Chaire Quetelet 2006 - 29/11 – 1/12 Louvain-la-Neuve (Belgium)
SECTION - Longitudinal studies: from retrospective approaches to follow-up

# The Use of Family Reconstructed Database in the Study of Human Longevity: The Population of Villagrande Strisaili (Sardinia)

Luisa Salaris[1,2] and Michel Poulain [1]
(1) GéDAP, Groupe d'études de démographie appliquée
Université Catholique de Louvain-la-Neuve
Place Montesquieu, 1, bte 4
1348 Louvain-La-Neuve (Belgium)
(2) DRES - Dip. Ricerche Economiche e Sociali
Facoltà di Scienze Politiche, Università degli Studi di Cagliari (Italy)

Corresponding author e-mail: salaris@sped.ucl.ac.be

**ABSTRACT**
The use of family reconstructed database represents a technique widely applied both in historical demography and in genetic. However in both fields few large family reconstructed databases are suitable for the study of longevity. The following communication aims to present a unique database build using data collected on a population experiencing exceptional longevity in Sardinia (Italy). Each individual is followed from birth to death, going through marriage and family formation. The Villagrande Longevity Database (VILD) includes all newborns in the village in a period of 50 years (1866-1915). The level of coverage reached is exceptionally high, thanks to the completeness of data sources used and the high endogamy registered. By considering family links within the whole population, this database enables to consider in the analysis also the effect of familial characteristics toward survival and extreme longevity. Preliminary results are presented in comparison with finding from other studies on human survival.

## Introduction

An important contribution in the study of human longevity and its determinants derives from the use of existing genealogical records. Gavrilova et al. (1999) proposed a review of available genealogical data resources for the study of human lifespan and pointed out a lack of appropriate data, underlying different issues concerning data quality and accessibility, biases, and critical levels of completeness. Parallel to existing genealogical records, there are family reconstructed databases collected on purpose by researchers in biodemography and historical demographers, which are derived from individual records on families and/or specific identified communities' members. The initiator of the family reconstruction methodology has been Henry (Fleury and Henry, 1956; Henry and Blum, 1988) who developed a detailed methodology in order to reconstruct all families among the population of one or more parishes. Originally these databases were not computerised. All the information for a given family was framed within a paper family sheet and this made difficult the establishment of more complex family links within the population concerned. When computers were used in the early '70s, the concept of family reconstruction was extended to larger populations. Among these databases, the most important is certainly the PRDH database which considered the whole population of Québec born since 1750 (Beauchamp et al., 1977). In general, to

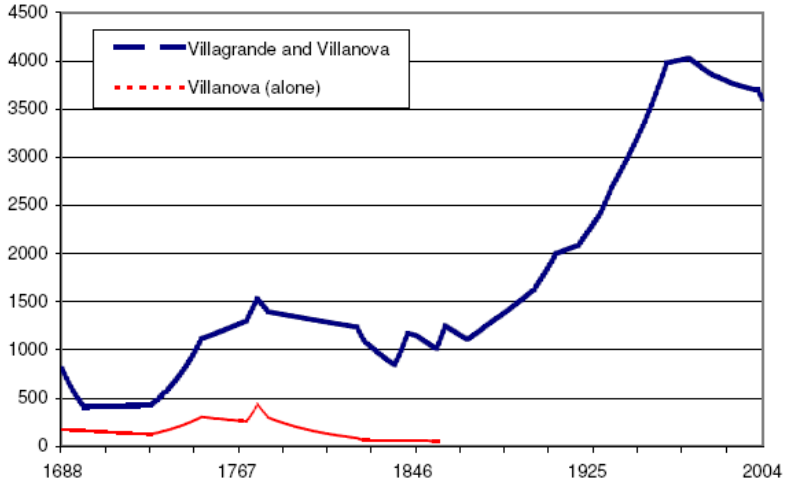
study more in details the mechanism behind mortality and longevity large populations are needed as a too small sample of oldest olds could bias the analysis. Accordingly, as the creation of a new database implies a great effort especially in data collection to cover large populations, only few databases exist nowadays that allows studying longevity. Among these, the most representative are the Balzac database for the recent population of Québec (Bouchard et al., 1989) or the Valserine database (Cournil, 2000).
The idea of building up a population database for the whole population of one village to study today longevity in Sardinia was stimulated by the results of previous researches on longevity in the region: a higher proportion of validated centenarians and an associated low sex ratio among them (Deiana et al., 1999; Salaris, 2005; Caselli et al., 2006). Centenarians in Sardinia are not equally distributed within its territory and a recent research identified in the island the so-called *Blue Zone* (BZ), an area characterized by the highest prevalence of male centenarians never met elsewhere (Poulain et al., 2004). Among all the municipalities included in the BZ, our attention was caught by the municipality of Villagrande Strisaili as it recorded the highest level of longevity with the probability for a newborn out of one hundred to reach 100 years old for males born between 1880 and 1901 (Poulain et al. 2006). A family reconstructed database was build for this village whose main feature are described in the next section.

