdf

Donovan, M. (1849). Report on the plague of Milain in 1630. *Dublin Quarterly Journal of Medical Science*, *8*(1), 217–232.

Drali, R., Shako, J.-C., Davoust, B., Diatta, G., & Raoult, D. (2015). A new clade of African body and head lice infected by Bartonella quintana and Yersinia pestis—Democratic Republic of the Congo. *The American Journal of Tropical Medicine and Hygiene*, *93*(5), 990.

Drevenstedt, G. L., Crimmins, E. M., Vasunilashorn, S., & Finch, C. E. (2008). The rise and fall of excess male infant mortality. *Proceedings of the National Academy of Sciences*, *105*(13), 5016–5021.

Duncan, C. J., & Scott, S. (2005). What caused the black death? *Postgraduate Medical Journal*, *81*(955), 315–320.

Durbin, J., & Watson, G. S. (1950). Testing for serial correlation in least squares regression: I. *Biometrika*, *37*(3/4), 409–428.

DURBIN, J., & WATSON, G. S. (1951). TESTING FOR SERIAL CORRELATION IN LEAST SQUARES REGRESSION. II. *Biometrika*, *38*(1–2), 159–178. https://doi.org/10.1093/biomet/38.1-2.159

Eckert, E. A. (1980). Seasonality of plague in early modern Europe: Swiss epidemic of 1628–1630. *Reviews of Infectious Diseases*, *2*(6), 952–959.

Farinelli, G. (1988). *Processo agli untori: Milano 1630: cronaca e atti giudiziari in edizione integrale*. Garzanti.

Faron, O. (1997). Le choléra à Milan en 1836. *Annales de Démographie Historique*, 89–114.

Fass, P. S. (2013). *The Routledge history of childhood in the western world*. Routledge.

FEBBRE BILIARE EMOGLOBINURICA - Dizionario medico. (n.d.). Retrieved September 29, 2023, from https://www.dica33.it/dizionario-medico/24584/febbre-biliare-emoglobinurica.asp

Ferrario, G. (1834). *Statistica delle morti improvvise e particolarmente delle morti per apoplessia nella città e nel circondario esterno di Milano dall'anno 1750 al 1834*. Imperial Regia stamperia.

Ferrario, G. (1840). *Statistica medica di Milano dal secolo 15. fino ai nostri giorni* (Vol. 2). Guglielmini & Redaelli.

Ferrario, G. (1846). *Cenni storici e regolamento della Accademia fisio-medico-statistica di Milano* (Redaelli, Ed.).

Finlay, R. (1981). Population and Metropolis: The Demography of London 1580–1650. In *Cambridge Geographical Studies*. Cambridge University Press. https://doi.org/DOI: 10.1017/CBO9780511735325

Fletcher-Lartey, S. M., & Caprarelli, G. (2016). Application of GIS technology in public health: successes and challenges. *Parasitology*, *143*(4), 401–415.

Forbes, T. (1974). The searchers. *Bulletin of the New York Academy of Medicine*, *50*(9), 1031–1038.

Foucault, M. (2014). The politics of health in the eighteenth century. *Foucault Studies*, 113–127.

Frith, J. (2012). The history of plague - part 1: The three great pandemics. *Journal of Military and Veterans Health*, *20*(2), 11–16. https://search.informit.org/doi/10.3316/informit.722091776908373

Galli, M., Nodari, R., Perini, M., Luconi, E., Fois, L., Vaglienti, F., Bandi, C., Biganzoli, E., & Comandatore, F. (2023). A spatiotemporal reconstruction of the 1630 plague epidemic in Milan. *Iscience*, *26*(5).

Galli, M., Oreni, L., Ridolfo, A. L., Formenti, A., Luconi, E., Boracchi, P., Antinori, S., Biganzoli, E., & Vaglienti, F. (2023). Milan's forgotten epidemic of summer 1629, a few months before the last great plague: An investigation into the possible cause. *Plos One*, *18*(6), e0279218.

Garcia, K. K. S., & Abrahão, A. A. (2021). Research Development Using REDCap Software. *Healthcare Informatics Research*, *27*(4), 341–349.

Gasser, T., Kneip, A., & Köhler, W. (1991). A flexible and fast method for automatic smoothing. *Journal of the American Statistical Association*, *86*(415), 643–652.

