INEQUALITY IN MORTALITY IN PRE-INDUSTRIAL SOUTHERN EUROPE DURING AN EPIDEMIC EPISODE: SOCIO-ECONOMIC DETERMINANTS (EIGHTEENTH - NINETEENTH CENTURIES SPAIN[*][†]

ESTE ARTÍCULO HA SIDO PUBLICADO EN LA REVISTA ECONOMICS AND HUMAN BIOLOGY: Víctor A. Luque de Haro, Joana M. Pujadas-Mora, José J. García-Gómez, Inequality in mortality in pre-industrial southern Europe during an epidemic episode: socio-economic determinants (eighteenth - nineteenth centuries Spain), Economics & Human Biology, Volume 40, 2021, 100941, ISSN 1570-677X, https://doi.org/10.1016/j.ehb.2020.100941.

[*] Funding: This work has been funded by the Ministry of Science, Innovation and Universities of the Spanish Government through the projects PGC2018-097817-B-C32 and RTI2018-095533-B-I00.

[†] Authors have no competing interests to declare.

ABSTRACT:

The objective of this study is to gain more comprehensive knowledge about social inequality in mortality in pre-industrial periods. With this aim, we have reconstructed the life courses of the inhabitants of the town of Vera in south-east Spain for the period 1797-1812 in order to estimate the influence of socio-economic status on ordinary and extraordinary mortality, given that, during this period, the town suffered from several epidemic outbreaks of yellow fever. As a result of these outbreaks, around a quarter of the town's population died. The results obtained indicate social inequality in mortality at least from the end of the eighteenth century. Although the differences are higher in mortality caused by non-infectious diseases or ill-defined causes, the coefficients also show a certain social gradient in mortality derived from infectious diseases. However, with respect to this latter type of mortality, the place of residence - seems to have a greater influence on the chances of survival than socio-economic status.

KEYWORDS: socioeconomic status, inequality, mortality, Mediterranean economies, nineteenth century.

JEL CODES: N33, I14, I18, J18

1 INTRODUCTION:

The debate on the evolution of inequality in the standard of living constitutes a central theme of economic history and remains open on both a biological and material level (Deaton, 2015; Escudero, 2002; Harris, 2009, 2019; Piketty, 2014). There is renewed interest regarding the importance of the determinants of mortality in populations of the past and regarding the evolution of the social inequality in mortality in the different stages of economic development (Bengtsson & Van Poppel, 2011; Debiasi & Dribe, 2019). Currently, there is relative consensus with respect to the many determinants that condition health, illness and mortality (Dahlgren & Whitehead, 1991; Harris, 2004; Harris & Helgertz, 2019; Schofield & Reher, 1991; Van Poppel et al., 2016).

The theoretical explanation of the long-term evolution of social inequalities in mortality is based on two principal visions. The constancy or fundamental cause theory maintains the existence of a social gradient in levels of mortality that persists over time. According to this perspective, the differences between social classes condition the access to resources such as education, money, power and prestige, among others. These factors enable the reduction of certain risks of contracting diseases and minimise their consequences once contracted (Freese & Lutfey, 2011; Link & Phelan, 1995, 1996, 2002; Marmot, 2004, 2005; Phelan et al., 2010). On the contrary, according to the divergence-convergence hypothesis, inequalities in mortality have experienced, in a first phase, a diverging trend –due to what we could call dual behaviour: a notable improvement in the standard of living of the more well-off classes and a stagnation or even worsening of the standard of living of the lower classes–. It was followed by a convergence phase –the diffusion of medical, sanitation and economic advances fostered a greater increase in the standard of living of those with a lower socio-economic status, reducing social differences– (Antonovsky, 1967; Deaton, 2015, pp. 103-108; Smith, 1983; Vallin, 2013; Vallin & Meslé, 2004). The increase in the social inequality in mortality during the twentieth century has led some authors to rename the hypothesis as "*divergence-convergence-divergence hypothesis*" (Bengtsson & Van Poppel, 2011).

Therefore, according to the divergence-convergence hypothesis, during the old demographic and epidemiological regime, the high levels of mortality, the predominance of infectious diseases and their high virulence, together with the limited medical knowledge, gave rise to low levels of inequality, putting everyone on the same level at the time of their death (Antonovsky, 1967; Pressat, 1985, pp. 53-59). However, according to the constancy or fundamental cause theory, the inequalities in accessing resources provided an advantage to those individuals with a higher socio-economic status, irrespective of the predominant epidemiological pattern (Freese & Lutfey, 2011; Link & Phelan, 2002).

The most recent articles on the social differences in mortality on a historical level offer diverse results. Some find significant distinctions in pre-transitional periods, even from the seventeenth or eighteenth centuries (Fornasin et al., 2016; Perrenoud, 1975, 1981; Schumacher & Oris, 2011) or during the early phases of the demographic transition (Breschi et al., 2011; Ferrie, 2003). Others have not found even a slight indication of inequality due to the socio-economic status until at least the end of the nineteenth century (Molitoris & Dribe, 2016; Razzell & Spence, 2006), beginning of the twentieth century (Edvinsson & Broström, 2017; Edvinsson & Lindkvist, 2011) or even until the second half of the twentieth century (Bengtsson et al., 2020; Bengtsson & Dribe, 2011; Debiasi & Dribe, 2019; Gagnon et al., 2011). The evidence obtained reveals that the differences depend on the individual characteristics of the different economies, their development processes and the state of medical science and the predominant epidemiological pattern.

The determining factors of mortality vary between the different causes of death, which means that the social differences could vary among the types of diseases causing death (Elo et al., 2014; Ferrie, 2003; «The Relationship of Nutrition, Disease, and Social Conditions: A Graphical Presentation», 1983). Furthermore, as the medical knowledge regarding the disease in question advances, the levels of social inequality tend to pass through different phases with a progression similar to that, on an aggregate level, described by the divergence-convergence hypothesis (Clouston et al., 2016; Vallin, 2013, p. 159). For these reasons, studying the social differences in depth and distinguishing between infectious and non-infectious diseases can provide valuable information so as to enhance our understanding of the impact of the different underlying mechanisms (Debiasi & Dribe, 2019; Erikson & Torssander, 2008; Toch-Marquardt et al., 2014).

Most of the studies analysing the social differences in mortality do not distinguish between specific causes of death. Studies on the existence of a social gradient analysing specific causes of mortality are scarce, particularly for periods before the twentieth century. Of the few studies that analyse inequalities related to specific causes of mortality, we can find examples such as the one conducted by Leonard, Robinson, Swedlund, & Anderton (2015) in which no socio-economic differences are observed in mortality derived from infectious diseases in the city of Holyoke (Massachusetts) between the second half of the nineteenth century and the first decades of the twentieth century. In other cases, such as the study by Debiasi & Dribe, (2019) for southern Sweden, differences are found related to infectious diseases and parasitic diseases from the beginning of the nineteenth century, but are only significantly higher for unqualified workers. On the other hand, in the study on the social differences in mortality in the mid nineteenth century in several eminently rural American counties, Ferrie (2003) analyses the deaths caused by tuberculosis. The results do not show a clear correlation between socio-economic status and mortality. Rather, they reflect the importance of the working environment and the lower risk of contagion associated to occupations carried out in the open air and with little social contact.

Particularly interesting for our study is the analysis carried out by Pritchett & Tunali (1995), examining the determinants of mortality caused by an epidemic outbreak of yellow fever in the city of New Orleans in 1853. The results show a higher mortality among poor individuals with respect to wealthier people. The authors allude to the fact that this group would be probably more exposed to the vector of the disease, the *Aedes aegyopti* mosquito, and lived in overcrowded conditions in small homes in densely populated neighbourhoods. Differences may also be observed in mortality between the natives and the immigrants, most likely derived from the higher prevalence of individuals with immunity to the disease among the native population due to the recurrence of the illness in the area.

It should be taken into account that the lack of studies analysing the social inequalities differentiating between causes of mortality is particularly intense with respect to southern European countries and during periods prior to demographic and health transitions. For this reason, by contributing new evidence related to this "knowledge gap", this study can enhance the knowledge of social inequalities on a historical level and their causes. In order to fulfil this objective, using longitudinal and individual data, we have analysed the behaviour of mortality in the town of Vera in south-east Spain at the end of the eighteenth century and beginning of the nineteenth century. The period studied was characterised by a mortality typical of the former demographic regime, a predominance of infectious diseases and recurrent crises of epidemic mortality. Contrary to the situation of the European countries that were pioneers in the demographic transition, in Spain the reduction in catastrophic mortality, particularly epidemic mortality, did not culminate until 1900 (Nadal, 1966, pp. 9-17). In fact, during the period under study (1797-1812), the town suffered from three large-scale yellow fever outbreaks[1]. The methodological approach is based on competing risk models, as these types of models enable us to analyse the influence of the socio-economic status in mortality, differentiating between causes of illness while controlling for variables such as age, sex, the location of the home and the number of people living in the household .

2 SOURCES, DATA AND METHODOLOGY

2.1 Vera as a case study

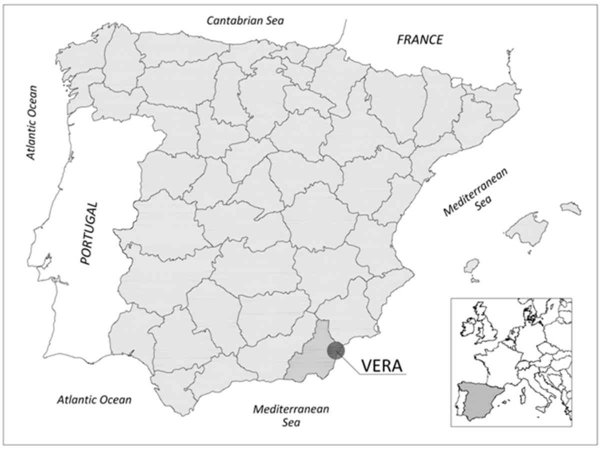
The town of Vera is located on the south-east coast of Spain in the province of Almería in the region of Andalusia (see Map 2.1). During the period analysed, it could be considered as having an average population with a production structure based on agriculture and livestock farming[2]. Furthermore, the good condition with which its demographic documentary sources have been conserved and the systematic recording of the demographic developments have enabled us to carry out an analysis using microdata, based on the reconstruction of the life courses of individuals, combining the information of different documentary sources. In addition, the wealth of information provided by its parish records enables us to disaggregate between the causes of

[1] During the eighteenth and nineteenth centuries, different populations in southern and eastern Spain located near port areas were affected by episodes of yellow fever. These outbreaks were related to trade openness and, specifically, to the configuration of an Atlantic trading system (Lobato Franco & Oliva Melgar, 2013; Martínez Shaw & Oliva Melgar, 2005; McNeill & Mcneill, 2004). To this effect, the frequency of yellow fever outbreaks on the peninsular coasts reveals a relationship with colonial trade. The first cases of yellow fever in the Iberian Peninsula appeared in Cádiz (south of Andalusia) throughout the eighteenth century. However, a few decades after the liberalisation of trade with the colonies in 1765 and 1778, which led to the opening of nine ports to free trade with the Caribbean regions, different coastal areas in the south and the east of Spain suffered different outbreaks of yellow fever (Delgado Ribas, 1986; Morillon et al., 2002; Nadal, 1966, pp. 100-108; Rodríguez Ocaña, 1987), mainly in the years 1800, 1804 and 1810-1812. The loss of the American continental empire between the French occupation and the Liberal Triennium closed privileged markets for Spanish trade, and considerably reduced Spanish exports and imports with these territories (Cuenca-Esteban, 2008; Fisher, 1993; García-Baquero, 1984; Prados de la Escosura, 1993). As might be expected, one of the consequences of this process was the reduction of yellow fever outbreaks.

[2] Like the rest of the country, these sectors employed the majority of the labour force. On the other hand, during the eighteenth century the economic structure diversified due to the growth of maritime activities and secondary sector activities such as the production of esparto and saltpetre (Conte Cazcarro, 2009). The distribution of land ownership, far from being considered egalitarian, did not have a very high degree of concentration, especially if we compare it with western Andalusia where large estates were the most common type of property. This feature is explained, among other factors, by the persistence of the type of land ownership from the repopulation after the expulsion of the Moors, which was based on the distribution of land plots, which included a house and land in order to favour the settlement of colonists (Bosque Maurel, 1973; Jiménez Alcázar, 1994).

mortality for a period prior to the time when, parish records in general began to include data regarding the cause of death[3] (Bernabeu-Mestre, 1992, pp. 29-30).