## 1. Why choosing to study Villagrande Strisaili?

The municipality of Villagrande Strisaili was selected since it registers the highest value of ELI (Extreme Longevity Index, the probability for a newborn to reach 100). This village may suspect to be an example of long-living population where a significant number of oldest olds aged 90 years and over may be found and analysed in order to identify on an appropriate way the familial determinants of longevity.
The municipality of Villagrande Strisaili which consists of the urban centre of Villagrande Strisaili itself and of the small fraction of Villanova Strisaili has an administrative territory of 21,000 hectares and it is located at 700 meters above the level of the sea in the region of Ogliastra in the central-eastern part of Sardinia (Cannas and Rubiu, 1977). At January 1, 2005 the municipality registered 3,577 inhabitants[1].
When looking at the evolution of the population size of the village in the long term, 3 major phases can be identified. From 1688 to almost the beginning of the last century (phase I), the population village appears almost stationary, not registering drastic growth periods (figure 1). However, this trend has to be interpreted with caution as the data available for that period prior to 1861 may not estimate the real situation, as data was collected for fiscal and military reasons and no civil registration was established (Corridore, 1902). Starting from the XIX$^{th}$ century the population known a continuous growth up to 1971 census (phase II), since when for the first time is registered a decreasing in the population stock (begin of phase III). This decreasing trend occurring since 60-70s years is not exclusively registered in Villagrande Strisaili, but is a trend involving most of the municipality from the inland areas as consequences of consistent out-migration of labour force. These out-migration flows are the result of a progressively marked gap between the economic development of rural and urban areas, which - starting after the Second World War - has trigged a phenomenon of depopulation of smaller centres in favour of bigger cities located mainly in the coast and along the main communication networks (Gatti and Puggioni,1998).

[1] Data source: Istat (http://demo.istat.it/pop2005/index.html)

**Figure 1 – Population of Villagrande Strisaili, 1688-2004**




## 2. Aim and strengths of the VILD Database

The major aim of Villagrande Longevity Database (VILD) is the identification of the possible determinants of longevity in its population. The choice of a single setting for population reconstruction is not new. It represents in fact a well-know approach in historical demography. However, in the building up procedure of such of database to the great effort to reach a satisfactory level of coverage and a sufficient number of oldest olds, it should be added accessibility constrains. These aspects discourage the data collection for enlarged databases developed for the main purpose of longevity studies. Nevertheless, this type of data collection presents several advantages compared to the classical "centenarian case-control studies" usually used in researches on longevity. In fact the latter only consider in the database long-lived individuals alive, their family and a limited number of control cases without family links. Moreover, in the case-control studies individuals under study, centenarians or nonagerians, are often distributed in large heterogeneous areas, making difficult for researchers to generalize results.

At the opposite the building up of a family reconstructed database presents several strengths which pay off of the great effort needed for the data collection. The first strength is certainly represented by the possibility to collect information not only on single individuals but also on all family links within the population concerned. This additional information makes possible to consider for each individual which familial variables may have some effect on his/her survival. Moreover using these family links may allow to identify whether or not long-living families exist and to detect possible relevant factors for heterogeneity among members of the same community.

The second strength of this approach is that all persons alive or not at the time of the study are considered in the data collection including those who died younger. Not exclusively longer survivors are included but both long-living and early dying individuals. Moreover, the adoption a longitudinal perspective helps to observe the complete extinction of the selected cohorts of newborns[2].

The third strength of studying a single setting is that it enables to control for a certain number of confounding variables. The influence of both physical environment and culture can be in

[2] There were recorded 43 still alive newborns in VILD at September, 2006.

fact excluded as individuals share the same location. Moreover, it makes possible to perform multi-level analysis as individual, familial and community variables may be considered.
Finally, in the case of Villagrande Strisaili, the high coverage level guarantees against selection biases and the database can be considered really representative of the entire population. Also in the case of Villagrande Strisaili the existence of a highly endogamic marriage market let us assume that the individuals in VILD share a common genetic pool and using the information available in the parish registers back to the XV$^{th}$ century it may be possible to detect families' origins determining from many generations they were settled in the village and – in case of migration - to assess when and from where they arrived.

## 3. Data

### 3.1. Data sources and Data Collection

The data contained in VILD refers to all individuals born in Villagrande Strisaili from 1866 to 1915. The choice of this specific period has not been casual. In fact 1866 is the first year since civil registers are available at municipality level. As the major interest is in measuring individuals' lifespan, we decided instead to close our observational window with the births occurred in 1915, as individuals belonging to this cohort may be today at least 90 years old. Currently, the VILD comprise a total of 6,557 individuals, among which 2,589 were born in Villagrande Strisaili during the period 1866-1915[3].
The major goal in the building procedure of the VILD was to include a significant number of oldest olds and to reach the highest possible level of known survival. For this purpose a multi-source approach was used. Our focus in fact was not in the community reconstruction as such, but on individuals and their survival history. Therefore the VILD was created in different steps and using different data sources: civil status registers, parish registers, population registers, military lists and in exceptional cases interviews of still living relatives of our newborns (Ego).
The main data source used has been the *civil status registers*. These registers consist of different volumes. Each volume reports on annual base all births, marriages and death occurred in the municipality (Del Panta and Rettaroli, 1994). They are available in the demographic office of the municipality of Villagrande Strisaili since 1866 till nowadays. Each record reports the following information:

> BIRTH RECORD: date of birth, place of birth with indication of the street or the neighbourhood, name/s and surname of the newborn, sex, parents' name and occupation.
>
> MARRIAGE RECORD: date of marriage, name and surname of the spouses, their age, place of birth and place of residence, names of parents', sometimes it reports previous marital status.
>
> DEATH RECORD: date of death, name and surname, age at death, place of birth and of residence, parents' name and surname, marital status and eventually the name of the partner.