Gavitt, P. (1990). *Charity and children in Renaissance Florence: the Ospedale degli Innocenti, 1410-1536*. University of Michigan Press.

Ghendon, Y. (1994). Introduction to pandemic influenza through history. *European Journal of Epidemiology*, *10*, 451–453.

Giovanelli Onida, F., & Marinai, E. (2001). *Gride dei Conservatori della Sanità dello Stato di Milano (1576 - 1706)*. http://turing.ittig.cnr.it/sanita/sanitaIntroduzione.php#_ftn18

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, *53*(3–4), 325–338.

Gräler, B., Pebesma, E. J., & Heuvelink, G. B. M. (2016). Spatio-temporal interpolation using gstat. *R J.*, *8*(1), 204.

Gras, P. (1939). Le registre paroissial de Givry (1334-1357) et la peste noire en Bourgogne. *Bibliotheque de l'Ecole Des Chartes*, 295–308.

Graunt, J. (1676). *Natural and Political Observation Mentioned in a following index, and made upon the Bills of Mortality.* (J. Martyn, Ed.; Fifth). Royal Society, at the Sign of the Bell in St. Paul's Church-yard.

Griffin, L. J., & Isaac, L. W. (1992). Recursive regression and the historical use of "time" in time-series analysis of historical process. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *25*(4), 166–179.

Haines, M. R. (1995). Socio-economic differentials in infant and child mortality during mortality decline: England and Wales, 1890–1911. *Population Studies*, *49*(2), 297–315.

Hanlon, G. (2000). *Early modern Italy, 1550-1800; three seasons in European history*. Macmillan.

Harrell, F. E. (2017). Regression modeling strategies. *Bios*, *330*(2018), 14.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., & Kirby, J. (2019). The REDCap consortium: building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208.

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, *28*(1), 100–108. https://doi.org/10.2307/2346830

Haycock, D. B. (2002). Exterminated by the bloody flux. *Journal for Maritime Research*, *4*(1), 15–39.

Heitman, K. (2020). Authority, autonomy and the first London Bills of Mortality. *Centaurus*, *62*(2), 275–284. https://doi.org/10.1111/1600-0498.12305

Herrmann, E. (1997). Local bandwidth choice in kernel regression estimation. *Journal of Computational and Graphical Statistics*, *6*(1), 35–54.

Houhamdi, L., Lepidi, H., Drancourt, M., & Raoult, D. (2006). Experimental model to evaluate the human body louse as a vector of plague. *The Journal of Infectious Diseases*, *194*(11), 1589–1596.

Hufthammer, A. K., & Walløe, L. (2013). Rats cannot have been intermediate hosts for Yersinia pestis during medieval plague epidemics in Northern Europe. *Journal of Archaeological Science*, *40*(4), 1752–1759.

Humphrey, L., Bello, S., & Rousham, E. (2012). Sex differences in infant mortality in Spitalfields, London, 1750–1839. *Journal of Biosocial Science*, *44*(1), 95–119.

Introduction to Public Health Surveillance|Public Health 101 Series|CDC. (2018). https://www.cdc.gov/training/publichealth101/surveillance.html

ISTAT. (2014). *Anno 2012. LE PRINCIPALI CAUSE DI MORTE IN ITALIA*. https://www.istat.it/it/files/2014/12/Principali_cause_morte_2012.pdf

James, N. A., & Matteson, D. S. (2013). ecp: An R package for nonparametric multiple change point analysis of multivariate data. *ArXiv Preprint ArXiv:1309.3295*.

Jullien, S., de Silva, N. L., & Garner, P. (2021). Plague transmission from corpses and carcasses. *Emerging Infectious Diseases*, *27*(8), 2033.

King, M. L. (2008). *Women of the Renaissance*. University of Chicago Press.

Korstanje, J. (2021, July 4). *Principal Coordinates Analysis*. Towards Data Science.

Krylova, O., & Earn, D. J. D. (2020). Patterns of smallpox mortality in London, England, over three centuries. *PLoS Biology*, *18*(12), e3000506.

Kwok, R. (2017). Historical data: Hidden in the past. *Nature*, *549*(7672), 419–421.