Map 2.1. Geographic location of Vera



Source: own elaboration

2.2 Sources and the construction of the database

The majority of the historical documents used in this study are stored in the Municipal Historical Archive of Vera. The information on the structure of the population has been drawn from two cross-sectional demographic sources; the Census of Godoy of 1797 and the municipal register of 1812, which is a local census[4]. Two other sources have also been used that provide information about demographic dynamics: the parish register records of burials for the years 1797-1812 and those of marriages for the period 1798-1812. A nominative and complete transcription was made of these demographic sources including all of the information shown in Table 2.1. After the literal transcription, the data was cleaned and standardised. The objective of cleaning the data was to eliminate or reduce spelling variations in the variables –e.g. first names, surnames, occupations, street names–. Meanwhile, the standardisation consists in replacing the text with a number, word or code that represents a common instance. The purpose of the latter task

[3] It was not until 1837 that an ecclesiastic regulation came into force which obliged parishes to collect information regarding deaths and the illness that caused the death (Bernabeu-Mestre, 1992; Rodríguez Ocaña & Martínez Navarro, 2008, p. 35).

[4] Given that we were interested in including the years with epidemic outbreaks in the analysis, we decided to use the Census of Godoy as a master file, together with the Census of Florida Blanca, regarded as one of the best censuses carried out in the whole of Europe in the Old Regime (García Pérez, 2007; Reher & Valero Lobo, 1995, pp. 22-23). Furthermore, although the political instability of the first half of the nineteenth century worsened the national statistics, in the case of Vera we had access to the municipal register of 1812, which included the whole population of the town and their occupations, and were able to use it to trace the demographic developments of the population.

was to simplify the data, reducing the variability and to facilitate and enable the nominative linkage of records and the subsequent statistical analysis (Clausen, 2015).

Table 2.1 Summary of the information contained in each of the demographic sources

	Census 1797	Municipal register 1812	Death register 1797-1800	Death register 1801-1812	Marriage register (1798-1811)	Database
Number of records	6479	6083	314	2343	960	3851
Individuals ≥ 7 years	5445				960	3747
Ego´s name	X	X	X	X	X	X
Ego´s surname	X	X	X	X	X	X
Adress	X	X				X
Name and surname of relative	+	+	X	X	X	X
Number of people living in the same home	X	X				X
Age	X	X		*		X
Occupation	X	X		X		X
Sex	+	+	+	+	X	X
Civil Status	X		X	X	X	X
Cause of Death				X		*
Date	X	X	X	X	X	X
Tenure status of household	X					X

Source: own elaboration. "X" means that the information appears across all of the records of this source or period. The sign "*" means that the information is found in some of the records but not all of them. The sign "+" means that the information is not explicitly found in all cases but, in general, can be inferred.

The individuals of the 1797 census were nominatively linked to the death and marriage registers of the period and to the municipal register of 1812 in order to reconstructed their life course from 1797 until 1812 or to the time of their death (Abramitzky et al., 2018; Fellegi & Sunter, 1969; Sayers et al., 2015; Winkler, 1999). The objective was to generate a longitudinal database that included more than 3800 individuals for whom we have the variables of age, occupation, the occupation of the head of the household, the number of individuals with which he/she lived, the tenure status of the dwelling (rented or owned), sex, status (survivor, deceased) and the date and cause of death (for those appearing in the parish burial records) or the date of the last documentary record of the individual alive (marriage book or municipal register of 1812). The observation time is expressed in weeks. The majority of the children under the age of seven cannot be identified, given that the census of 1797 referred to them with names such as "son", "daughter", "youngest boy" or "youngest girl". Due to the impossibility of identifying the individual in question, even when the names of the parents and the characteristics of the family were known, we decided to exclude them from the study.

After cleaning and standardising the demographic data, the record linkage process was carried out following these steps: first, a nominative linkage was conducted between the individuals included in the 1797 census and those who got married in the period 1798-1812. The linkage was made using the *Jaro-Winkler* string distance and those results with coincidences below a 0.8 threshold have been left out in the case of both surnames and first names (Christen, 2012, pp. 109-111; Jaro, 1989; Winkler, 1990). After completing the linkage of marriages with the 1797 census the outcomes were linked with the municipal register of 1812. In this case, the objective was to locate the individuals who had formed a new family. In this way, we identified many individuals of the 1797 census in the 1812 municipal register which confirmed their survival during the period studied. Subsequently, we proceeded to identify the deceased by linking the 1797 census, previously linked with the marriage records, with the data drawn from the burial records.

The identification of the survivors in the 1812 municipal register enabled us to determine that these people did not die during the period studied. If we would have considered as survivors all those that had not been located as deceased, we would have over-estimated the survival of the groups most prone to migration (Christen et al., 2015; Efremova et al., 2015; Ruggles, 1992, 2002). The fact that parish records provided information about the name and surname of the next of kin[5] has been a fundamental element for identification, given that it enabled us to significantly reduce the overlinks derived from shared onomastics. In this way, based on the above-described procedures, 3747 individuals aged seven or over were identified in the 1797 census in at least two sources, reaching the 70% threshold level of individuals commonly considered as a reference in the nominative linkage processes (Ruggles et al., 2018). All of the links have been individually reviewed with the objective of reducing false matches that could be derived from the use of a deterministic approach[6] (Christen, 2012, pp. 163-185). The final result was the reconstruction of the demographic trajectory of 3851 individuals, 3747 over the age of seven, forming a longitudinal database enabling the tracking of these people during the period 1797-1812.

The possible inclusion of biases in the record linkage processes due to difficulties in identifying the population most prone to migrations has received much attention by the literature (Ruggles, 1992, 2002). In our case, the prevalence of temporary migrations and movements of people due to the War of Independence (1808 – 1814) adds another potential source of bias in the database, leading to a possible

---

[5] In the case of single people, these were the parents and in the case of married people, the spouse and in that of widows and widowers that of the deceased spouse.

[6] All of the information available for each individual has been used to verify the nominative linkage through the Jaro-Winkler distance (age, occupation, location of home, name and surnames, information about relatives, etc.). In this way, the basis for the linkage was the names, surnames and age of each individual, together with the nominative information of their relatives. However, all of the variables have not been given the same importance: occupation and neighbourhood have only been used to confirm the link but it is established that these two variables could have changed over time.

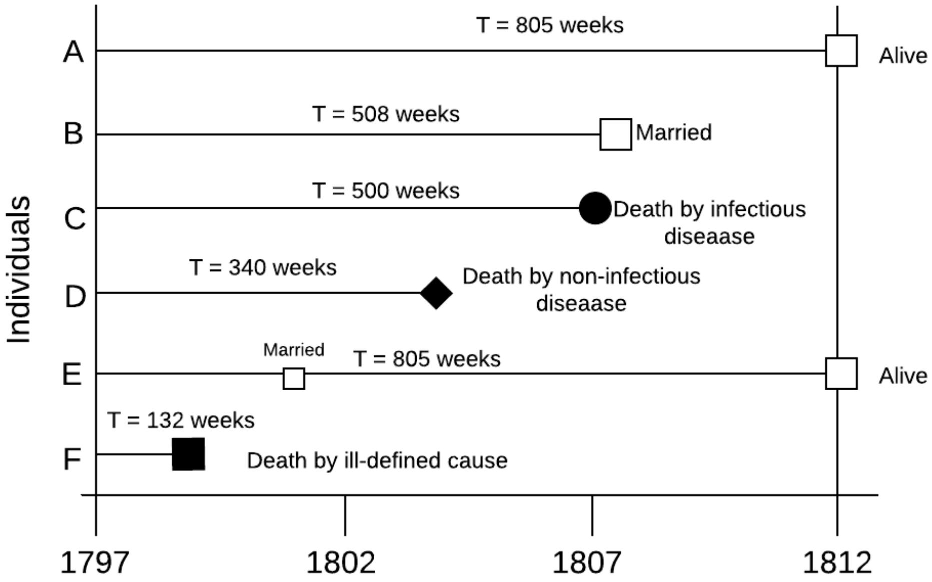
overestimation of their levels of survival. This is due to the fact that these people could have died during the time they were away and therefore their deaths do not appear in the parish records. Consequently, in the case of having a previous observation –derived from a marriage–, the status of the individual would be that of *survivor* and the time passing until the date of the last record would be used as the observation time. In the case in which there are no other observations, the individual would not appear in our database. However, having successfully overcome the migration or military campaign and returned, they would be identified as survivors as they would appear in the 1812 municipal register[7].

The individuals for whom there are not at least two observations and, therefore, have not been considered in this analysis can be classified into two groups: 1) those who only appear in the 1797 census and for whom there is no information available in any other records and, 2) those who share their first name and surname with other individuals and, although appearing one or several subsequent records, we have not been able to determine, with the additional information, which of the coinciding individuals it refers to (overlinks). This latter group is very small (<2%). As the marriage and burial records provide information about relatives, in most cases we have been able to identify the person in question despite having the same first name, surname and age as another individual[8]. Therefore, this analysis only includes individuals for whom there is certainty regarding their death or survival until a specific date. In diagram 2.1. we can observe examples of different individuals included in our database.

Diagram 2.1. Graphic representation of the demographic events identified for different individuals

[7] The studies on temporary migration in Vera in the mid-nineteenth century have shown how between 700 and 1,000 passports were issued annually, most of which were free, suggesting the low level of income of the applicants. These figures represent a high percentage (more than 15%) of the total population of the town, equal to more than a third of the casual workers (Sánchez Picón, 1988). Therefore, the effect of the greater propensity of individuals belonging to the lower social class to emigrate temporarily would introduce an underestimation of their mortality levels in our results, giving greater significance to the social gradient in mortality than our results show.

[8] However, it is probable that this means that the individuals who lived alone and did not marry during the period are under-represented (Antonie et al., 2015). However, this greater difficulty to link these people who appear in the 1797 census will be more or less similar in the case of the survivors and the deceased. In other words, it makes it difficult to identify an individual dying and appearing in the parish burial records or identifying a person who survived and appeared in the 1812 municipal register. Therefore, *a priori,* one effect can compensate the other, and as we have only included those individuals for which we have at least two observations, this should not influence the estimates of the probability of survival or death.



Source: own elaboration.

In order to classify the individuals in occupational categories, we used the *Historical International Standard Classification of Occupations* (HISCO)[9]. The socioeconomic status of the individuals in a family unit has been assigned based on the occupation of the head of the family in the 1797 census[10]. In accordance with the different codes that HISCO assigns to each occupation, we used the HISCLASS scheme to transform the occupations into social groups. This scheme was chosen so that the results could be more easily compared on an international level (Van Leeuwen & Maas, 2011). The 12 HISCLASS categories were regrouped into three groups. The first includes categories 1 to 5 and represents those with the highest socio-economic status. The intermediate status group includes the individuals in categories 6 to 9 while the low status group is formed by categories 10 to 12. This modification responds to the need to better adapt to the production structure in Vera at that time as the range of occupations was limited and relatively concentrated in certain occupations. Therefore, grouping the individuals into larger socio-economic categories enabled us to obtain more robust results than those that we would have obtained if we were to analyse each HISCLASS category separately.

With respect to the causes of death, there are difficulties involved in studying a period between the end of the eighteenth century and the beginning of the nineteenth century. These limitations are derived from 1) the problem regarding the exactness and

---

[9] See Pujadas Mora, Romero Marín, y Villar Garruta (2014) y Van Leeuwen, Maas y Miles (2002).

[10] In the case of the servants, we have considered the same geographical variable (neighbourhood) as that of the family for whom they worked and also the number of people with whom they lived. However, we have maintained the category HISCLASS corresponding to their occupation.

reliability of the cause-of-death diagnoses[11]; 2) the fact that the same cause can have different diagnoses due to different diachronic reasons; 3) the records of the cause of death were not always kept according to strictly scientific criteria, being influenced by the popular conceptions of certain causes of death and 4) the influence that the history of the many classifications and nomenclatures had on the "medical" causes of death (Arrizabalaga, 1993; Barona Vilar, 1993, p. 49; Bernabeu-Mestre, 1993, pp. 14-15; Perdiguero Gil, 1993; Ramiro Fariñas et al., 2003, p. 3).

The period under study is prior to the development of microbiological theory. However, during this period, the causes of disease were attributed to combined elements of the dominant scientific trends of the day; the miasmatic theory and contagionism (Barona Vilar, 1993; Preston & Haines, 1991, pp. 6-11; Urquía, 2006). This means that it is advisable to use a historical equivalence scheme in order to carry out an adequate classification of the diseases. In this way, with the objective of conducting the analysis differentiating between types of disease, it was necessary to go beyond the diagnosis appearing in the documentary source so as to find a nomenclature with which to place it in a category within a classification preferably made according to aetiological criteria[12]. This enabled us to analyse the relationship between the causes of death and their principal determinants as the classification carried out by Ramiro Fariñas, Sanz Gimeno, Bernabeu Mestre y Robles González (2003)[13] facilitates.

2.3 Method of analysis

As we only have information about the causes of death for those who died from 1800 onwards, we chose to use two survival analysis methods[14]: 1) the Cox Proportional Hazards (COX PH) model for the whole of the period (1797-1812) as this method does not take into account the type of cause of death, and 2) an analysis of competing risks for the period 1800-1812, given that this enables the effect of the different covariables for each type of disease to be evaluated. The objective of the Cox regression is to detect

---

[11] In certain circumstances and at certain moments, the death records could have been kept for purposes that were not exclusively medical. The causes of the illness may have been concealed for different reasons such as maintaining public order, attempting to avoid social stigmatisation or covering up professional negligence (Arrizabalaga, 1993, pp. 33-35).