Name of the individuals in birth, marriage and death records are alphabetically reported at the end of each registers together with the act number. Moreover, decennial registers were created with the same logic, which allow easier research of the records.
It should be mentioned that another very useful source of information may be found in the birth registers. The so-called "marginal annotations" are additional notes hand written in front

[3] They were excluded from the totality of newborns in Villagrande only those babies (101 out of 2,690) that clearly were accidentally born in the village, as their parents were present at the moment for working reasons. In the birth act in fact was clearly reported the origins of the both parents and their occupation (miners, workers involved in the construction of road and train networks, police officers).

of a birth record and reporting marriage and death of the individual. Unfortunately the procedure of marginal annotations was not systematically performed but still provides – when present - additional information on individual life course.
To integrate the information collected from the civil status registers with those concerning migrants or individuals that died outside the administrative territory of Villagrande Strisaili the population registers were used. These registers - called *anagrafe* – are available in Villagrande starting from August 1930 and represent a continuous registration of the resident population in the municipality. Their use appeared particularly useful in tracing death occurred outside the municipality as any death is communicated to the municipality of residence. Until 1980s the registers were organized according to two criteria. The first type of registers consisted of individual paper sheet, alphabetically ordered. While a second type of register was structured according to the address and composition of single households. However, since the beginning of the 80s the information is organized into modern computer databases, which made the system more efficient. The coverage of the date of death known reached using these two sources was satisfactory, however not complete.
It was decided to use further the *parish registers* as the presence of the Catholic Church is very strong in Sardinia. The availability of the data from parish registers goes really back in the past. The municipality of Villagrande Strisaili belonged to the Diocese of Ogliastra and data is available since 1633 (Anatra and Puggioni, 1983). These registers consist of 5 volumes, each one of which refers to Baptism, Marriage, Burial, Confirmation, *Stato delle Anime*[4]. In the past, these registers were written in Spanish, in Latin and only starting from 1810 priests started to use Italian language. The priest has still today the duty to fill the registers and they are kept partially in the Bishop of Lanusei (from 1633 to 1936) and in the parish of Villagrande Strisaili (from 1937 to 2006). However, for privacy concerns it is not possible to access to data after 1936. Each act reports the following information:

> BAPTISM RECORD: baptism date; name; surname; affiliation; birth date; paternity; maternity; parental origins, name and surname of godfather/mother and their origins.
>
> MARRIAGE RECORD: marriage date; spouses' name and surname; marital status; origins; paternity and maternity; impediments, witnesses: name and surname; origins.
>
> BURIAL RECORD: burial date; name; surname; date of death (generally); age; paternity and maternity; marital status; origins; cause of death (always for violent, improvise and accidental deaths), place of burial, child/adult marginal annotation (not always); testament.
>
> CONFIRMATION LIST: confirmation date; name; surname; name and surname of parents; parents' origins; name and surname of godfather/mother; godfather/mother origins.
>
> *STATO delle ANIME*: it consisted of a list of parish members grouped by household done by the priest on annual base when visiting the families during Easter. It is normally used in comparison to population registers, since it gives an indication of the population present each year in the municipality and of the composition of the household.

In addition to conventional sources for demographic research at individual level described above, data gathered from the *military registers* was also collected. In this manner deaths occurred during the world conflicts were traced. Moreover, the information collected enable to take into analysis the anthropometrics measures registered among male newborns at adulthood (20 years old). In Italy, military service has been universal and compulsory for all men born from 1861 until 1985. With regard to Villagrande Strisaili, the military registers are partially presents at the barrack "Carlo Elderle" located in Cagliari (since 1880) and partially at the State Archive of Cagliari (from 1857 to 1879). They are organized by volumes, by

[4] The literally translation of "*stato delle anime*" is *soul status*.

military districts and by '*circondario*' (county)[5]. For all individuals were reported the following variables:

> MILITARY SHEET: number of matriculate, name, surname, father name, mother name, date of birth, height, thoracic perimeter, state of teeth, occupation, date of the medical examination, result of the medical checks and possible cause of postponement or rejection, activities during the military service, date of expiration from military duties[6].

As underlined previously, in building up VILD our major aim was to cover the largest proportion of known survival. In exceptional cases, where for some individuals no additional information either than their birth acts were found, it was decided to interview the oldest members of the municipality. Additional information was recuperated through the still existing links in the families in the village. People were asked about the death of their relatives as they probably participated to the burial ceremony.
From what described in this section, it should pointed out that data collection was developed in different steps mainly to meet the need for high level coverage of known survival. In an initial phase the collection of the data was conducted systematically, later it was driven by the results and difficulties faced in the linkage procedure.

### 3.2. Data Quality

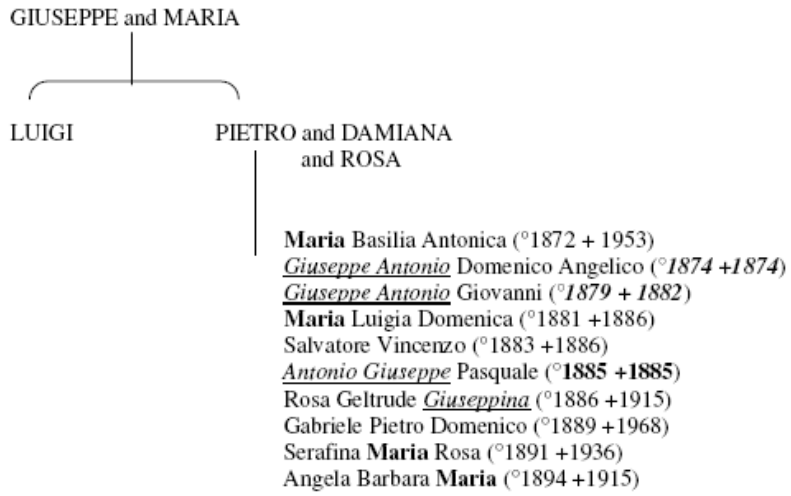
With concerns to the quality of the data sources used, there are several general points to be discussed. In general, the level of completeness and coverage of the data is high. The use of different data sources enable to overcome the problems related to the lack of information, communication and updating procedure that the municipalities' demographic system had in the past in the reporting procedure of death occurred outside their administrative territory.
The information reported is generally precise with exception of few cases of age misreporting in the death and marriage records. This confirms the well know attraction on round numbers (Livi Bacci, 1999). It may generate errors in estimating the exact age at death of individuals especially when the exact date of birth is unknown. However, this regarded exclusively information on parents and grandparents of the VILD newborns.
Most of the difficulties faced in the family reconstruction emerged in the linkage procedure and were related to certain local uses in giving names to their children. Among these uses, we find the use to give to the children the grandparents' name. This could not be itself a really a problem. However, often the same inherited name is given to more than one child in the same family, occupying different order (e.g. the same name is given as first name, second name etc). Generally, the newborn was given more than one name. Lately in their life, the same individual would keep in most of the cases only one of these names, generally, the first or the last name. This aspect complicates the linkage procedure when linking more events to the same individual. An example is given in figure 2.
The names of grandparents – *Giuseppe and Maria* – are given to 8 of the 10 children born to one of their son: *Pietro*. Moreover, the early death of the first *Giuseppe Antonio* born in 1874 implied that the same name is given to following children, turning finally to the feminine *Giuseppina* in 1886.