La Croce, P. (1730). *Memorie delle cose notabili successe in Milano intorno al mal contaggioso l'anno 1630: del riccorso da signori della città a padri Capuccini per il governo del Lazzaretto...: de'portamenti d'essi padri in quelle calamità, e come entrasse la peste* *ne'conventi loro...: con in fine tre capitoli in compendio della purga della cose infette, e sospette usata*. Nelle stampe di Giuseppe Maganza.

Lathem, W. W., Crosby, S. D., Miller, V. L., & Goldman, W. E. (2005). Progression of primary pneumonic plague: a mouse model of infection, pathology, and bacterial transcriptional activity. *Proceedings of the National Academy of Sciences*, *102*(49), 17786–17791.

Lazzari, G., Colavizza, G., Bortoluzzi, F., Drago, D., Erboso, A., Zugno, F., Kaplan, F., & Salathé, M. (2020). A digital reconstruction of the 1630–1631 large plague outbreak in Venice. *Scientific Reports*, *10*(1), 17849.

Lewis, M. E., & Gowland, R. (2007). Brief and precarious lives: Infant mortality in contrasting sites from medieval and post-medieval England (AD 850–1859). *American Journal of Physical Anthropology: The Official Publication of the American Association of Physical Anthropologists*, *134*(1), 117–129.

MacMoreno. (n.d.). *Sestieri e Contrade di Milano*. Https://It.Wikipedia.Org/Wiki/File:Contrade_di_Milano.Svg#file.

Manfredi, M. F., & Getis, A. (2010). Spatial autocorrelation. In *Handbook of applied spatial analysis*. Spinger. Berlin.

Marinozzi, S. (2010). Ageing in the early modern age. *Medicina Nei Secoli*, *22*(1–3), 531–552.

Matteson, D. S., & James, N. A. (2014). A nonparametric approach for multiple change point analysis of multivariate data. *Journal of the American Statistical Association*, *109*(505), 334–345.

McEvedy, C. (1988). The bubonic plague. *Scientific American*, *258*(2), 118–123.

Meli, D. (1837). *Trattato Delle Febbri Biliose*.

Miarinjara, A., Bland, D. M., Belthoff, J. R., & Hinnebusch, B. J. (2021). Poor vector competence of the human flea, Pulex irritans, to transmit Yersinia pestis. *Parasites & Vectors*, *14*(1), 1–15.

Moore, Z. S., Seward, J. F., & Lane, J. M. (2006). Smallpox. *The Lancet*, *367*(9508), 425–435.

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society. Series B (Methodological)*, *10*(2), 243–251.

Morrison, A. S., Kirshner, J., & Molho, A. (1985). Epidemics in Renaissance Florence. *American Journal of Public Health*, *75*(5), 528–535.

Munoz-Tuduri, M., García-Moro, C., & Walker, P. L. (2006). Time series analysis of the epidemiological transition in Minorca, 1634–1997. *Human Biology*, *78*(5), 619–634.

Murianni, L., & Tinto, A. (2019). *Salute*. ISTAT. https://www.istat.it/it/files/2019/12/1.pdf

Naghavi, M., Makela, S., Foreman, K., O'brien, J., Pourmalek, F., & Lozano, R. (2010). Algorithms for enhancing public health utility of national causes-of-death data. *Population Health Metrics*, *8*, 9. http://www.pophealthmetrics.com/content/8/1/9

Nicolas, X., Granier, H., & Le Guen, P. (2007). Shigellose ou dysenterie bacillaire. *La Presse Médicale*, *36*(11), 1606–1618.

Nikiforov, V. V, Gao, H., Zhou, L., & Anisimov, A. (2016). Plague: Clinics, diagnosis and treatment. *Yersinia Pestis: Retrospective and Perspective*, 293–312.

Orzack, S. H., Stubblefield, J. W., Akmaev, V. R., Colls, P., Munné, S., Scholl, T., Steinsaltz, D., & Zuckerman, J. E. (2015). The human sex ratio from conception to birth. *Proceedings of the National Academy of Sciences*, *112*(16), E2102–E2111.

Otis, J. (2017). Death by Numbers: Bills of Mortality in Early Modern London. *DH*.

Patridge, E. F., & Bardyn, T. P. (2018). Research electronic data capture (REDCap). *Journal of the Medical Library Association: JMLA*, *106*(1), 142.

Pebesma, E. J. (2004). Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*, *30*(7), 683–691.