[12] The aetiological classification of the causes of death is particularly useful for explaining cause-effect relationships principally when we are faced with a delayed phase of the demographic or epidemiological transition predominated by infectious diseases (Ramiro Fariñas et al., 2003, p. 35).

[13] These authors base their categories on the disease classification proposed by Bertillon (1899), which constituted the starting point for constructing the different versions of the International Classification of Diseases published by the World Health Organisation and McKeown's (1976) proposal given that it differentiates between diseases according to their transmission mechanism. For these reasons, this classification is ideal for studying the impact of the different determinants in the evolution of mortality due to specific causes.

[14] By *survival analysis* we understand the series of statistical techniques used to analyse data whose objective is to model the time expected for a certain event to happen (Kleinbaum & Klein, 2012, pp. 1-43; Llopis Pérez, 1996).

the possible existence of a relationship between the risk that a certain event of interest will occur –in our case death– and one or several independent explanatory variables –such as age, sex, socio-economic status, location of the home or the number of people living together– (Cox, 1972; Llopis Pérez, 2014). In this method, the dependent variable is dichotomous –survivor or deceased–. Therefore, as we do not distinguish between types of death we are able to use the information corresponding to the period 1797-1800.

The risk function of the COX PH model for an individual with the characteristics of vector X can be formulated as follows:

$$\lambda(t|X) = \lambda_0(t)e^{\sum_i^n \beta_i X_i}$$

i = [socioeconomic status, age, neighbourhood of residence, ... n], with n being the number of variables

$X_i = \begin{bmatrix} X_{SES} \\ X_{age} \\ \vdots \\ X_n \end{bmatrix}$, where X_i corresponds to the value of variable i.

$\beta_i = [\beta_{SES} \quad \beta_{age} \quad \cdots \quad \beta_n]$, where β_i corresponds to the value "beta" of variable i.

The hazard function is the product of two elements; the baseline hazard function $\lambda_0(t)$ and a second term that represents the influence of a *covariate vector* in the probability that the event in question will occur. The principal output of this regression are the relative risks $exp(\beta_i)$. If the variable is categorical, the relative risk represents the ratio between two risk functions; that associated to the value of the variable and that of the reference category. When the variable is continuous, the relative risk shows the effect of the increase of one unit in the value of this variable, with the rest of the variables remaining unchanged (Blossfeld & Rohwer, 2002; Kleinbaum & Klein, 2012).

One of the principal advantages of this method with respect to others, such as minimum least squares regression, is that the Cox regression enables us to use the information regarding the cases in which there is a bias towards the right - those people for which we only have information until a certain point. In recent years, the COX PH model has been one of the most frequently used in studies analysing social differences in mortality (Bengtsson & Lindström, 2000; Edvinsson & Broström, 2012; Edvinsson & Lindkvist, 2011; Gagnon, Tremblay, Vézina, & Seabrook, 2011; Mamelund, 2006; Schenk & Van Poppel, 2011).

On the other hand, given that the importance of the different determinants of mortality can vary significantly according to the type of disease (Debiasi & Dribe, 2019; Elo et al., 2014), we have also applied another survival analysis model to conduct the study differentiating between mortality caused by infectious diseases, non-infectious diseases and ill-defined causes. By differentiating between these three types of death, we find a typical case of competing risk (Kleinbaum & Klein, 2012, p. 430; Pintilie, 2011, p. 600). We find competing risk when there is more than one type of event (surviving, dying from an infectious disease, from a non-infectious disease or due to an ill-defined cause) and the occurrence of one of them modifies the probability of observing the event of interest.

Therefore, we decided to use a multivariable regression analysis based on the semi-parametric analysis and proportional hazard models proposed by Fine & Gray (1999). The aptness of this analysis for studying the influence of the different variables in situations in which different competing events can occur has been shown in different studies (Lau et al., 2009; Scrucca et al., 2010). The method proposed by Fine & Gray (1999) is a sub-distribution hazard model of the cumulative incidence function. Similarly to the COX PH model, the assumption of the proportionality of the hazards continues to be enforceable but referring to the sub-distribution hazards (Pintilie, 2011, p. 603). The model is formulated as follows:

$$\lambda^c(t|X) = \lambda_0^c(t)e^{\sum_i^n \beta_i^c x_i}$$

Where "$\lambda_0^c(t)$ is the baseline sub-distribution hazard of cause c and β_i is the vector of coefficients for the covariates. The estimation for this model follows the partial likelihood approach used in a standard Cox model" (Scrucca et al., 2010, p. 1389). The interpretation of the exponents of the coefficients will be similar to that previously described for the COX regression, except in this case we refer exclusively to the hazard associated to cause *c* (Austin & Fine, 2017; Scrucca et al., 2010).

We have verified the fulfilment of the assumption of proportional hazards based on the estimate of Schoenfeld's residuals and the analysis of their graphical representation (Cleves et al., 2008; Kleinbaum & Klein, 2012). The results, particularly in the models that distinguish between sub-periods and types of diseases do not reveal serious violations.

3 RESULTS

3.1 Descriptive analysis

Table 3.1 describes the principal characteristics of the sample which includes 3747 inhabitants over the age of seven who appeared in the 1797 census and for whom we have information about their death or survival throughout the period analysed. The first two columns represent, respectively, the number of individuals in the 1797 census within each category and the percentage of them that have been linked. In our database, with respect to the HISCLASS category of the head of the household, we can observe that there is information for more than 90% of the individuals. The largest group is that formed by categories 6-9, which includes the "*lower-skilled workers*"[15]. The following group, in relative terms, represents the occupations with a lower social category (HISCLASS 10-12), among which we can highlight the "*unskilled farm workers*"[16]. Meanwhile, the group made up of the categories with a higher status in the classification system (HISCLASS 1-5) would be the smallest in percentage terms, despite including almost 600 individuals[17]. More than a half of the individuals analysed survived during the time period studied (58.1% survivors as opposed to 41.9% deceased). Among the deceased, the sample includes 968 who died from infectious diseases, 167 who died from non-infectious diseases and 445 due to ill-defined causes. The average age of the individuals in the 1797 census who finally survived was 24.7 years, that of those who died from infectious diseases was 37.2, that of those who died due to non-infectious diseases was 40.2 and that of those who died from ill-defined causes was 52.6[18].

Table 3.1. Descriptive analysis of our database

	1797 census (individuals ≥ 7 years)		**Our database (≥ 7 years)**					
Population by HISCLASS	N	% links	N	%	Average age	Sd	Survivals	Deceased
1. Higher managers	250	66	165	4	37.0	19.0	58	42
2. Higher professionals	70	44	31	1	39.5	18.4	35	65
3. Lower managers	284	75	213	6	31.4	18.3	66	34
4. Lower professionals	71	68	48	1	33.5	16.8	44	56
5. Lower clerical and sales	178	76	135	4	28.5	16.6	54	46
6. Foremen	2	50	1	0	45.0	-	0	100
7. Medium-skilled workers	321	72	232	6	30.0	17.6	60	40
8. Farmers	483	77	371	10	29.7	17.2	67	33
9. Lower-skilled workers	1543	65	1001	27	30.5	19.9	60	40

[15] Within this category, the most common occupations were that of "muleteer", (repeated 210 times) and that of "in campaign", (repeated 211 times).

[16] This category is principally formed by families whose main breadwinner was a day labourer. Specifically, this occupation is repeated 377 times.

[17] The occupation of the head of the household most repeated among these individuals is that of "landlord" or "landlady" (53 times).

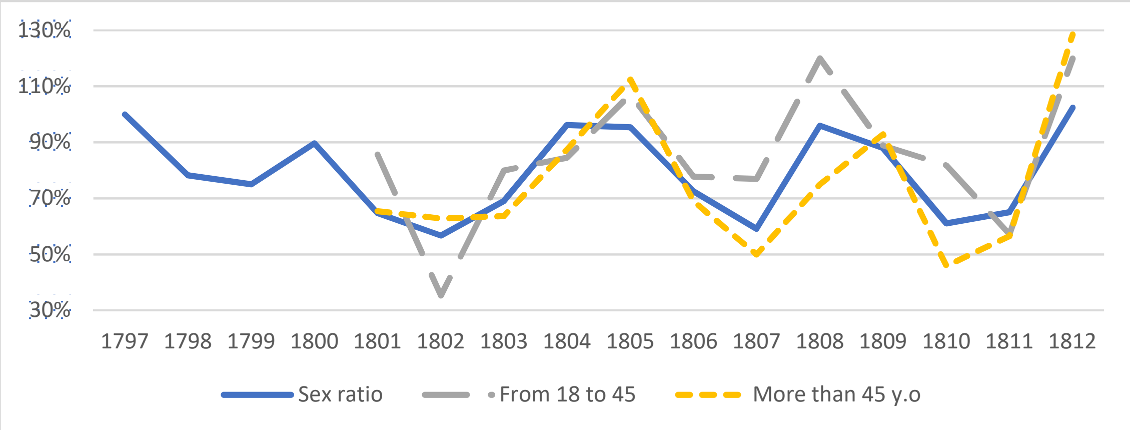
[18] Throughout the period 1797-1812, an important transformation in the occupational structure took place. Between these two dates, we can observe a drop in the percentage of individuals classified as being Upper class and Middle class. On the contrary, the proportion of individuals belonging to the lowest social class increased. In general terms, there was a downward social mobility due to the epidemic outbreaks of yellow fever, together with the economic consequences of the War of Independence which, among other factors, could have negatively affected the resources of many families. In any case, this mobility should not affect our results since only the socioeconomic status in 1797 is taken into account, regardless of whether or not there was social mobility. The reason why we have chosen this criterion is that the marriage and some death registers do not include information on the occupation of individuals and, although the 1812 census does include it, we cannot know when the change of occupation took place, and we would only have this information for the survivors.

10. Lower-skilled farm workers	376	80	299	8	33.3	18.8	51	49
11. Unskilled workers	113	53	60	2	33.9	21.5	57	43
12. Unskilled farm workers	1194	69	825	22	30.6	17.2	55	45
Unknown	560	65	366	10	32.2	19.9	57	43
Population by HISCLASS								
Upper Class (1-5)	853	69	592	16	32.9	18.3	57	43
Middle Class (6-9)	2349	68	1605	43	30.2	17.7	62	38
Lower Class (10-12)	1683	70	1184	32	31.4	17.9	54	46
Unknown	560	65	366	10	32.2	19.9	57	43
Sex								
Female	2818	73	2063	55	31.5	18.2	56	44
Male	2627	64	1684	45	30.8	18.0	60	40
Neighbourhoods								
1) Intramuros	1452	65	942	25	33.4	18.4	54	46
2) Arrabal	964	72	691	18	30.7	18.1	60	40
3) Crecimiento Sur	906	65	591	16	29.9	17.7	62	38
4) Crecimiento Oeste	744	70	520	14	31.4	18.8	54	46
5) Ensanche Oeste	943	74	697	19	30.6	17.9	58	42
6) Barrio de Jesús	266	71	188	5	29.0	16.2	63	37
7) Outskirts	170	69	118	3	29.8	17.4	64	36
Events								
0 = Survivor			2177	58	23.9	13.5		
1 = Deceased by infectious disease			968	26	37.2	17.1		
2 = Deceased by non-infectious disease			167	4	40.2	18.0		
3 = Deceased by ill-defined causes			445	12	52.6	17.9		
Total	5445	69		3747	100	31.1	2177	1570

Source: own elaboration (see text).

The percentage of women included in our database is higher than that of men (55.1 % as opposed to 44.9 %). Furthermore, the average age of the women is also slightly higher. In the 1797 census the number of women was already higher than that of men. However, the lower percentage of men that can be linked across the censuses and death certificates (64% as opposed to 73% in the case of women) increases the relative difference. The greater success obtained in the linkage of women across time is due to the fact that the percentage of women identified in the death certificates is higher than that of men. If we take into consideration the sex composition of the death records, we can see that with the exception of 1812, more women died than men, which would be consistent with the greater propensity of men to temporarily migrate. In the years of epidemic outbreaks - 1804, 1811 and 1812 - the sex ratio increases and displays higher values that are approach levels that are close to “normality”. This could be consequence of a lower intensity of temporary migrations during these years due the exceptional circumstances of epidemics and the sanitary measures implemented that limited movement. We will return to this question later when explaining survival models. Moreover, in the evolution of the sex ratio between the established subperiods, we can appreciate that there are hardly any differences. In both cases the level of male deaths over female deaths is between 81.5% and 82%. This similarity in the percentage renders the results of the two subperiods comparable.