---

[5] Up to the Second World War Sardinia was split into 2 big military districts, namely the district of Cagliari and the district of Sassari. The first included the province of Cagliari, the province of Oristano and part of the province of Nuoro. The district of Sassari included instead the province of Sassari and the remaining northern part of the province of Nuoro. Villagrande Strisaili was part of county of Lanusei and of the district of Cagliari.

[6] In some formats were also reported the colour and the shape of hairs, colour of eyes, shape of nose and of the chin. In some cases, they annotated also the date of death, as it occurred during some military activities.

**Figure 2 – Example of names transmission in a family from Villagrande Strisaili**




Some problematic linkages are also generated by the use to give the name of a premature dead child to the next newborn, generating sometimes incorrect attribution of date of birth between the two children later in life. One good example of this type of errors is given by the case of one supercentenarian of Villagrande Strisaili who died at 110 in 1985 while she was only in reality 107 years old: Damiana Sette[7]. However this is the only case of age invalidation we have found among the oldest olds included in the VILD (Poulain et al., 2006).

Another peculiarity that is worth to point out is that some individuals appeared in the registers bearing two surnames. The first refers the father side and the second to the mother side. This double-surname attribution is the result of some remaining rule deriving from the Spanish domination. With regard to surnames, it appears necessary a procedure of standardisation of surnames, since they show some changes across time. For example the surname *Mugiri* was found also as *Mugeri* and *Muggiri*. The standardization procedure will take into consideration the actual version of surnames.

And finally, it was found in the village a limited number of surnames as consequence of a persistent closed marriage market that has lead across time to numerous cases of homonymy which may complicate the linkage procedure. However, this high endogamy implies also that since few individuals would exit from observation because of out migration as consequence of marriage.

---

[7] The meticulous reconstruction of the family composition of Damiana Sette, based on the civil status registers and the *anagrafe* allowed to discover that the person who died in 1985 was not Maria Angelica Damiana Sette, born 8 August 1874, but her younger sister called Maria Monserrata Damiana Sette who was born on 5 May 1877. The "exchange" between the two sisters occurred when the family forms were established around 1930. In fact the birth date of her older sister who died at a young age was erroneously attributed to Damiana Sette. As she never married there was no possibility to find the mistake with other age checkings.

**Family composition of Damiana Sette**

**Parents:** Sette Pietro (◦1835 + 2.10.1905)**;** Pirroni Monserrata (◦1844 + 21.11.1913) **-** Married in Villagrande on 29 November 1866

**Brothers and sisters:** Sette, Maria Barbara Lucia (◦7.5.1867 + 14.8.1945)**;** Sette, Maria Agostina (◦4.2.1870 + 5.4.1898)**;** Sette, Maria Luigia Vittoria (◦7.4.1872 + 16.7.1922)**;** Sette, Maria Angelica Damiana (◦8.8.1874 + 10.6.1876)**;** Sette, Maria Monserrata Damiana (◦5.5.1877 + 25.2.1985)**;** Sette, Domenico Antonio Daniele (◦2.5.1880 + 8.12.1968)**;** Sette, Serafino Giovanni Francesco (◦18.2.1883 + 19.8.1946)**;** Sette, Tomaso Salvatore Angelo (◦12.3.1886 + 17.8.1981)**.** *Data source:* Poulain et al., 2006 pg. 13

### 3.3 Individual linkage procedure

The data collected from the civil state registers has been the starting point of the data collection phase. All births occurred in Villagrande Strisaili from January 1st, 1866 to December 31st, 1915 constituted the initial body of the VILD. The data was collected using the genealogical software Heredis 8.0. In this software for genealogical data all individuals are framed in their family of origin and all family links are fixed. Information such as date and place of birth, date and place of marriage, date and place of death might also implement into the system. The most important advantage of genealogical software is that all information can be checked and updated at any time when finding additional information. Data improvements are permanent and cumulative. Another advantage of this type of software lays in the possibility to export back the information exported into electronic sheet, suitable for ex-post statistical analysis including reconstructed family sheets.

Secondly, each newborn (Ego) was linked to his/her exact date of death. Case of homonymy and age misreporting were handle using as reference the name of the parents and in some cases whenever available also the name of grandparents. Information on the exact death was collected for Ego's parents and siblings to complete the familial history. Data on deaths covered all civil status register from 1866 till today (September, 2006). As explained in the previous section, the coverage was maximised using population registers, parish registers, military lists and oral reported information.

Consistency in family formation was checked through marriage acts. In addition, the reported age of the spouses at marriage helped for identification and linkage procedures.