Peller, S. (1947). Studies on mortality since the renaissance. *Bulletin of the History of Medicine*, *21*(1), 51–101.

Peña, D. (2000). Outliers, influential observations, and missing data. *A Course in Time Series Analysis*, 136–170.

Perry, R. D., & Fetherston, J. D. (1997). Yersinia pestis--etiologic agent of plague. *Clinical Microbiology Reviews*, *10*(1), 35–66.

Pobst, P. (2013). Should we teach that the cause of the black death was bubonic plague? *History Compass*, *11*(10), 808–820.

Pollitzer, R. (1954). Plague. *Geneva: WHO*, 409–482.

Pongou, R., Kuate Defo, B., & Tsala Dimbuene, Z. (2017). Excess male infant mortality: The gene-institution interactions. *American Economic Review*, *107*(5), 541–545.

Queen Mary University of London. (n.d.). *REDCap Overview - REDCap Docs @ QMUL*. Retrieved December 27, 2023, from https://docs.redcap.qmul.ac.uk/

Randremanana, R., Andrianaivoarimanana, V., Nikolay, B., Ramasindrazana, B., Paireau, J., Ten Bosch, Q. A., Rakotondramanga, J. M., Rahajandraibe, S., Rahelinirina, S., & Rakotomanana, F. (2019). Epidemiological characteristics of an urban plague epidemic in Madagascar, August–November, 2017: an outbreak report. *The Lancet Infectious Diseases*, *19*(5), 537–545.

Raoult, D., Woodward, T., & Dumler, J. S. (2004). The history of epidemic typhus. *Infectious Disease Clinics*, *18*(1), 127–140.

Ratovonjato, J., Rajerison, M., Rahelinirina, S., & Boyer, S. (2014). Yersinia pestis in Pulex irritans fleas during plague outbreak, Madagascar. *Emerging Infectious Diseases*, *20*(8), 1414.

Riley, J. C. (1987). *The Eighteenth-Century campaign to avoid disease*. Macmillan.

Ripamonti, G. (n.d.). *Iosephi Ripamontii Canonici Scalensis Chronistae Vrbis Mediolani Historiae Patriae Decadis V libri VI*. Ex Regio Palatio Apud Io. Baptistam et Iulium Caesarem Malatestam Regios ….

Ripamonti, G. (1640). *Iosephi Ripamontii cononici scalensis chronistae urbis mediolani de peste quae fuit anno MDCXXX: Libri V : desumpti ex annalibus urbis quos LX. decurionum autoritate scribebat* . Apud Malatestas, Regios ac Ducales Typoraphos.

Rotta, P. (1891). *Passeggiate storiche, ossia Le chiese di Milano dalla loro origine fino al presente*. Tipografia del Riformatorio Patronato.

Rubini, M., Gualdi-Russo, E., Manzon, V. S., Rinaldo, N., & Bianucci, R. (2016). Mortality risk factors show similar trends in modern and historic populations exposed to plague. *The Journal of Infection in Developing Countries*, *10*(05), 488–493.

Salima, B. A., & Bellefon, M. D. (2018). Spatial autocorrelation indices. *Handbook of Spatial Analysis: Theory Aplication with R*, 51–68.

Schroth, C., Siebert, J., & Groß, J. (2021). *Time Traveling with Data Science: Focusing on Change Point Detection in Time Series Analysis (Part 2) - Blog des Fraunhofer IESE*. https://www.iese.fraunhofer.de/blog/change-point-detection/

Skog, L., & Hauska, H. (2013). Spatial Modeling of the Black Death in S weden. *Transactions in GIS*, *17*(4), 589–611.

Smith, R. S. (1936). Barcelona" bills of mortality" and population, 1457-1590. *Journal of Political Economy*, *44*(1), 84–93.

Spyrou, M. A., Tukhbatova, R. I., Feldman, M., Drath, J., Kacki, S., de Heredia, J. B., Arnold, S., Sitdikov, A. G., Castex, D., & Wahl, J. (2016). Historical Y. pestis genomes reveal the European Black Death as the source of ancient and modern plague pandemics. *Cell Host & Microbe*, *19*(6), 874–881.

Stillé, A. (1885). *Cholera: its origin, history, causation, symptoms, lesions, prevention, and treatment*. Lea Bros. & Company.