Graph 3.1. Annual sex ratio in death records 1797-1812 (male/female)



Source: See text

Another of the variables included for the whole group of individuals of the sample is the location of their dwelling in the 1797 census[19]. We believe that this variable can be particularly important to explain the behaviour of mortality caused by infectious diseases due to the contagion mechanisms (Recaño & Esteve, 2006). Furthermore, it permits us to capture the inequality in mortality derived from the hygiene and sanitation characteristics of each of the neighbourhoods. Finally, it enables us to better isolate the effect of the socio-economic status, as we can control the possible "confounding effect" that introducing the location of the dwelling could cause. We have grouped the streets into seven areas, based on the classification conducted by Luis Cano (2015). In table 5.1 we can observe the relative weight that each of them have in the total population of the municipality, while table 3.2 shows their socio-economic composition. Their territorial location and the outline of their streets can be seen in the map of the town of 1798 shown below.

[19] Calculations have been made regarding the spatial mobility of the population. The results reveal that the vast majority of the surviving population remained in the same neighbourhood between 1797 and 1812. The location of the home in the 1812 census has not been taken into account in the analyses given that we only know the social mobility of those individuals who we know survived the period studied. Therefore, given that we cannot know whether there was a change of residence prior to the death of those for whom we have death records, we have opted to treat all of the individuals homogeneously, using as a variable the place of residence in 1797.

Table 3.2. Distribution of the socioeconomic status of the population by neighbourhood

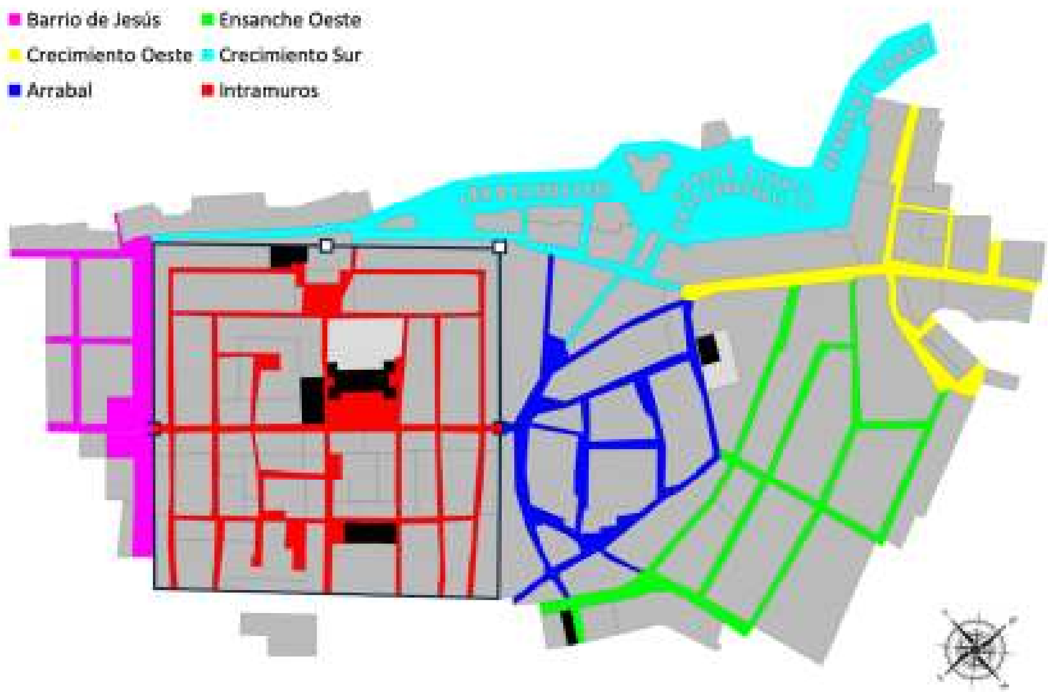
	Upper Class (1-5)	Middle Class (6-9)	Lower Class (10-12)	Unknown	Total
1) Intramuros	26 %	39 %	27 %	8 %	1452
2) Arrabal	15 %	49 %	27 %	9 %	964
3) Crecimiento Sur	10 %	44 %	33 %	13 %	906
4) Crecimiento Oeste	15 %	31 %	36 %	18 %	744
5) Ensanche Oeste	8 %	47 %	34 %	10 %	943
6) Barrio de Jesús	14 %	45 %	38 %	3 %	266
7) Outskirts	5 %	68 %	23 %	4 %	170
Total	853	2349	1683	560	5445

Source: own elaboration

The area with the largest population is that of *Intramuros*, where a quarter of the inhabitants lived and, as its name (intra-walls) indicates, comprises the walled area of the town. This is the oldest neighbourhood, dating back to the first half of the sixteenth century. The average age of its inhabitants is the highest and incorporates a high percentage of higher class neighbours. The socio-economic improvements occurring after the end of the seventeenth century stimulated a rapid population growth, requiring "as well as the densification of the existing stretch (...) a planned urban extension in the north-east of the town" (Cano, 2019, p. 278). Until the eighteenth century, the growth of the town outside of the walls had only been permitted on the north-east side for defensive reasons. As a result of this urban extension, the population residing in the Zanja neighbourhood (included in map 5.1 within the Arrabal area and the Crecimiento Oeste) grew larger than that of the intra-wall area from the mid eighteenth century. These neighbourhoods have a "high density, higher than one hundred dwellings per hectare, with low rise buildings, due to an intensive subdivision of the land, with no free plots or spaces, with narrow streets and rectangular blocks" (Cano, 2015, pp. 95-96).

The *Barrio de Jesús* neighbourhood dates back to the last quarter of the eighteenth century when, with the reinforcement of security on the coast, the relaxation of the defensive functions enabled the eastern part of the town to be developed. Although the types of plot proposed were similar to those in the Zanja neighbourhood, in this case, the majority of them had two or more modules, which considerably reduced the density of the inhabitants of this area. Finally, the "outskirts" area included the neighbours who lived outside of the urban area. This is made up of the Garrucha area, which is on the coast, in addition to the "llano y las huertas" and the "Jara". It is the area with the lowest population density and was mostly formed by isolated houses and farmhouses. This is reflected in the labour distribution of its inhabitants who largely appear with occupations such as "*labrador*" or "*labrador mediero*" (Cano, 2015).

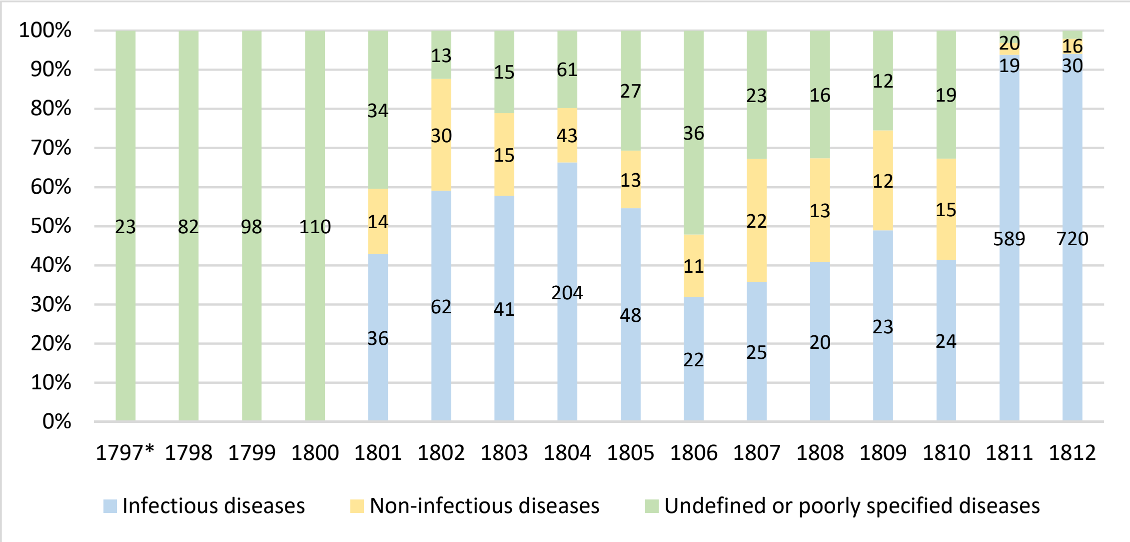
Map 3.1: Map of the location of the neighbourhoods of Vera 1797



Source: Own elaboration based on the map drawn up by Cano (2015) based on the Godoy census. The outskirts zone is all the municipal area not included in this map.

The period of time analysed in this study (1797-1812) is prior to the demographic transformations and the epidemiological pattern described by the demographic and epidemiological transition theories. This was a time characterised by high mortality rates combined with years in which several yellow fever outbreaks gave rise to catastrophic mortality rates. With respect to the epidemiological pattern (Graph 3.1), we can observe the predominance of infectious diseases with respect to non-infectious diseases and ill-defined or non-specified diseases. Due to the magnitude of the epidemic outbreaks of yellow fever in the years 1804, 1811 and 1812 the structure of the causes of death differ substantially between the years in which these outbreaks occurred and those when they did not. As could be expected, the years in which the percentage of deaths due to infectious diseases was higher are those in which the yellow fever outbreaks occurred (1804, 1811 and 1812). However, in the years when there were no epidemic outbreaks the number of deaths due to infectious diseases was higher than those due to non-infectious diseases. As mentioned in the methodology section, until May 1801, we have no information about the cause of death so its classification was not possible.

Graph 3.1 Distribution of mortality by type of disease



Source: Own elaboration based on the information drawn from parish burial records. * For the year 1797 only the deaths from November are included given that this is when the census ended.

3.2 Survival analysis[20]

Table 3.3 shows the results of the survival analysis based on the COX PH model. Due to the absence of information regarding the causes of mortality during the four first years of the period studied, we decided first, to conduct an analysis of the effect of the different covariables on mortality for the whole period without differentiating between types of death. This enabled us to use the data available for the first four years (1797-1800) and provided us with information about the effect of the socio-economic status and the rest of the independent variables –age, sex, and number of people living in the home and neighbourhood of residence– on mortality in general.

Model 1 of Table 3.3 includes both men and women. The results show that belonging to the social group with the highest social status and the intermediate group constituted a clear advantage in terms of survival. Both categories had significantly lower probabilities of death than that of the individuals of the lowest socioeconomic status, used as a reference category. As could be expected, the variable of age was positively associated with the risk of death. An increase of one year in age increased the relative risk of death by approximately 4.9%. With respect to sex, men had a lower risk of dying than women. We will attempt to find feasible explanations for this abnormal result in the following pages, based on the estimates broken down by causes of death and sub-periods. Meanwhile, the spatial variable is not significant. Finally, we can

[20] As well as the survival models included in the article, stepwise models have also been carried out, progressively adding the different control variables. No significant changes are observed in the relationship between occupational status and survival. Furthermore, the effect of possible interactions between the different variables has been tested (between Class and nº of people; and between sex and nº of people), with the results showing scarce or no significance and a worsening of the goodness of fit of the estimates.

observe how a higher number of people in the dwelling seemed to be associated to a higher risk of death. We should remember that we are analysing mortality as a whole, therefore, the highest risk of contagion derived from living with other people could have been compensated by the advantages of greater social support in the form of care, food safety, the reduction of certain hazards, etc. (Bengtsson, 2004; Cassel, 1976).

Table 3.3. Cox regression for the whole period (1797-1812) Exp (coef).

	Model 1 (males and females)	Model 2 (females)	Model 3 (males)
Upper Class (1-5)	0.809**	0.858	0.767*
Middle Class (6-9)	0.811***	0.907	0.712***
Lower Class (10-12)	1	1	1
Age	1.049***	1.051***	1.046***
Female	1.000	-	-
Male	0.887*	-	-
Neighbourhoods			
1) Intramuros	1.016	1.145	0.870
2) Arrabal	0.896	1.013	0.760*
3) Crecimiento Sur	0.837•	0.932	0.727*
4) Crecimiento Oeste	1	1	1
5) Ensanche Oeste	0.935	0.966	0.895
6) Barrio de Jesús	0.835	0.919	0.747
7) Outskirts	0.920	0.944	0.878
Number of people living in the same home	0.980•	0.989	0.971•
Individuals (N)	3747	2063	1684
Events	1570	900	670

• p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,001

Model 2 shows the effect of the different covariables on female mortality. The age variable is significant, showing a higher risk of death as the age of the individuals increased. With respect to the differences between socio-economic categories, we can observe the possible existence of a social gradient in mortality, whereby there was a lower risk for individuals with the highest status, followed by the intermediate category, although the significance was low. Both the spatial variable and the number of people with whom individuals shared a home are not significant for the case of women.

In model 3 we can observe how, in the case of men, the social differences seemed to have had a greater effect on mortality levels. The individuals with the highest socioeconomic status and those with an intermediate status had a significantly lower risk of death than those with the lowest status. The lower risk related to living with a greater number of people is significant in the case of men. Finally, we can observe some differences related to the variable of place of residence. This effect will be shown in more detail and is better explained in the following tables, given that they differentiate between types of diseases and sub-periods.

Next we will comment on the results of the survival analysis based on the competing risk analysis. As mentioned in the methodology section, this type of analysis enables us to study the effect of the covariables on the different categories of death. For this reason, we begin the analysis from the year 1801, given that this is the first year that the death records include information about the causes of mortality.