### 3.4 Coverage and censoring

The VILD include information on the 2589 newborns identified between 1866 and 1915 (table 1) and information on duration of survival has been found for 94.7% of the population[8]. For an additional 2.6 % it was possible to identify a precise censoring time. Finally only for 22 females (1.7%) and 45 males (3.3%) no additional information on their survival was found[9].

Among those individuals for whom the complete information was available, it is possible to distinguish two categories:

a) Individuals who died in Villagrande Strisaili itself;
b) Individuals who died outside Villagrande Strisaili;

All individuals with no precise information on their survival were treated differently. A main distinction was done between those for whom some information was known (which implied censoring) and those for whom no information on death was found. These uncompleted cases were classified into three groups:

Group 1 – Individuals who are still alive at September, 2006.
Group 2 – Individuals for whom at least another recorded event is known as out-migration, marriage, birth of a child, medical checks for the military service (only for males) or confirmation;
Group 3 – Individuals for whom no other event is known except their birth in Villagrande Strisaili.

[8] The percentage considers individuals with known survival and still living newborns.
[9] The figures for unknown survival are calculated separately on the total of females and of males.

**Table 1 – Data coverage on survival for individuals born in Villagrande Strisaili by 5-year cohort, 1866-1915**

| FEMALES | 1866-1875 | 1876-1885 | 1886-1895 | 1896-1905 | 1906-1915 | Total |
|---|---|---|---|---|---|---|
| Known Survival | 190 | 225 | 244 | 265 | 268 | ***1192*** |
| Alive | - | - | - | 2 | 23 | ***25*** |
| Censored | 3 | 4 | 1 | 3 | 6 | ***17*** |
| Unknown | 3 | 4 | 5 | 5 | 5 | ***22*** |
| **Total** | ***196*** | ***233*** | ***250*** | ***275*** | ***302*** | ***1256*** |

| MALES | 1866-1875 | 1876-1885 | 1886-1895 | 1896-1905 | 1906-1915 | Total |
|---|---|---|---|---|---|---|
| Known Survival | 184 | 184 | 249 | 305 | 296 | ***1218*** |
| Alive | - | - | - | - | 18 | ***18*** |
| Censored | 15 | 13 | 7 | 9 | 8 | ***52*** |
| Unknown | 9 | 7 | 9 | 14 | 6 | ***45*** |
| **Total** | ***208*** | ***204*** | ***265*** | ***328*** | ***328*** | ***1333*** |

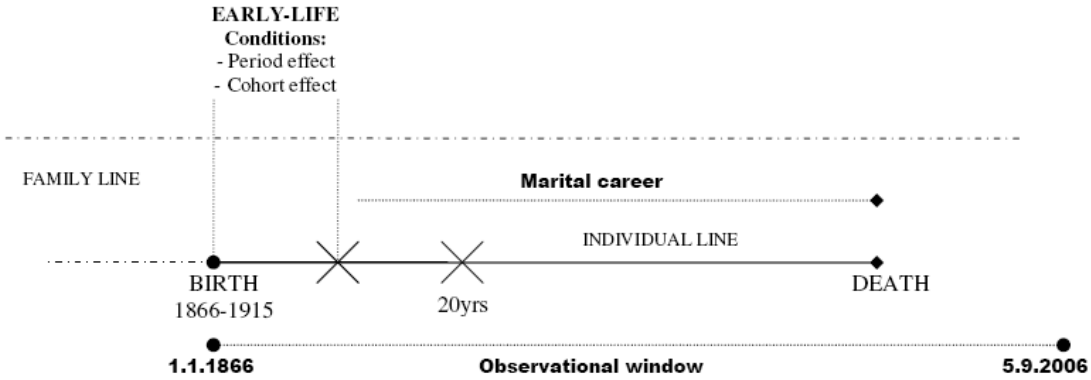
The individuals in the first two groups were left censored as no complete information on their survival was available. For group 1 the censoring time is the date of last data collection section (September 5, 2006) as it is known that there are alive at least at that time.
For the second group different assumptions for more precise censoring time were formulated according to the available information. We have further identified five subgroups, which are the following:
- '*simple out-migrants*' for who is known the exact timing of out-migration from Villagrande Strisaili. Their migration occurred when they were adult and was regularly registered in the population registers. The censoring time is the date of migration.
- '*marriage out-migrants*'. To this subgroup belong those individuals for which the last information available was the occurrence of marriage with a no-local spouse. Moreover, no births occurred in Villagrande Strisaili in the former family. These two conditions let us assumed that the couple out-migrated as soon after marriage maybe in the village of the no-local spouse. The censoring time considered is the date of marriage.
- '*last born child*', this category includes all individuals for which date of marriage is know, he/she had one or more children in Villagrande, but still the exact date of death is missing. The censoring time is given by the date of birth of the last child.
- '*military visit*'. The date of military visit is considered for males that did not registered other events in the civil status registers or in the population registers. As they were visited for the military service, we know that they were alive at least at the time of the medical checks. The censoring time is the date of military visit.
- '*confirmation date*'. When ever it was not possible to trace a marriage, a former family and/or the date of military visit, the parish registers of confirmation were checked. The date of confirmation was considered the censoring time.
Finally for a total of 67 individuals (group 3), no other event is known excepted from their date of birth. They are classified as unknown. However their proportion is really limited (2.58%) that the results from the analysis should not be affected by this lack of information.

## 4. Potential use of VILD for longevity studies

The family reconstruction methodology provides a powerful database for analysis as it enables to work both at aggregate and individual level.