Stinson, S. (1985). Sex differences in environmental sensitivity during growth and development. *American Journal of Physical Anthropology*, *28*(S6), 123–147.

Swellengrebel, N. H., & Otten, L. (1914). *Experimentelle Beiträge zur Kenntnis der Uebertragung der Pest durch Flöhe und Läuse*. Fischer.

Tadino, A. (1648). Raguaglio dell'origine et giornali successi della gran peste contagiosa, venefica, & malefica seguita nella Città di Milano, & suo Ducato dall'anno 1629. sino all'anno 1632. Con le loro successiue prouisioni, et ordini. Aggiontoui un breue compendio delle più segnalate specie di peste in diuersi tempi occorse... Con diuersi antidoti descritti da Alessandro Tadino... In *per Filippo Ghisolfi*.

Tasca, G. (2013). *Storia della medicina. Dalla preistoria alla fine dell'Ottocento.* Temi Editrice.

Terpstra, N., & Rose, C. (2016). *Mapping Space, Sense, and Movement in Florence: Historical GIS and the early modern city*. Routledge.

Tissot, S. (1772). *Dissertazione intorno le febbri biliose o sia storia dell'epidemia biliosa di Losanna ne'Svizzeri dell'anno 1755* (G. Vitto, Ed.).

Tognotti, E. (2013). Lessons from the history of quarantine, from plague to influenza A. In *Emerging Infectious Diseases* (Vol. 19, Issue 2, pp. 254–259). Centers for Disease Control and Prevention (CDC). https://doi.org/10.3201/eid1902.120312

Tran, T.-N.-N., Signoli, M., Fozzati, L., Aboudharam, G., Raoult, D., & Drancourt, M. (2011). High throughput, multiplexed pathogen detection authenticates plague waves in medieval Venice, Italy. *PLoS One*, *6*(3), e16735.

Tsay, R. S. (1988). Outliers, level shifts, and variance changes in time series. *Journal of Forecasting*, *7*(1), 1–20.

Tucci, M. (2011). *For an urban history of Milan, Italy: The role of GIScience*. Texas State University-San Marcos.

Tulchinsky, T. H. (2018). John Snow, Cholera, the broad street pump; waterborne diseases then and now. *Case Studies in Public Health*, 77.

UK Health Security Agency. (2023). *Weekly all-cause mortality surveillance: 2022 to 2023*. https://www.gov.uk/government/statistics/weekly-all-cause-mortality-surveillance-2022-to-2023

University of Nebraska Medical Center. (2019). *A Comprehensive Guide to REDCap*. University of Nebraska Medical Center. https://www.unmc.edu/vcr/_documents/unmc_redcap_usage.pdf

Vaccaro, G. (2021). Marzo 1348. La fine del mondo tra paure e prevenzione nelle cronache volgari coeve. *RiMe. Rivista Dell'Istituto Di Storia Dell'Europa Mediterranea (ISSN 2035-794X)*, 139–164.

Vaglienti, F. (2020). Marginalia. Esempi di umane miserie nei Registri dei Morti di età sforzesca. *Studi Di Storia Medioevale e Di Diplomatica-Nuova Serie*, 383–400.

Vaglienti, F. I. M. (2013). Noluit ire ad visum: i segreti dei Mortuorum Libri di Milano da Francesco Sforza a Leonardo da Vinci:(1452-1485). *La Popolazione Di Milano Dal Rinascimento: Fonti Documentarie e Fonti Materiali per Un Nuovo Umanesimo Scientifico*, 25–59.

Vicentini, C. B., Guidi, E., Lupi, S., Maritati, M., Manfredini, S., & Contini, C. (2015). L'influenza nelle ondate epidemiche del XIX secolo. *Infez Med*, *23*(4), 374–389.

Visconti, A. (1911). *Il magistrato di sanità nello stato di Lombardia, in «Archivio Storico Lombardo»*.

Warnes, G., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., Schwartz, M., & Venables, B. (2022). *gplots: Various R Programming Tools for Plotting Data* (R package version 3.1.3).

Welaga, P., Moyer, C. A., Aborigo, R., Adongo, P., Williams, J., Hodgson, A., Oduro, A., & Engmann, C. (2013). Why are babies dying in the first month after birth? A 7-year study of neonatal mortality in northern Ghana. *PloS One*, *8*(3), e58924.