Table 3.4 represents the value of the coefficients of the hazard ratio for the period (1801-1812), differentiating between mortality derived from ill-defined diseases (model 4), that derived from non-infectious diseases (model 5) and that derived from infectious diseases (models 6 and 7). With respect to mortality derived from ill-defined causes (model 4), we can observe a distinct social gradient: a clear advantage of individuals classified in the categories with the highest status followed by the intermediate category. As we could expect, the age variable is highly significant and is correlated with the probabilities of death. In these types of death, the effect of age is highest and, almost certainly, is due to the fact that a higher percentage of deaths due to ill-defined causes appeared with the diagnosis of "old age", which was a cause obviously concentrated among the older population segment. Moreover, according to the results, the higher the number of people living in a same household, the lower probability of death. With respect to the differences between the sexes due to ill-defined causes, we can observe that women experienced higher levels of mortality than men. Finally, regarding the location of the home, no significant differences can be observed among neighbourhoods.

Table 3.4. Competing risk regression (1801-1812). Individuals over the age of seven Exp (coef).

	Mortality derived from ill-defined diseases	Mortality derived from non-infectious diseases	Mortality derived from infectious diseases		
	Model 4	Model 5	Model 6	Model 7 (HISCLASS right side)	
1 (HISCLASS 1-5)	0.580*	0.730	0.949	0.764*	1 (HISCLASS 1, 3)
2 (HISCLASS 6-9)	0.616**	0.873	0.9	0.912	2 (HISCLASS 2, 4-11)
3 (HISCLASS 10-12)	1	1	1	1	3 (HISCLASS 12)
Age	1.076***	1.03***	1.025***	1.026***	Age
Females	1	1	1	1	Females
Males	0.712*	0.687*	1.023	1.029	Males
Neighbourhoods					Neighbourhoods
1) Intramuros	1.059	1.043	0.967	0.982	1) Intramuros
2) Arrabal	1.018	0.831	0.856	0.859	2) Arrabal
3) Crecimiento Sur	0.824	1.467	0.818	0.811	3) Crecimiento Sur
4) Crecimiento Oeste	1	1	1	1	4) Crecimiento Oeste
5) Ensanche Oeste	0.734	1.037	0.991	0.98	5) Ensanche Oeste
6) Barrio de Jesús	0.834	0.924	0.912	0.909	6) Barrio de Jesús
7) Outskirts	1.723	0.972	0.528*	0.516**	7) Outskirts
Number of people living in the same home	0.949	0.961	1.011	1.012	Number of people living in the same home
Individuals (N)	3481	3481	3481	3481	Individuals (N)
Events	205	167	968	968	Events

• p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,00

In model 5, which includes mortality derived from non-infectious diseases, we can observe a social gradient in mortality. However, the significance level is low. An increase in age, again can be seen to be associated with a higher risk of death, although its effect is lower than that observed in ill-defined diseases. With regard to the differences between the sexes, we can find that men had a significantly lower risk of death than women. This circumstance could explain, at least partly, the high level of mortality associated to giving birth (Alter et al., 2004; Chamberlain, 2006; Schofield, 1986) together with the differences in living conditions between the sexes (Beltrán Tapia & Gallego-Martínez, 2017; Humphries, 1991). However, in this case, the value is excessively high to be attributed exclusively to these factors. Among other causes, it could also be due to a possible bias in the database as a result of the displacement of people due to military campaigns and seasonal migrations[23]. Therefore, these individuals, mostly men, will only have been included in the database in the case of having survived these displacements, therefore, enabling them to be identified in the municipal register of 1812. However, those who died during the migration or the military campaign will not have been identified as deceased, as they do not appear in the town's parish burial records. The lower percentage of men who we have been able to link (64% as opposed to 73% in the case of women) would seem to confirm this possibility. Furthermore, on the contrary to what the results of the survival analysis suggest, the higher percentage of women in the census and the fact that they had a higher average age would seem to suggest a higher level of survival with respect to the men[24].

Meanwhile, living with a higher number of people constitutes a factor that reduced the possibilities of dying from these types of causes. Finally, regarding the location of the home, no significant differences can be observed between neighbourhoods.

Model 6 of this table shows the values of the coefficients for infectious mortality. We should remember that in this category the deaths derived from yellow fever are particularly prominent due to their relative weight. In this case, the risk of death did not differ significantly between men and women, although it was slightly higher in the case of men[25]. Similarly, the advantage related to the wealthier classes is not clearly and significantly reflected. Apparently, for these types of diseases, the differences in the

---

[23] In relation to the importance in this area of seasonal migrations as a mechanism for complementing family incomes, see Gómez Díaz (1995) and Sánchez Picón (1988, 2011)

[24] We have analysed the average sex ratio for the different months using the death records and the results coincide with the seasonality of temporary migration seen in this area. In this way, the lowest levels of male death occur between late spring and the end of summer, months in which (male) casual workers tended to be unemployed and took advantage of the time to supplement their family income by going to harvest in the Andalusia baja area (Sánchez Picón, 1988).

[25] The lower propensity to temporarily migrate in the epidemic years, together with the fact that a large percentage of infectious mortality (derived from the epidemic) occurred during these years, also helps to explain that our results on infectious mortality do not show a greater risk of female mortality compared with non-infectious or ill-defined mortality.

resources associated to the socio-economic status did not influence the possibilities of survival so much. At this point, we should remember that, according to the literature, the influence of the nutritional status on the resistance to the effects of yellow fever (mortality or morbidity) was low («The Relationship of Nutrition, Disease, and Social Conditions: A Graphical Presentation», 1983). This circumstance could help to explain why, in a period in which medical knowledge was insufficient, the social differences in this diseases were low. Despite having a positive correlation with the probabilities of dying, the age has a lower effect on mortality associated with infectious diseases than that associated with non-infectious diseases or ill-defined ones. With respect to the variable of the number of people living together at home, we can see how for this type of cause of death, the variable had the opposite effect than that of the non-infectious or ill-defined diseases: the higher the number of individuals in the home the higher the probability of death. This increase in the risk of infectious mortality related to living with a higher number of people was probably due to the consequent higher exposure to contagion. Moreover, with regard to the neighbourhood of residence, the coefficients show how living in the outskirts of the town was related to a substantially lower risk of dying from infectious diseases. The situation of isolation and the lower population density that characterised the dwellings of this area probably constituted a protective element against contagion.

The difference of model 7 with respect to model 6 is that we have regrouped the socio-economic categories, classifying HISCLASS categories 2, 4, 5, 10 and 11 in the intermediate socio-economic status group. The objective is to analyse whether the lack of evidence of social differences between categories is due to the inclusion in the wealthier group of certain occupations that, during epidemics, could have had a higher risk of death due to a higher exposure to the diseases through their work -such as doctors or clergymen. In this case, we can observe how the population group classified in the highest class had an advantage with respect to the individuals with the lowest SES. The value of the coefficients of the rest of the variables is fairly similar to that of those in model 6.

Next, we will present the calculations broken down into sub-periods. We have divided the period for which we have information on the causes of death into two sub-periods of six years (1801-1806 and 1807-1812). The objective is to analyse whether the catastrophic experience of the epidemic outbreak of yellow fever in 1804 could have influenced the response –and, in turn, the probabilities of death– of the different socioeconomic categories to the successive outbreaks of 1811 and 1812. With this breakdown into subperiods we also seek to analyse whether the war situation characterising much of the second sub-period and the associated short-term economic stress caused a variation in social differences in mortality (Campbell et al., 2004).

Table 3.5. Competing risk regression (1801-1806). Individuals over the age of seven

	Mortality derived from ill-defined diseases	Mortality derived from non-infectious diseases	Mortality derived from infectious diseases		
	Model 8	Model 9	Model 10	Model 11 (HISCLASS right side)	
1 (HISCLASS 1-5)	0.550*	0.951	1.095	0.913	1 (HISCLASS 1, 3)
2 (HISCLASS 6-9)	0.557**	0.964	0.929	0.984	2 (HISCLASS 2, 4-11)
3 (HISCLASS 10-12)	1	1	1	1	3 (HISCLASS 12)
Age	1.08***	1.037***	1.039***	1.039***	Age
Females	1	1	1	1	Females
Males	0.790	0.693•	1.145	1.153	Males
Neighbourhoods					Neighbourhoods
1) Intramuros	1.457	1.039	2.269***	2.315***	1) Intramuros
2) Arrabal	1.529	0.649	0.963	0.964	2) Arrabal
3) Crecimiento Sur	1.052	1.154	1.224	1.214	3) Crecimiento Sur
4) Crecimiento Oeste	1	1	1	1	4) Crecimiento Oeste
5) Ensanche Oeste	0.932	1.036	1.101	1.079	5) Ensanche Oeste
6) Barrio de Jesús	1.449	1.214	1.597	1.587	6) Barrio de Jesús
7) Outskirts	3.416**	1.712	0.826	0.795	7) Outskirts
Number of people living in the same home	0.939	0.961	1.023	1.024	Number of people living in the same home
Individuals (N)	3481	3481	3481	3481	Individuals (N)
Events	125	93	968	968	Events

• p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,00

Table 3.5 shows the values of the coefficients for the sub-period 1801-1806. A yellow fever outbreak occurred during this period, in 1804, which was of a smaller scale than those of the years 1811 and 1812. In model 8, referring to mortality derived from ill-defined diseases, we can observe a higher risk of death among the segments with the lowest socio-economic status. The differences are considerable. The individuals with the inferior socio-economic status had almost twice the risk than their wealthier neighbours. Age shows a positive correlation with the odds of death and, similarly to what we have observed for the period as a whole, men had a lower risk than women and living with a higher number of people diminished the possibilities of death due to ill-defined causes. With respect to the location of the home, the propensity of the neighbours living in the outskirts to die due to an ill-defined cause is higher than that of the rest of the neighbourhoods. With respect to mortality derived from non-infectious diseases –model 9– we can see how age again has a significant correlation with the risk of death. The socio-economic status, did not have a significant effect, nor did the place of residence. Again, men had lower levels of risk of death than women.

Model 10 shows the risk associated with the different independent variables in relation to the mortality derived from infectious diseases. In this case, men had a higher risk although the degree of significance is not high. With respect to the differences associated to the SES, we cannot observe any advantage derived from belonging to the higher status categories. On the contrary, although it is not significant, the value of the hazard ratio of the group with the highest status is higher than that of the reference category. Age is significant again, although with a lesser effect than that shown for mortality derived from ill-defined diseases. While not being significant, the number of people living in the same dwelling seems to have increased the probabilities of death.

In this model, the coefficients related to the geographic variable are those that experience the greatest change with respect to the values shown for the whole period. In this first sub-period, the neighbourhood that exhibited a higher hazard level was the *Intramuros* neighbourhood, followed by the *Barrio de Jesús*. At first glance, the result may seem surprising, given that, as commented in the descriptive analysis, these neighbourhoods were not the closest to the water sources –which could foster the reproduction of the mosquito vector of yellow fever– or those with the highest population density. However, after analysing the Minutes of the Local Healthcare Board of 1804, we know that, in this yellow fever outbreak, the first deaths occurred in homes located in these neighbourhoods and therefore a cordon sanitarie was established there[26]. For this sub-period, model 11 reflects the coefficients of the analysis after the afore-described reclassification of occupations. In this case no significant social

[26] The main internal cordon sanitarie was established on 7 November 1804 through the agreement of the Local Healthcare Board in the intra-wall area of the town. Specifically, in the block where the majority of the contagions were located, and the limits were set within the streets of *Padre Santiago, Calle del Mayordomo and Placeta del Hospital.* (AMV, Minutes Book of the Local Healthcare Board of Vera, file: 665).

differences in mortality derived from infectious diseases can be observed either. However, the value of the coefficients related to the highest and medium status categories is lower than that of the reference category.

Finally, Table 3.6 includes the exponents of the coefficients of the different variables for the sub-period 1807-1812. As previously explained, in these years the town of Vera suffered from two epidemic outbreaks –in the years 1811 and 1812–. Furthermore, between 1808 and 1814, Spain was immersed in the War of Independence against Napoleon's troops. The province of Almería was under French occupation from March 1810 to October 1812. As we can see in model 12, in this sub-period, a new social gradient can be observed in mortality derived from ill-defined diseases, although in this case it is not significant. The category made up of individuals with a higher social status reveals a greater advantage, followed by that with a medium status. The effect of age on the probability of dying also seems to have been greater than in the previous period and once more, is highly significant. Again, men had lower levels of risk of death than women. The number of people sharing a dwelling, while having little significance, can be observed again as a variable that reduced the odds of dying. Finally, with respect to the impact of the neighbourhood of residence, despite there being differences between the values of the coefficients, only the risk associated with the inhabitants living in the *Ensanche Oeste* is significantly lower. For this sub-period it has not been possible to estimate the importance of each of the variables of non-infectious mortality, given that the number of cases was very low (only 74 deaths).