At aggregate level, it is possible to describe the structure and the composition of the population identifying the major demographic trends across time. Conclusions can be drawn on the level of fertility, nuptiality, mortality and mobility.
However, at individual level the investigation assumes more interesting features as the effect of a set of variables on individual survival can be estimated applying *life history analysis* methodologies (Allison, 1984; Yamaguchi, 1991; Blossfeld and Rohwer, 2002). It may consist of a multi-scale analysis, whereas individual, family and contextual variables are included in the models. At a contextual level, it is possible to take into consideration for example the disease load considering the effect of the level of infant and child mortality on average individual lifespan. In addition, the family structure and composition may contribute to describe and investigate the possible influence of variables such as family size, parity order, child spacing, age of the parents at birth, marriage or celibate, parents and siblings' survival and so on. The effects of all the mentioned variable may be assessed considering their interaction with the individuals characteristics such as sex, month of birth, month of death, height and thoracic perimeter at age 20 (the latter variables only for males).
A synthetic idea of the structure of the data into a life course perspective is given in figure 3.

**Figure 3 – A life course approach: Context, Familial line and Individual line**



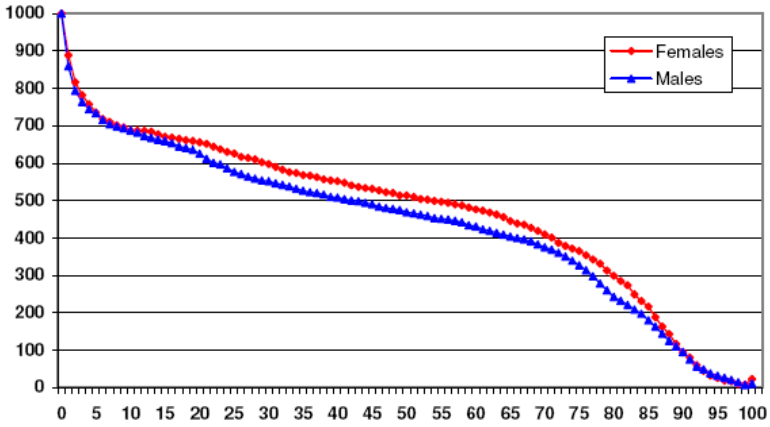

The list of possible variables for analysis appears to be important and relevant indications can be derived from the existing literature. Their organization into the conceptual model should attempt to integrate the contribution of different research groups and to compare previous studies on human survival.
Finally, it can be added that the present study may be replicable in another setting with a partially or totally different context. This could represent in future a successful way for the investigation differential longevity in different settings.

## 5. Some results

### 5.1. A general overview on the evolution of mortality

A first investigation of the exceptionality of Villagrande Strisaili was conducted looking at the structure of mortality level at different ages, starting from infant and child mortality and going to later survival.
Following Villagrande newborns survival across time (figure 4), it was estimated that 78% of them reached 5 years old and 65% lived at 18 years old, just before military obligations for males and childbearing for females.

**Figure 4 - Survival curves for females and males for VILD newborns (1866-1915)**




The village registered in the first World conflict 40 losses. In addition, for deaths occurring before 50 years old, certainly we should consider the role of maternal mortality and accidental deaths, especially for males. However, the study of adult mortality (from 18 to 50 years) appears complex to analyze. However, a further investigation through the death acts could give some more insight for example on the causes of deaths helping to estimate specific disease patterns. After 50 years old, survivors are 50% of males and 55% of females. Starting from that age – from 50 to 75 years old – males and females seem to be exposed to the same mortality risk. Only from 75 to 85 years men record higher risk than females, while at oldest old ages (after 85 years) differences between males and females tend to disappear with men recording reduced mortality risk. And at the end, 10% of the newborns – both males and females - are surviving at 90 years old, meaning that 1 newborn out of 10 reached 90 years old. Moreover, Villagrande Strisaili males show to live as long as women. This result is really exceptional compared to other populations born during the same period. However, there is no straight explanation for it. Despite genetics, we believed that further investigation should focus on the social role of males and the oldest old, who still cover a privileged position in the community as it was in traditional societies.

Summarizing, the analysis of their survival *tout court* does not give an answer on why males and females in Villagrande are long-living and at older ages males are living as long as women. However, it helps us to identify some significant periods of higher mortality as well as which ages show to be critical for later survival. In order to find the possible determinants of exceptional survival, similar studies should be carried out in other villages, registering lower level of longevity. Moreover, selected and relevant variables should be compared. It should be possible in this manner to control for example for some similarities in order to estimate their effect and whether or not some discriminate factors can be identified which lead to improved survival.

### 5.2. The effect of Parental age at Death and Parental age at Birth

The inheritability of longevity has been investigated already at the beginning of last century (Beeton and Pearson, 1899; 1901) and since then many authors have aimed to estimate its contribution toward longevity. However, when talking of familiar transmission of longevity, two meanings of this term should be considered. There are: a) Familiar genetic component, meaning that the favourable treats for survival are inherited and part of the shared family DNA; b) Familiar environmental component, which refers instead to the influence that

specific characteristics of the household such as shared familiar environments, daily habits, family size and so on, have on individuals' survival.
According to the existing literature, there is a positive association between parental age at death and survival of offspring, representing about a quarter of the variability of lifespan (Cournil, 2000; Mitchell et al., 2001). However, there is not common agreement among researchers. In the one hand, some studies did not propose their results as definitive, presenting sample biases. While on the other hand, in some cases it was also found a positive association also for partners. This latter result suggests that familiar environmental component could also have an effecting role on human longevity. There are different possibilities for investigation. However, generally authors agree that only later survival should be considered in the analysis as the impact of infant and child mortality and accidental deaths could bias the estimations. We decided therefore to explore the VILD data according to 3 possible combinations which are the following:

a) Offspring reaching 50 years old and over and no selection according to the age at death of their parents;
b) Offspring 50 years old and over with fathers 50 years old and over calculated separately from offspring 50 years old and over with mothers 50 and over;
c) Offspring 50 years and over with both parents 50 years and over.