Welford, M. R., & Bossak, B. H. (2009). Validation of inverse seasonal peak mortality in medieval plagues, including the Black Death, in comparison to modern Yersinia pestis-variant diseases. *PloS One*, *4*(12), e8401.

Wheelwright, S., Makridakis, S., & Hyndman, R. J. (1998). *Forecasting: methods and applications*. John Wiley & Sons.

WHO. (2022). *Plague*. https://www.who.int/news-room/fact-sheets/detail/plague

Wong, D. W. S. (1999). Several fundamentals in implementing spatial statistics in GIS: Using centrographic measures as examples. *Geographic Information Sciences*, *5*(2), 163–174.

Woods, R. (2006). Mortality, Childcare and Mourning. In *Children Remembered. Responses to Untimely Death in the Past.* Liverpool University Press.

Younis, N. M. (2014). Assessment For Mortality Rate Children Under Five Years In Mosul City. *Journal of Kufa for Nursing Science Vol*, *4*(1).

Zanetti, D. E. (1976). La morte a Milano nei secoli XVI-XVIII. *Rivista Storica Italiana*, *4*, 803–851.

Zanoboni, M. P. (2020). *PESTE A MILANO: DAI PROVVEDIMENTI DI GIAN GALEAZZO VISCONTI ALLA PESTE DEL MANZONI* . Https://Milano.Biblioteche.It/Sites/Milano/Assets/Sormani/Allegati/LaPesteAMilano.Pdf.

Zarulli, V., Barthold Jones, J. A., Oksuzyan, A., Lindahl-Jacobsen, R., Christensen, K., & Vaupel, J. W. (2018). Women live longer than men even during severe famines and epidemics. *Proceedings of the National Academy of Sciences*, *115*(4), E832–E840.

Zeileis, A., & Hothorn, T. (2002). *Diagnostic checking in regression relationships* (Vol. 2, Issue 3). na.

Zhao, Q. F., & Yin, J. X. (2016). Analysis for the compositions of plague natural foci and their Roles. *Mod Prev Med*, *43*(2), 370–372.

SOFTWARE USED:

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

QGIS.org (2021). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.org

Berthold, Michael R., et al. "KNIME-the Konstanz information miner: version 2.0 and beyond." *AcM SIGKDD explorations Newsletter* 11.1 (2009): 26-31.

Acknowledgements

I would like to thank Prof. Elia Biganzoli, my Ph.D. tutor, for his support in these three years. I am also grateful to Prof. Patrizia Boracchi, my mother, for her support morally and in the analysis. I am thankful also to Giuseppe Marano for our endless discussion in the analyses. I appreciate your patience. I am grateful to Dott. Cristiana Panella for her precious collaboration on this project. I want to thank Prof. Massimo Galli, Dott. Folco Vaglienti, Dott. Luca Fois, Dott. Riccardo Nodari, Dott. Francesco Comandatore, and the team of 1630.

I am grateful to Dott. Luigi Compagnoni for let us the access to the *Archivio di Stato di Milano*.

I am grateful to Dott. Alex Bamji and Prof. Ann Carmichael agree to evaluate this dissertation.

I am grateful to Alessio Midali, my boyfriend, for his support in this period of my life (he stood by me in all my stress and my victories during the Ph.D.).

I am grateful to my dad, Daniele Luconi.

I thank my friends: Federica Gervasoni, Agnese Omini, Alessia Marelli, Valeria Porta, Walter Di Pinto, Andrea Magni, and Livia Lorenzon.

I am grateful to all member of All Saints Anglican Church in Milan for the support. I am grateful also to *Amici del Sidamo* and to San Egidio.