The effect of the different variables during the period 1807-1812 on infectious mortality, which mainly consisted in deaths derived from yellow fever, is summarised in models 14 and 15. Model 14 shows the existence of a social gradient in mortality, belonging to both the group with the highest socio-economic status –HISCLASS 1-5–, and that with the medium status –6-9–, and reflects the advantage of these population segments with respect to the reference category –10-12–. As previously mentioned, resistance to the disease should not have been very different between classes due to the low effect that the nutritional status had on the effects of this disease. Therefore, the differences observed can be explained by an unequal capacity to reduce the exposure to the epidemic. In this sense, the documentary evidence available reveals that the opportunities of "escaping" from the disease were clearly different between social classes. The minutes of the Local Healthcare Board show that this strategy was used more frequently by the wealthier classes, particularly during the 1811 and 1812 outbreaks, which suggests the possibility of a kind of "learning" process after the recent experience of 1804. While for dates prior to 1804 we have no evidence that the town had suffered an epidemic of such a scale, in the years 1811 and 1812 the memory of the outbreak of 1804 was still very recent. On this occasion, the doctors warned of the risk of staying overnight in the town and it was common for the wealthier residents to leave the town centre to stay in the area of Garrucha which formed part of the outskirts of the municipality[27]. So, different reasons can explain why it was mainly the inhabitants with a higher status who temporarily abandoned the town centre. Among the causes giving rise to this situation, we can highlight the possession of economic resources so as to be able to stay in a property outside of the town (many were owners of property or farms in the country), better access to information, the different treatment received by the authorities or, simply, the facility to obtain the documents necessary to do so.

In this case, the differences between the sexes are not significant, while age is highly significant, but with a lower impact than in the case of mortality derived from ill-defined diseases. For these years, living with other individuals does not reveal significant results. Meanwhile, in the same way as the previous sub-period, the location of the dwelling can be observed to be an essential variable in relation to the possibilities of death due to infectious diseases. On this occasion, the neighbourhood with the highest risk of mortality was the *Crecimiento Oeste* - used as a reference. Again, in this sub-period, the geographic area that had a higher risk of death due to infectious disease was, once more, that in which the first infected people lived and where the first internal cordon sanitaire was imposed[28]. Living in the outskirts is again, the variable that most reduced the possibility of dying, followed by the *Intramuros* and *Crecimiento Sur* areas. Although we do not have data on morbidity, it is probable that the high incidence of the disease in the *Intramuros* district during the outbreak of 1804 favoured the presence of a high percentage of the population with immunity[29]. This circumstance would contribute to explaining the significantly lower probability of death among the population of this district during the outbreaks of 1811 and 1812.

Finally, in model 15, after the reclassification of the occupations, the differences between the risk related to the individuals in the category with the highest status and those with the lowest status, the category of reference, became more acute. The coefficients of the rest of the covariables are very similar to those in the second model. It is noteworthy that on this occasion, the inequality in mortality derived from infectious diseases was substantially higher than the previous sub-period. This seems to confirm the learning effect and the efficiency of the escaping strategy which, as we have seen, was used with more intensity by the wealthier segments of the town's population.

---

[27] By way of example, see points 6, 7 and 8 of the minutes of Vera's Healthcare Board of 2 October 1811, file 671, AMV. For greater detail on the sanitation and public healthcare measures implemented during the yellow fever outbreaks and the different economic and social problems, see Luque de Haro (2020, pp. 376-409).

[28] In the year 1811, a cordon sanitaire was imposed in *la Calle Ancha* and the surrounding avenues located in the *Crecimiento Oeste area.* The first death had been registered in this street one week before and, it seems that from then, the number of those infected in the area multiplied (AMV, Minutes Book of the Local Healthcare Board of Vera, file: 671).

[29] The individuals with yellow fever develop antibodies that are able to eliminate the virus and provide a long-lasting immunity (Restrepo, 2004, p. 74). The importance of the effect of immunity to diminish the intensity in the repetition of the outbreaks has been indicated by many authors (Iglesias Rodríguez, 1987; McNeill, 2010; Pritchett & Tunali, 1995).

Table 3.6. Competing risk regression (1807-1812) Individuals over the age of seven

	Mortality derived from ill-defined diseases	Mortality derived from non-infectious diseases	Mortality derived from infectious diseases		
	Model 12	Model 13	Model 14	Model 15 (HISCLASS right side)	
1 (HISCLASS 1-5)	0.644	0.493•	0.815•	0.648**	1 (HISCLASS 1, 3)
2 (HISCLASS 6-9)	0.716	0.686	0.845•	0.857•	2 (HISCLASS 2, 4-11)
3 (HISCLASS 10-12)	1	1	1	1	3 (HISCLASS 12)
Age	1.094***	1.036***	1.028***	1.029***	Age
Females	1	1	1	1	Females
Males	0.624*	0.686	0.937	0.939	Males
Neighbourhoods					Neighbourhoods
1) Intramuros	0.792		0.643***	0.648***	1) Intramuros
2) Arrabal	0.604		0.851	0.86	2) Arrabal
3) Crecimiento Sur	0.629		0.715*	0.712*	3) Crecimiento Sur
4) Crecimiento Oeste	1		1	1	4) Crecimiento Oeste
5) Ensanche Oeste	0.510•		0.927	0.924	5) Ensanche Oeste
6) Barrio de Jesús	0.378		0.73	0.73	6) Barrio de Jesús
7) Outskirts	0.414		0.5**	0.493*	7) Outskirts
Number of people living in the same home	0.946		0.995	0.996	Number of people living in the same home
Individuals (N)	2822	2822	2822	2822	Individuals (N)
Events	80	74	665	665	Events

• p < 0,1; * p < 0,05; ** p < 0,01; *** p < 0,00

In any case, when interpreting the results, it is important to take into account that the fact that the differences in mortality between socio-economic categories are not significant for some sub-periods or categories does not confirm the null hypothesis (the absence of social differences in mortality). We should not make the mistake of concluding that there is no link between the socio-economic category and the standard of living or mortality in a certain place or period just because the p-value of this variable is higher than 0.05 or because the confidence interval includes the value 0 (Amrhein et al., 2019).

4 CONCLUSIONS AND ELEMENTS FOR DISCUSSION

In this study we have analysed the social differences in adult mortality between the end of the eighteenth century and the beginning of the nineteenth century, using the data of the town of Vera in south-east Spain. During this period, the transformations in mortality described by the demographic and epidemiological transitions had not yet been experienced. The results obtained for the whole period reveal the existence of a social gradient in mortality in the three cause-of-death groups considered: infectious diseases, non-infectious diseases and ill-defined causes. However, the social differences were particularly intense in the mortality derived from ill-defined diseases. This phenomenon could be related to the predominance of poorer quality diagnoses in the death registers of the individuals of the lower classes. This probably reflects their limited access to medical services.

In relation to mortality derived from non-infectious diseases, the differences in the resources associated to the socio-economic status seem to have been important in determining the probabilities of death. These differences were especially intense during the 1807-1812 subperiod. Among the factors that could have influenced mortality, we can refer to the poorer living conditions, the deficient nutritional status or the physical deterioration associated to labour (Link & Phelan, 2002; Phelan et al., 2010). In this respect, the strong price increases experienced during the epidemic outbreaks and the following months, together with the institutional vacuum prevailing during certain moments of the War of Independence (Luque de Haro, 2019)[30] would have mainly affected the day labourers as they had a lower capacity to purchase food due to their low wages (Bengtsson, 2004, pp. 45-50; Fogel, 1989). This became even more pronounced if we add the economic and social shock generated by the outbreaks of 1811 and 1812. In this way, the individuals who remained in the town were not only exposed to the

[30] From the information contained in the *Minutes*, we know that in the months following the epidemic outbreak of 1804, the prices of cereals as much as doubled the average prices of the period. Meanwhile, during the first months of 1812, between the second and third yellow fever outbreaks, the prices of cereals increased by six times the average price (AMV, book 150-157 of the *Minutes*; AMV, state of fruits and manufactured goods, file 375).

epidemic, but they were also immersed in a situation of institutional absence in terms of welfare, sanitation and order, which made the situation even worse, particularly affecting the less well-off. Furthermore, in this second sub-period, the war situation, with the obligation of offering shelter and food to the troops that passed through, could have increased the mortality among the civil population, as it gave rise to a worsening of the sanitation conditions and a decrease in economic resources (Pérez Moreda, 1980, p. 84).

In relation to the social differences in the mortality derived from infectious diseases, our results show a low social gradient which is only significant when the groups of greater and lesser socio-economic status are compared, in line with the conclusions of the literature (Debiasi & Dribe, 2019; Leonard et al., 2015). The reasons that explain this circumstance are probably related to the great ease of contagion of the illnesses and their high virulence, due to the absence of appropriate treatments, among other factors, meaning that for a large proportion of these deaths the mortality was not substantially related to the nutritional level (Antonovsky, 1967, p. 67; Deaton, 2015, pp. 103-104; Lee, 1997; Smith, 1983). However, when analysing the results by sub-periods, we find noteworthy differences. In the first period, the social differences related to mortality derived from infectious diseases were not significant. On the other hand, in the second sub-period, particularly after the regrouping of occupations included in model 15, we can observe an important advantage of the category with the highest status. In any event, in both sub-periods, the majority of deaths derived from infectious diseases were due to epidemic outbreaks of yellow fever where the nutritional status of those infected was of little or no importance. Consequently, exposure to infectious diseases would have had a greater effect on mortality than the nutritional status. Therefore, the differences should be explained by another type of resources aimed at avoiding this exposure. In this sense, we have confirmed that the opportunities of "escaping" from the disease were clearly different between social classes.

With respect to the location of the household, we have observed that this had a considerable influence on mortality, particularly in relation to infectious diseases. In fact, for this type of deaths, in each sub-period the variable that had the greatest effect on the probabilities of death was the location of the home in the area where the first deaths by yellow fever occurred and where the first internal cordon sanitaire and quarantine measures were established –*Zona intramuros* in the outbreak of 1804 and *Crecimiento Oeste* in that of 1811–. In fact, it had an even higher effect than that of the socio-economic status, in line with the results of other studies (Edvinsson & Lindkvist, 2011). Meanwhile, living in the outskirts was the option that most increased the possibilities of survival with respect to the mortality derived from infectious diseases, unlike the case of non-infectious diseases or ill-defined diseases. Therefore, protective isolation against infections seemed to go hand in hand with a penalty in relation to other causes of mortality.

The results obtained reflect a higher mortality among women than men. These excess mortality rates among women are typical of pre-twentieth century societies, in which there was a high incidence of infectious diseases (Alter et al., 2004; Barford et al., 2006; Goldin & Lleras-Muney, 2018; Harris, 2008; Pinnelli & Mancini, 1998). However, our differential analysis shows that the higher risk for women was concentrated in deaths related to non-infectious diseases. The possible existence of a bias in our database, generated by a higher incidence of displacement among men due to military campaigns and seasonal migrations, means that we should take caution when drawing conclusions about the differences between the sexes.

For all of the disease categories taken together, an older age was associated to a higher risk of death. However, the value of the coefficient of this variable is lower in the mortality related to infectious diseases. The inclusion of cancer and chronic and degenerative diseases in the non-infectious diseases category, together with the importance of the cause *senectitude* among the ill-defined causes can help to explain this value.

The number of people in the household had the opposite effect on mortality derived from non-infectious diseases or ill-defined diseases than in that derived from infectious diseases. The value of the coefficients indicates that while a higher number of people living in a home increased the risk of dying from an infectious disease, this risk diminished in the case of non-infectious or ill-defined diseases. Therefore, the greater exposure to contracting infectious diseases related to sharing a home with a higher number of individuals would increase the odds of dying from this type of causes. On the contrary, the benefits associated to living in a family unit, such as lower income volatility or the possibility of receiving support at times of need, would have implied a lower risk of dying from non-infectious diseases (Lee et al., 2004).

Finally, we should point out that the effect of the different variables included in the models on the risk of death due to mortality derived from ill-defined diseases seems to coincide considerably with the results obtained for non-infectious mortality. This could reflect the prevalence of these causes of death among the diagnoses that we have not been able to classify. In any case, we assume that, in the more than two hundred cases of deaths due to ill-defined causes included in our calculations, there will be deaths of both categories –infectious and non-infectious mortality–. For this reason, we should warn that the probability that the social gradient observed in both types of diseases is greater than that reflected by the respective econometric models.

The relationship observed between the socio-economic status and mortality during the old demographic regime differs with other historical research, which predominantly studies populations from northern and central Europe, in which the presence of a social gradient in mortality is a relatively recent phenomenon (Bengtsson et al., 2020; Ekamper & van Poppel, 2019; Razzell & Spence, 2006; Schenk & Van Poppel, 2011). In this sense, our results coincide with the little evidence available for some southern European countries (Breschi et al., 2011; Fornasin et al., 2016). This fact could point to the possible existence of a regional pattern of social inequality in mortality during the first phases of the demographic and health transition, corresponding to those places where salaries were close to the subsistence level (Allen, 2001).