From the estimates of the three groups the relation between parental age at death and age at death of their offspring is in general positive with exception of the relation between fathers and sons. It is worth noting that the mother effect on survival is always greater than that of father and among the offspring it shows to be greater for sons than for daughters; however it is not statistically significant. This prominent mother effect was found also in prior genetic studies, meaning that the results derived from VILD database go in the same direction. More importantly the identification of a mother effect calls for reflection on the role of mtDna (Herskind et al., 1996; Koperlainen, 1999), which should be further investigated.

**Table 2a – Correlation estimates for offspring 50 and over with no parental age selection**

| Age at Death | Sons | | Daughters | |
|---|---|---|---|---|
| | R2 | no. Cases | R2 | no. Cases |
| Father | -0,026 | 545 | 0,032 | 569 |
| Mother | 0,067 | 560 | 0,066 | 585 |

**Table 2b – Correlation estimates for offspring 50 and over with fathers 50 and over mothers 50 and over**

| Age at Death | Sons | | Daughters | |
|---|---|---|---|---|
| | R2 | no. Cases | R2 | no. Cases |
| Father | -0,025 | 483 | 0,045 | 501 |
| Mother | 0,074 | 483 | 0,059 | 493 |

**Table 2c –Correlation estimates for offspring 50 years and over with both parents 50 years and over**

| Age at Death | Sons | | Daughters | |
|---|---|---|---|---|
| | R2 | no. Cases | R2 | no. Cases |
| Father | -0,006 | 421 | 0,044 | 438 |
| Mother | 0,085 | 433 | 0,054 | 449 |

With regard instead to the effect of parental age at birth, it emerges that increases in parental age at reproduction showed to have a decreasing effect on offspring lifespan (Kemkes-

Grottenthaler, 2001). Both maternal and paternal ages at conception were studied and different effects emerge on sons and daughters (Gavrilov and Gavrilova, 2000).
The effect of this variable was tested among all Villagrande Strisaili newborns surviving 50 years or longer, whose parents reached 50 years old and over (table 3). The mean age of fathers at the birth of their child is 42.1 years old, while mother on average have their child at 32.1 years old. Differently from the expectations deriving from the literature, increases in father age at birth leads to increases in the mean lifespan duration for both sexes. This increase is almost linear for daughters. With regard to the effect of maternal age, no significant results were found.

**Table 3 -Average age at death of offspring (50 years and over) of father and mother (50 years or over) and age at birth**

| PATERNAL AGE | SONS | | | DAUGHTERS | | |
|---|---|---|---|---|---|---|
| | Mean age at death | Number | Stand. Error | Mean age at death | Number | Stand. Error |
| Less than 30 | 77,3 | 34 | 2,0 | 78,6 | 34 | 2,3 |
| 30-35 | 78,8 | 76 | 1,3 | 79,7 | 91 | 1,3 |
| 35-40 | 80,2 | 107 | 1,2 | 79,7 | 103 | 1,2 |
| 40-45 | 80,3 | 105 | 1,2 | 80,4 | 113 | 1,1 |
| 45 and over | 79,5 | 178 | 0,9 | 80,6 | 184 | 0,8 |

| MATERNAL AGE | SONS | | | DAUGHTERS | | |
|---|---|---|---|---|---|---|
| | Mean age at death | Number | Stand. Error | Mean age at death | Number | Stand. Error |
| Less than 25 | 80,3 | 86 | 1,3 | 81,5 | 91 | 1,2 |
| 25-30 | 81,6 | 107 | 1,1 | 79,1 | 101 | 1,3 |
| 30-35 | 79,7 | 124 | 1,0 | 81,8 | 124 | 1,0 |
| 35-40 | 78,4 | 102 | 1,2 | 79,9 | 117 | 1,1 |
| 40 and over | 79,2 | 58 | 1,6 | 78,9 | 64 | 1,4 |

## Conclusions

The present contribution supports the choice of family reconstructed database for the study of human longevity. In this sense, the VILD represents an attractive and fruitful application of the family reconstruction in a population experiencing exceptional longevity in Sardinia (Italy). Historical data is used to study a today population establishing a unique bridge between historical demography and a current level of longevity. Each individual is followed from birth to death, going through marriage and family formation.
The use of this type of database in the study of longevity has several advantages. It enable to control for a certain number of confounding variables such as environmental and cultural factors as all individuals share the same location. By considering family links, the database considers not only individual variables such as sex, birth cohort, month of birth, height, but enables to include in the analysis familial characteristic, testing both the familial genetic and familial environmental components of longevity. Among the possible variables there are: parental age at birth, parental age at death, siblings' age at death, family size, birth intervals etc.
The list appears really important. From such a database we expect to answer to several recurrent research questions in the study of longevity such as: Why in the BZ men live as long as women? What is the contribution of the familial component on human survival? Do contextual variable matter on longer lifespan? The high coverage reached guarantees from selection biases since not only long survivors are included. The VILD is really representative of the entire population of the selected community. This finally permits the comparison of the explanatory analysis results with numerous previous studies in the field of longevity.

## REFERENCES


Allison P.D. (1984), *Event history analysis: regression for longitudinal event data*, (Quantitative applications in the social sciences, no. 46), Sage, Newbury Park, California.

Anatra B. and Puggioni, G. (1983), *Fonti ecclesiastiche per lo studio della popolazione della Sardegna centro-meridionale. Inventario dei registri parrocchiali di sette Diocesi della Sardegna Centro-Meridionale*, CISP, Roma, Italy.