9 APPENDIX 1

The algorithm:
step 1: set $h_0 = (b-a)/n$
step 2 iterate $g_i = ch_{i-1}n^{2/(5\times 7)}$ and $h_i = \left(\frac{I_1 C(K)}{I_2(m,g_i)}\right)^{1/5}$ for i=,…,15
with $C(K) = \frac{\int (K(x))^2 dx}{\{x^2 K(x) dx\}^2}$
$I_1 = \sum_{i=1}^{n} w(x_i)\widehat{\sigma^2}(x_i) n (s_i - s_{i-1})^2$ $\widehat{\sigma^2}(x_i) = \frac{1}{a_i^2 + b_i^2 + 1}(Y_i - a_i Y_{i-1} - b_i Y_{i+1})^2$
where $a_i = \frac{x_{i+1} - x_i}{x_{i+1} - x_{i-1}}$ $and\ b_i = 1 - a_i$ $c = 1.1$
the local variance estimator is then smoothed.
$\hat{\sigma}^2(x_i) = \sum_{j=1}^{n} \widehat{\sigma^2}\ (x_j) \int_{s_{j-1}}^{s_j} \frac{1}{h_\sigma} K\left(\frac{x_i - u}{h_\sigma}\right) du$ $x_i \in [a + h_\sigma, b - h_\sigma]$

$$I_2 = \int w(x)[\widehat{m}\ \ (x, g_i)]^2 dx$$

Step3 set $\boldsymbol{h = h_{15}}$ the global plug-in bandtwidth.
Step 4 Set the pilot bandtwith $g_i = c_i h n^{2/35}$ for i=1,2
Where $c_1 = 0.86\ and\ c_2 = 1.4$

The local plug-in estimator is $\boldsymbol{h(x) = \varphi(x, g_1)\left(\frac{S(x)C(K)}{n\widehat{m}^2(x, g_1, g_2)}\right)^{\frac{1}{5}} + (1 - \varphi(x, g_1))h}$

With $\boldsymbol{S(t) = n\sum_{i=1}^{n} \hat{\sigma}^2(x_i)(s_i - s_{i-1}) \int_{s_{i-1}}^{s_i} \frac{1}{h} L\left(\frac{x-u}{h}\right) du}$ $\boldsymbol{x \in [a + h, b - h]}$ $\boldsymbol{and\ L(x) = \frac{\{K(x)\}^2}{\int \{K(y\}^2 dy}}$

With $\boldsymbol{\widehat{m}^2(x, g_1, g_2) = \int \frac{1}{g_2} K\left(\frac{x-u}{g_2}\right)\{\widehat{m}\ (u, g_1\}^2 du}$ and $\boldsymbol{\varphi}$ is a weight function

Details on the plugin estimator are reported in Herrmann, Eva. "Local bandwidth choice in kernel regression estimation." *Journal of Computational and Graphical Statistics* 6.1 (1997): 35-54.

Identification of change points

Regression residuals $R(X_i) = Y_i - \widehat{m}(X_i)$ are supposed to be independent observations with cumulative distribution F_1,…,F_n. The aim si to identify points $\tau_1, \ldots \tau_m$ (change points) that divide the residuals into homogeneous segments such that.
$F_1 = \cdots = F_{\tau_1} \neq F_{\tau_{1+1}} = \cdots = F_{\tau_2} = \cdots \neq \cdots \neq F_{\tau_m} = \cdots \neq F_{\tau_m + 1} = \cdots = F_n$
To simplify the notation in the rest of tha paragrah $R(X_i)$ will be indicated as R_i
The method assume the independence of R_i (i=1..n) and requires the absolute r-th moment exist $r \in (0,2)$ (the moment of the distribution is $E(R_i) = \int_{-\infty}^{\infty} {R_i}^r f(R_i) dr_i$)
Change points could result from distributions which change in mean, variance, shape of the tails.

Firstly, a change point $\boldsymbol{\tau}$ is searched with unknown location and the location is estimated.
The set of R_i (i=1..n) is suvided into two subsets:
$\boldsymbol{X}_\tau = \{R_1, \ldots R_\tau\}$ and $\boldsymbol{Y}_{\tau(k)} = \{R_{\tau+1}, \ldots, R_k\}$ where $1 \leq \tau < k \leq n$
k is introduced because of possible weakness of bisection method. In fact, fixing k=T, the presence of different distributions F in $\boldsymbol{Y}_\tau$ (resulting in a mixture of ditributions) could mask the presence of a valid change point.