References:

Abramitzky, R., Mill, R., & Pérez, S. (2018). *Linking Individuals Across Historical Sources: a Fully Automated Approach* (N.º 24324). https://doi.org/10.3386/w24324

Allen, R. C. (2001). The great divergence in European wages and prices from the middle ages to the First World War. *Explorations in Economic History*, *38*(4), 411-447. https://doi.org/10.1006/exeh.2001.0775

Alter, G., Manfredini, M., & Nystedt, P. (2004). Gender Differences in Mortality. En J. Z. Lee, T. Bengtsson, & G. Alter (Eds.), *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900* (pp. 327-358). MIT press. https://doi.org/10.7551/mitpress/4227.001.0001

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*, 305-307. https://doi.org/10.1038/d41586-019-00857-9

Antonie, L., Inwood, K., & Ross, J. A. (2015). Dancing with Dirty Data: Problems in the Extraction of Life-Course Evidence from Historical Censuses. En *Population Reconstruction* (pp. 217-241). Springer International Publishing. https://doi.org/10.1007/978-3-319-19884-2_11

Antonovsky, A. (1967). Social Class, Life Expectancy and Overall Mortality Reviewed work. *The Milbank Memorial Fund Quarterly*, *45*(2), 31-73.

Arrizabalaga, J. (1993). La identificación de las causas de muerte en la Europa pre-industrial: algunas consideraciones historiográficas. *Boletín de la ADEH*, *XI*(3), 23-48.

Austin, P. C., & Fine, J. P. (2017). Practical recommendations for reporting Fine-Gray model analyses for competing risk data. *Statistics in Medicine*, *36*(27), 4391-4400. https://doi.org/10.1002/sim.7501

Barford, A., Dorling, D., Smith, G. D., & Shaw, M. (2006). Life expectancy: women now on top everywhere. *BMJ*, *332*, 808. https://doi.org/https://doi.org/10.1136/bmj.332.7545.808

Barona Vilar, J. L. (1993). Teorías médicas y la clasificación de las causas de muerte. *Boletín de la ADEH*, *XI*(3), 49-64.

Beltrán Tapia, F. J., & Gallego-Martínez, D. (2017). Where are the missing girls? Gender discrimination in 19th-century Spain. *Explorations in Economic History*, *66*, 117-126. https://doi.org/10.1016/j.eeh.2017.08.004

Bengtsson, T. (2004). Living Standards and Economic Stress. En T. Bengtsson, C. Campbell, & J. Z. Lee (Eds.), *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900* (pp. 27-59). MIT Press. https://doi.org/https://doi.org/10.7551/mitpress/4227.001.0001

Bengtsson, T., & Dribe, M. (2011). The late emergence of socioeconomic mortality differentials: A micro-level study of adult mortality in southern Sweden 1815-1968. *Explorations in Economic History*, *48*(3), 389-400. https://doi.org/10.1016/j.eeh.2011.05.005

Bengtsson, T., Dribe, M., & Helgertz, J. (2020). When did the Health Gradient Emerge? Social Class and Adult Mortality in Southern Sweden, 1813-2015. *Demography*. https://doi.org/https://doi.org/10.1007/s13524-020-00877-5

Bengtsson, T., & Lindström, M. (2000). Childhood misery and disease in later life: The effects on mortality in old age of hazards experienced in early life, southern Sweden, 1760-1894. *Population Studies*, *54*(3), 263-277. https://doi.org/10.1080/713779096

Bengtsson, T., & Van Poppel, F. (2011). Socioeconomic inequalities in death from past to present: An introduction. *Explorations in Economic History*, *48*(3), 343-356. https://doi.org/10.1016/j.eeh.2011.05.004

Bernabeu-Mestre, J. (1992). *Las estadísticas demográfico-sanitarias*. Instituto de Salud Carlos III, Centro Nacional de Epidemiología.

Bernabeu-Mestre, J. (1993). Expresiones diagnósticas y causas de muerte. Algunas reflexiones sobre su utilización en el análisis demográfico de la mortalidad. *Revista de Demografía Histórica*, *11*(3), 11-22.

Bertillon, J. (1899). Reglas que deben seguirse para la clasificación de las enfermedades en la nomenclatura. En *Nomenclatura de las enfermedades (causas de las defunciones por...)* (pp. 3-55). Imprenta de la Dirección General del Instituto Geográfico y Estadístico.

Blossfeld, H.-P., & Rohwer, G. (2002). *Techniques of event history modeling: New approaches to casual analysis*. Lawrence Erlbaum Asssociates.

Breschi, M., Fornasin, A., Manfredini, M., Mazzoni, S., & Pozzi, L. (2011). Socioeconomic conditions, health and mortality from birth to adulthood, Alghero 1866-1925. *Explorations in Economic History*, *48*(3), 366-375. https://doi.org/10.1016/j.eeh.2011.05.006

Campbell, C., Lee, J. Z., & Bengtsson, T. (2004). Economic Stress and Mortality. En T. Bengtsson, C. Campbell, & J. Z. Lee (Eds.), *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900* (pp. 61-84). MIT Press. https://doi.org/10.7551/mitpress/4227.001.0001

Cano, L. (2015). *Historia Urbana de Vera. Castillo en la traza*. Universidad de Granada.

Cano, L. (2019). Historia Urbana de Vera. Castill@ en la Traza. En V. Luque de Haro & M. Caparrós Perales (Eds.), *La tierra de Vera: Nuevas contribuciones sobre la historia de un territorio de frontera* (pp. 245-304). Edual.

Cassel, J. (1976). The contributution of the social environment to host resistance: The fourth Wada Hampton Frost lecture. *American Journal of Epidemiology*, *104*(2), 107-123. https://doi.org/10.1093/oxfordjournals.aje.a112281

Chamberlain, G. (2006). British maternal mortality in the 19th and early 20th centuries. *Journal of the Royal Society of Medicine*, *99*(11), 559-563. https://doi.org/https://doi.org/10.1177/014107680609901113

Christen, P. (2012). *Data Matching. Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Springer. https://doi.org/10.1007/978-3-642-31164-2

Christen, P., Vatsalan, D., & Fu, Z. (2015). Advanced Record Linkage Methods and Privacy Aspects for Population Reconstruction—A Survey and Case Studies. En G. Bloothooft, P. Christen, K. Mandemakers, & M. Schraagen (Eds.), *Population Reconstruction* (pp. 87-110). Springer International Publishing. https://doi.org/10.1007/978-3-319-19884-2_5

Clausen, N. F. (2015). The Danish Demographic Database—Principles and Methods for Cleaning and Standardisation of Data. En *Population Reconstruction* (pp. 3-22). Springer International Publishing. https://doi.org/10.1007/978-3-319-19884-2_1

Cleves, M., Gould, W., Gould, W. W., Gutierrez, R., & Marchenko, Y. (2008). *An introduction to survival analysis using Stata*. Stata press.

Clouston, S. A. P., Rubin, M. S., Phelan, J. C., & Link, B. G. (2016). A Social History of Disease: Contextualizing the Rise and Fall of Social Inequalities in Cause-Specific Mortality. *Demography*, *53*(5), 1631-1656. https://doi.org/10.1007/s13524-016-0495-5

Cox, D. R. (1972). Regression Models and Life-Tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, *34*(2), 187-220. https://doi.org/10.2307/2985181

Dahlgren, G., & Whitehead, M. (1991). *Policies and strategies to promote social equity in health*. Institute for future studies.

Deaton, A. (2015). *El gran escape: salud, riqueza y los orígenes de la desigualdad*. Fondo de Cultura Económica.

Debiasi, E., & Dribe, M. (2019). Social class inequalities in cause-specific adult mortality in a long-term perspective. *Lund Papers in Economic Demography*, *4*.

Dopico, F., & Rowland, R. (1990). Demografia del censo de Floridablanca. Una aproximacion. *Revista de Historia Económica / Journal of Iberian and Latin American Economic History*, *8*(3), 591-618. https://doi.org/10.1017/S0212610900002391

Edvinsson, S., & Broström, G. (2012). Old age, health, and social inequality: Exploring the social patterns of mortality in 19th century northern Sweden. *Demographic Research*, *26*, 633-660. https://doi.org/10.4054/DemRes.2012.26.23

Edvinsson, S., & Broström, G. (2017). Life course and long-term perspectives of social inequality in mortality among elderly and adults in Northern Sweden 1801-2013. *diva-portal.org*.

Edvinsson, S., & Lindkvist, M. (2011). Wealth and health in 19th Century Sweden. A study of social differences in adult mortality in the Sundsvall region. *Explorations in Economic History*, *48*(3), 376-388. https://doi.org/10.1016/j.eeh.2011.05.007

Efremova, J., Ranjbar-Sahraei, B., Rahmani, H., Oliehoek, F. A., Calders, T., Tuyls, K., & Weiss, G. (2015). Multi-Source Entity Resolution for Genealogical Data. En *Population Reconstruction* (pp. 129-154). Springer International Publishing. https://doi.org/10.1007/978-3-319-19884-2_7

Ekamper, P., & van Poppel, F. (2019). Infant mortality in mid-19th century Amsterdam: Religion, social class, and space. *Population, Space and Place*, *25*(4), e2232. https://doi.org/10.1002/psp.2232

Elo, I. T., Martikainen, P., & Myrskylä, M. (2014). Socioeconomic status across the life course and all-cause and cause-specific mortality in Finland. *Social Science & Medicine*, *119*, 198-206. https://doi.org/10.1016/J.SOCSCIMED.2013.11.037

Erikson, R., & Torssander, J. (2008). Social class and cause of death. *The European Journal of Public Health*, *18*(5), 473-478. https://doi.org/10.1093/eurpub/ckn053

Escudero, A. (2002). Volviendo a un viejo debate : el nivel de vida de la clase obrera británica durante la Revolución Industrial. *Revista de Historia Industrial*, *21*, 13-60.

Fellegi, I. P., & Sunter, A. B. (1969). A Theory for Record Linkage. *Journal of the American Statistical Association*, *64*(328), 1183-1210. https://doi.org/10.1080/01621459.1969.10501049

Ferrie, J. P. (2003). The Rich and the Dead Socioeconomic Status and Mortality in the U. S., 1850-1860. En D. L. Costa (Ed.), *Health and Labor Force Participation Over the Life Cycle: Evidence from the Past* (pp. 11-50). National Bureau of Economic Resarch.

Fine, J. P., & Gray, R. J. (1999). A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*, *94*(446), 496-509. https://doi.org/10.1080/01621459.1999.10474144

Fogel, R. W. (1989). *Second Thoughts on the European Escape from Hunger: Famines, Price Elasticities, Entitlements, Chronic Malnutrition, and Mortality Rates* (N.º 1; NBER Historical Working Paper). https://doi.org/10.3386/h0001

Fornasin, A., Breschi, M., & Manfredini, M. (2016). Environment, Housing, and Infant Mortality: Udine, 1807–1815. En D. Ramiro-Fariñas & M. Oris (Eds.), *New Approaches to Death in Cities during the Health Transition* (pp. 43-54). Springer. https://doi.org/10.1007/978-3-319-43002-7

Freese, J., & Lutfey, K. (2011). Fundamental Causality: Challenges of an Animating Concept for Medical Sociology. En B. A. Pescosolido, J. K. Martin, J. D. McLeod, & A. Rogers (Eds.), *Handbook of the Sociology of Health, Illness, and Healing* (pp. 67-81). Springer New York. https://doi.org/10.1007/978-1-4419-7261-3_4

Gagnon, A., Tremblay, M., Vézina, H., & Seabrook, J. A. (2011). Once were farmers: Occupation, social mobility, and mortality during industrialization in Saguenay-Lac-Saint-Jean, Quebec 1840-1971. *Explorations in Economic History*, *48*(3), 429-440. https://doi.org/10.1016/j.eeh.2011.05.008

García Pérez, M. (2007). El padrón municipal de habitantes: origen, evolución y significado. *Hispania Nova: Revista de historia contemporánea*, *7*(7), 11.

Goldin, C., & Lleras-Muney, A. (2018). *XX>XY?: The Changing female advantage in life expectancy* (NBER Working Paper Series, Número 24716).

Gómez Díaz, D. (1995). *Las migraciones almerienses. Una historia económica hasta 1910*. Instituto de Estudios Almerienses. Diputación Provincial de Almería.

Harris, B. (2004). Public health, nutrition, and the declinein mortality: The McKeown Thesis Revisited. *Social History of Medicine*, *17*(3), 379-407.

Harris, B. (2008). Gender, health, and welfare in England and Wales since industrialisation. *Research in Economic History*, *26*, 157-204. https://doi.org/10.1016/S0363-3268(08)26003-9

Harris, B. (2009). Anthropometric history, gender and the measurement of well-being. En L. Gálvez (Ed.), *Gender and Well-Being in Europe: Historical and Contemporary Perspectives*. Routledge. https://doi.org/https://doi.org/10.4324/9781315584034

Harris, B. (2019). Anthropometric Historiy and the Measurement of Wellbeing. *Anual Conference of the Economic History Society*.