Beauchamp P., Charbonneau, H., Desjardins, B., Légaré, J. (1977), "La reconstruction automatique des familles: un fait acquis", *Population*, Numero special : La mesure des phénomènes démographiques: Hommage a Louis Henry, pp. 375-399.

Beeton M. and Pearson, K. (1899), Data for the problem of evolution in man. II. A first study on the inheritance of longevity and the selective death rate in man, in *Proceeding of the Royal Society*, vol. 65, pp. 290-305.

Beeton M. and Pearson, K., (1901), On the inheritance of the duration of life, and on the intensity of natural selection in man, in *Biometrika*, n° 1, pp. 50-89.

Blossfeld H.P and Rohwer, G. (2002), *Techniques for event history modelling: New approaches to casual analysis*, Lawrence Erlbaum Associates, Mahwah, New Jersey.

Bouchard G., Roy R., Casgrain B., Hubert. (1989). Fichier de population et structures de gestion de base de données: le fichier-réseau BALSAC et le système INGRES/INGRID », *Histoire & Mesure*, Vol. IV, no. 1/2, pp. 39-57.

Cannas A. and Rubiu A. (1977), *Villagrande Strisaili: tra storia e leggenda*, Fossataro, Cagliari, Italy.

Caselli G. and Lipsi, R.M. (2006), "Survival differences among the oldest old in Sardinia: who, what, where, and why?" *Demographic Research*, vol. 14, n° 13, pp. 267-294.

Corridore F. (1902), *Storia documentata della popolazione di Sardegna (1479-1901)*, Anarldo Forni Editore 1990 (ristampa dell'edizione di Torino 1902), Italy.

Cournil A. (2000), *Contribution à l'étude de l'hérédite de la longévite humaine. Analyse des ressemblances familiales dans la vallee de la valserine (jura francais) du XVIIIè siecle a nos jours*, PhD, Biométrie-Biologie évolutive UMR CNRS 5558 Centre d'études démographiques Institut des Sciences, Université Claude Bernard - Lyon 1, 147 p.

Deiana L., Ferrucci L., Pes G.M., Carru C., Delitala G., Ganau A., Mariotti S., Nieddu A., Pettinato S., Putzu P., Franceschi C., Baggio G. (1999), "AKEntAnnos. The Sardinia study of extreme longevity", *Ageing Clinical and Experimental Research*, vol. 11, pp. 77-84.

Del Panta L. and Rettaroli, R. (1994), *Introduzione alla demografia storica*, Editori Laterza, (ML), Roma, Italy.

Fleury M. and Henry L. (1956), *Des registres paroissiens à l'histoire de la population,* INED, Paris, France.

Gatti A.M. and Puggioni G. (1998), “Storia della popolazione dal 1847 a oggi” in: *Storia d’Italia, Le regioni, dall’Unità ad oggi*, Berlinguer L. and Mattone A. (eds), vol. *La Sardegna,* pp. 1039-1079, Einaudi, Torino, Italy.

Gavrilov L.A. And Gavrilova N.S. (2000), Human longevity and Parental age at Conception, in Robine J.-M., Kirkwood, T.B.L., Allard, M. (ed.) *Sex and Longevity: Sexuality, Gender, Reproduction, Parenthood*, Berlin, Springer-Verlag, pp. 7-31.

Gavrilova N.S., Gavrilov, L.A. (1999), "Data resources for biodemographic studies on family clustering of human longevity", *Demographic Research*, vol. 1, n° 4, pp. 1-48.

Henry L. and Blum, A. (1988), *Techniques d'analyse en démographie historique*, INED, Paris, France.

Herskind A.M., Mcgue, M., Holm, N.V., Sorensen, T.I., Harvald, B., Vaupel, J.W. (1996), "The heritability of human longevity, a population-based study of 2872 Danish twins pairs born 1870-1900", in *Human Genetics*, n° 97, pp. 319-323.

Kemkes-Grottenthaler A. (2001), Parental effects on offspring longevity--evidence from 17th to 19th century reproductive histories, in *Annals of Human Biology*, vol. 31, n° 2, pp. 139-158.

Korpelainen H. (1999), "Genetic maternal effects on human life span through the inheritance of mitochondrial DNA", in *Human Heredity*, vol. 49, pp. 183-185.

Livi Bacci M. (1999), *Introduzione alla demografia*, Loescher Editore, Torino, Italy

Mitchell B.D., Hsueh W-C., King T.M., Pollin T.I., Sorkin J., Agarwala R., Schäffer A.A., Shuldiner A.R. (2001), Heritability of life span in the Old Order Amish, in *American Journal of Medical Genetics*, vol. 102, n° 4, pp. 346-352.

Poulain M., Pes G.M., Carru C., Baggio G., Franceschi C., Ferrucci L., Deiana L. (2006) “The Validation of Exceptional Male Longevity in Sardinia”, in *Human Longevity, Individual Life Duration and the Growth of the Oldest-Old Population*, ROBINE J.M. et al. (eds.) Chapter 7, Springler and Kluwer .

Poulain M., Pes G.M., Grasland C., Carru C., Ferrucci L., Baggio G., Franceschi C., Deiana L. (2004), "Identification of a geographic area characterized by extreme longevity in the Sardinia island: the AKEA study", *Experimental Gerontology*, vol. XX, pp. 1-7.

Salaris L. (2005), "Geography and longevity: a demographic study on Sardinian centenarians". Population Research Centre, *Master Thesis Series* 05-3, University of Groningen, The Netherlands.

Yamaguchi K. (1991), *Event history analysis*, Sage, Newbury Park, California.