The location of the change points(and the value of k) are estimated $(\widehat{\tau}, \widehat{k})$ as the values that maximise the distance between $\boldsymbol{X}_\tau$ and $\boldsymbol{Y}_{\tau(k)}$ for a given moment r $(Q(\boldsymbol{X}_\tau, \boldsymbol{Y}_{\tau(k)}, r)$

$$Q\left(\boldsymbol{X}_{\tau},\boldsymbol{Y}_{\tau(k)},r\right) = \frac{mo}{m+o}\frac{2}{mo}\sum_{i=1}^{o}\sum_{j=1}^{m}\left|X_i-Y_j\right|^r-\binom{o}{2}^{-1}\sum_{1\le i<k\le o}\left|X_i-Y_k\right|^r - \binom{m}{2}^{-1}\sum_{1\le j<k\le m}\left|Y_j-Y_k\right|^r$$

Where o is the number of R in $\boldsymbol{X}_{\tau}$, m is the number of R in $\boldsymbol{Y}_{\tau(k)}$, $\binom{o}{2}=\frac{o!}{2!(o-2)!}$ and $\binom{m}{2}=\frac{m!}{2!(m-2)!}$

A test statistic $\widehat{q_k}$ is calculated after using $(\widehat{\tau}, \widehat{k})$ in the equation $Q\left(\boldsymbol{X}_{\tau},\boldsymbol{Y}_{\tau(k)},r\right)$

The permutation test is applyed for the null hypotesis of no change point. The observation R are permuted to construct 499 random permutations. For each permutation the procedure for change point location above described is applyed and the test statistic is calculated. In this way the distribution of the test statistic under the null hypotesis is obtained. The p value is calculated as the ratio between the number of test statistics of the null distribution greater than $\widehat{q_k}$ and the total number of permutations.

If p value is less than 0.05 the null hypotesis is rejected and the algorithm search for additional change points.

Second hierarchically estimated multiple change points

In each one of the subset a new change point is found and $\hat{\tau}_i, \hat{k}_i$ (i=1,2) is estimated by the equation $Q\left(\boldsymbol{X}_{\tau},\boldsymbol{Y}_{\tau(k)},r\right)$ within the subset and the correpsonding value of $\hat{q}_{\hat{k}_i}$ is obtained. The statistic $\hat{q}_k$ is calculated after using in the equation $Q\left(\boldsymbol{X}_{\tau},\boldsymbol{Y}_{\tau(k)},r\right)$ within the subset in which the values of $\hat{\tau}_i, \hat{k}_i$ corresponds to the max ($\hat{q}_{\hat{k}_1}, \hat{q}_{\hat{k}_2}$)

An additional change point is then proposed for the i-th subset, and a permutation test is performed. The null hypotesis is that no additional change point is needed. The observation within each subset is permuted to construct a new sequence of length n. The sequence above reported to select the change point within the cluster is repeated for 499 permutations, for each permutation the statistic $\hat{q}_k$ is calculated and the distribution under null hypotesis is obtained. The p value is then calculated. If p value is less than 0.05 the null hypotesis is rejected and the algorithm search for additional change point. The procedure is applyed until a p value greater than 0.05 is obtained.

Presence of outliers in subsets identifyed by change points

The presence of outliers may result in subsets with a "small" number of residuals. The definition of "small" is given after fixing a pre-determined minimal number of observations a subset has to contain. The variance of the residuals on a subset with a small number of residuals is compared with the variance of the residuals of the neighbouring set (after removing the residual with the largest value in the "small" set) by a robust test on the homogeneity of variances. If the null hypotesis is not rejected the two sets are merged.

After this check, within each subset defined by changepoints, the outliers of the residuals are found by Chebyshev's inequality. Maximally $1/L^2$ residuals fall outside the interval -Ls, +Ls where s is the standard deviation of the distribution of residuals.

L is found for each interval by an iterative procedure.

1)The starting L (L_{start}) is 2.5 and a value of L_{max} is calculated such as the number of residuas greater than L_{max}*s is equal to 1.0

If L_{max} < no outliers were detected

2)Otherwise L is set to $L_{new}=L_{start}$*s and the set of the non-outliers residuals (< L_{new}*s) are ordered and the largest difference between couples of ordered residuals is calculated (the GAP)

3)L_{new} is then increased by a value e=(L_{max}-L_{start})/100) and the value of GAP for L_{new}+e (GAP_{new}) is calculated on the ordered non-outliers residuals

The procedure 3) on increased L_{new} values for each step is repeated untill no residuals are greater than L_{new}

4) the value of L corresponding to the maximum GAP among the GAP_{new} value calculated in each step is retained and residuals in the set greater than L are outliers