Harris, B., & Helgertz, J. (2019). Urban sanitation and the decline of mortality. *The History of the Family*, *24*(2), 207-226. https://doi.org/https://doi.org/10.1080/1081602X.2019.1605923

Humphries, J. (1991). 'Bread and a pennyworth of treacle': excess female mortality in England in the 1840s. *Cambridge Journal of Economics*, *15*, 451-473. https://doi.org/10.1093/oxfordjournals.cje.a035183

Iglesias Rodríguez, J. J. (1987). *La epidemia gaditana de fiebre amarilla de 1800*. Diputación provincial de Cádiz.

Jaro, M. A. (1989). Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. *Journal of the American Statistical Association*, *84*(406), 414-420.

Kleinbaum, D. G., & Klein, M. (2012). *Survival Analysis, A Self-Learning Text* (Third Edit). Springer. https://doi.org/10.1007/0-387-29150-4

Lau, B., Cole, S. R., & Gange, S. J. (2009). Competing risk regression models for epidemiologic data. *American Journal of Epidemiology*, *170*(2), 244-256. https://doi.org/10.1093/aje/kwp107

Lee, C. (1997). Socioeconomic Background, Disease, and Mortality among Union Army Recruits: Implications for Economic and Demographic History. *Explorations in Economic History*, *34*(1), 27-55. https://doi.org/10.1006/EXEH.1996.0661

Lee, J. Z., Bengtsson, T., & Campbell, C. (2004). Family and Community. En T. Bengtsson, C. Campbell, & J. Z. Lee (Eds.), *Life under Pressure: Mortality and Living Standards in Europe and Asia, 1700-1900* (pp. 85-105). MIT Press.

Leonard, S. H., Robinson, C., Swedlund, A. C., & Anderton, D. L. (2015). The effects of wealth, occupation, and immigration on epidemic mortality from selected infectious diseases and epidemics in Holyoke township, Massachusetts, 1850−1912. *Demographic Research*, *33*, 1035-1046. https://doi.org/10.4054/DemRes.2015.33.36

Link, B. G., & Phelan, J. C. (1995). Social Conditions as Fundamental Causes of Disease. *Journal of Health and Social Behavior*, *35*, 80-94. https://doi.org/10.2307/2626958

Link, B. G., & Phelan, J. C. (1996). Understanding sociodemographic differences in health--the role of fundamental social causes. *American Journal of Public Health*, *86*(4), 471-473. https://doi.org/10.2105/AJPH.86.4.471

Link, B. G., & Phelan, J. C. (2002). McKeown and the Idea That Social Conditions Are Fundamental Causes of Disease. *American Journal of Public Health*, *92*(5), 730-732. https://doi.org/10.2105/AJPH.92.5.730

Llopis Pérez, J. (1996). La estadística: una orquesta hecha instrumento. *Ariel, Madrid*.

Llopis Pérez, J. (2014). *La Estadística: Una Orquesta Hecha Instrumento*. https://estadisticaorquestainstrumento.wordpress.com/

Luque de Haro, V. (2019). Veratenses ante la muerte: Transición demográfica y epidemias a principios del siglo XIX. En V. Luque de Haro & M. Caparrós Perales (Eds.), *La tierra de Vera: Nuevas contribuciones sobre la historia de un territorio de frontera* (pp. 345-378). Edual.

Luque de Haro, V. (2020). *Análisis histórico del nivel de vida en Andalucía. Desigualdad social y mortalidad: ensayo interpretativo y estudio empírico de un caso en el sureste español*. Universidad de Almería.

Mamelund, S.-E. (2006). A socially neutral disease? Individual social class, household wealth and mortality from Spanish influenza in two socially contrasting parishes in Kristiania 1918–19. *Social Science & Medicine*, *62*(4), 923-940. https://doi.org/10.1016/J.SOCSCIMED.2005.06.051

Marmot, M. (2004). Status syndrome. *Significance*, *1*(4), 150-154. https://doi.org/10.1111/j.1740-9713.2004.00058.x

Marmot, M. (2005). Public Health Social determinants of health inequalities. *Lancet*, *365*, 1099-1104. https://doi.org/10.1016/S0140-6736(05)71146-6

McKeown, T. (1976). *The modern rise of population*. Edward Arn.

McNeill, J. R. (2010). Mosquito empires: Ecology and war in the greater Caribbean, 1620–1914. En *Mosquito Empires: Ecology and War in the Greater Caribbean, 1620-1914*. https://doi.org/10.1017/CBO9780511811623

Molitoris, J., & Dribe, M. (2016). Industrialization and inequality revisited: Mortality differentials and vulnerability to economic stress in Stockholm, 1878-1926. *European Review of Economic History*, *20*(2), 176-197. https://doi.org/10.1093/ereh/hev023

Nadal, J. (1966). *La población española: siglos XVI a XX*. Ariel.

Perdiguero Gil, E. (1993). Causas de muerte y relación entre conocimiento científico y conocimiento popular. *Boletín de la ADEH*, *XI*(3), 65-88.

Pérez Moreda, V. (1980). Las crisis de mortalidad en la España interior. Siglos XVI-XIX. En *Madrid* (Vol. 36). Siglo Veintiuno Editores.

Perrenoud, A. (1975). L'inégalité sociale devant la mort à Genève au XVIIIe siècle. *Population*, *30*(1), 221-243.

Perrenoud, A. (1981). Surmortalité féminine et condition de la femme (XVIIe-XIXe siècles). Une vérification empirique. *Annales de démographie historique*, 89-104.

Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social Conditions as Fundamental Causes of Health Inequalities: Theory, Evidence, and Policy Implications. *Journal of Health and Social Behavior*, *51*(1_suppl), S28-S40. https://doi.org/10.1177/0022146510383498

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Pinnelli, A., & Mancini, P. (1998). Mortality Peaks in Italy in the Late 19th and Early 20th Centuries: Trends by Age and Sex. *European Journal of Population/ Revue europenne de Dmographie*, *14*(4), 333-365. https://doi.org/10.1023/A:1006182029835

Pintilie, M. (2011). Análisis de riesgos competitivos. *Revista Española de Cardiología*, *64*(7), 599-605. https://doi.org/10.1016/j.recesp.2011.03.017

Pressat, R. (1985). *Introducción a la Demografía*. Ariel.

Preston, S., & Haines, M. (1991). *Fatal Years*. Princeton University Press.

Pritchett, J. B., & Tunali, I. (1995). Strangers′ Disease: Determinants of Yellow Fever

Mortality during the New Orleans Epidemic of 1853. *Explorations in Economic History*, *32*(4), 517-539. https://doi.org/10.1006/EXEH.1995.1022

Pujadas-Mora, J. M., Romero Marín, J., & Villar Garruta, C. (2014). Propuestas metodológicas para la aplicación de HISCO en el caso de Cataluña, siglos XV-XX. *Revista de Demografía Histórica*, *32*(1), 181-219.

Ramiro Fariñas, D., Sanz Gimeno, A., Bernabeu Mestre, J., & Robles González, E. (2003). De expresiones diagnósticas a causas de muerte: Una propuesta metodológica para el análisis de la mortalidad. *Revista de Demografía Histórica*, *21*(1), 167-193.

Razzell, P., & Spence, C. (2006). The Hazards of Wealth: Adult Mortality in Pre-Twentieth-Century England. *Social History of Medicine*, *19*(3), 381-405. https://doi.org/10.1093/shm/hkl048

Recaño, J., & Esteve, A. (2006). Revisitando Garcia Faria: Un estudio de los factores espaciales y medioambientales de la mortalidad en la Barcelona de finales del siglo XIX. *Revista de Demografía Histórica*, *XXIV*(1), 121-180.

Reher, D. S., & Valero Lobo, A. (1995). *Fuentes de información demográfica en España* (Vol. 13). CIS.

Restrepo, B. N. (2004). Fiebre Amarilla. *Revista CES MEDICINA*, *18*(1), 69-82.

Rodríguez Ocaña, E., & Martínez Navarro, F. (2008). Salud pública en España: de la Edad Media al siglo XXI. *Nueva Salud Pública*, 1-158.

Ruggles, S. (1992). Migration, Marriage, and Mortality: Correcting Sources of Bias in English Family Reconstitutions. *Population Studies*, *46*(3), 507-522. https://doi.org/10.1080/0032472031000146486

Ruggles, S. (2002). Linking Historical Censuses: a New Approach. *History and Computing*, *14*(1-2), 213-224. https://doi.org/10.3366/hac.2002.14.1-2.213

Ruggles, S., Fitch, C. A., & Roberts, E. (2018). Historical Census Record Linkage. *Annual Review of Sociology*, *44*(1), 19-37. https://doi.org/10.1146/annurev-soc-073117-041447

Sánchez Picón, A. (1988). Marchar a las Andalucías. Un episodio migratorio en la Almería del siglo XIX. En *Homenaje al Padre Tapia*. Unicaja.

Sánchez Picón, A. (2011). Ejes de desarrollo en el Sureste y la Andalucía mediterránea (1800-2000). *Historia Contemporánea*, *42*, 261-302.

Sayers, A., Ben-Shlomo, Y., Blom, A., & Steele, F. (2015). Probabilistic record linkage. *International journal of epidemiology*, *45*(3), 954-964. https://doi.org/https://doi.org/10.1093/ije/dyv322

Schenk, N., & Van Poppel, F. (2011). Social class, social mobility and mortality in the Netherlands, 1850-2004. *Explorations in Economic History*, *48*(3), 401-417. https://doi.org/10.1016/j.eeh.2011.05.010

Schofield, R. (1986). Did the mothers really die? Three centuries of maternal mortality in the world we have lost. En L. Bonfield, R. M. Smith, & K. Wrightson (Eds.), *The world we have gained: histories of population and social structure* (pp. 231-260). Basil Blackwell.

Schofield, R., & Reher, D. S. (1991). The decline of mortality in Europe. En R.

Schofield, D. S. Reher, & A. Bideau (Eds.), *The decline of mortality in Europe* (pp. 1-17). Clarendon Press.

Schumacher, R., & Oris, M. (2011). Long-term changes in social mortality differentials, Geneva, 1625-2004. *Explorations in Economic History*, *48*(3), 357-365. https://doi.org/10.1016/j.eeh.2011.05.011

Scrucca, L., Santucci, A., & Aversa, F. (2010). Regression modeling of competing risk using R: An in depth guide for clinicians. *Bone Marrow Transplantation*, *45*(9), 1388-1395. https://doi.org/10.1038/bmt.2009.359

Smith, D. S. (1983). Differential Mortality in the United States before 1900. *The Journal of Interdisciplinary History*, *13*(4), 735-759. https://doi.org/10.2307/203888

The Relationship of Nutrition, Disease, and Social Conditions: A Graphical Presentation. (1983). *The Journal of Interdisciplinary History*, *14*(2), 503-506. https://doi.org/10.2307/203718

Toch-Marquardt, M., Menvielle, G., Eikemo, T. A., Kulhánová, I., Kulik, M. C., Bopp, M., Esnaola, S., Jasilionis, D., Mäki, N., Martikainen, P., Regidor, E., Lundberg, O., & Mackenbach, J. P. (2014). Occupational Class Inequalities in All-Cause and Cause-Specific Mortality among Middle-Aged Men in 14 European Populations during the Early 2000s. *PLoS ONE*, *9*(9), 1-11. https://doi.org/10.1371/journal.pone.0108072

Urquía, M. L. (2006). *Teorías dominantes y alternativas en epidemiología* (II). Universidad Nacional de Lanús, Departamento de Salud Comunitaria.

Vallin, J. (2013). Inequalities in Life Expectancy Between and Within European Countries. En *The Demography of Europe* (pp. 139-173). Springer Netherlands. https://doi.org/10.1007/978-90-481-8978-6_7

Vallin, J., & Meslé, F. (2004). Convergences and divergences in mortality. A new approach to health transition. *Demographic Research*. https://doi.org/10.4054/DemRes.2004.S2.2

Van Leeuwen, M., & Maas, I. (2011). *Hisclass: A Historical International Social Class Scheme*. Universitaire Pers Leuven. https://doi.org/10.1080/01615440.2012.721340

Van Leeuwen, M., Maas, I., & Miles, A. (2002). *HISCO: Historical international standard classification of occupations*. Universitaire Pers Leuven.

Van Poppel, F., Bijwaard, G., Van Lieburg, M., Van Lieburg, F., Hoekstra, R., & Verkade, F. (2016). The Life Expectancy of Medical Professionals in the Netherlands, Sixteenth to Twentieth Centuries. *Population*, *71*(4), 619-640. https://doi.org/https://doi.org/10.3917/popu.1604.0659

Winkler, W. E. (1990). *String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage.* Proceedings of the Section on Survey Research Methods.

Winkler, W. E. (1999). The State of Record Linkage and Current Research Problems. *Statistical Research Division, U.S. Census Bureau*.