Longevity pattern in the Italian region of Emilia Romagna: a dynamic perspective

Giulia Roli[1], Alessandra Samoggia[1], Rossella Miglio[1], Rosella Rettaroli[1]

[1]*Department of Statistical Sciences, University of Bologna, 40126 Bologna, Italy*

Abstract. The pattern of longevity in the Italian north-eastern region of Emilia Romagna was investigated at the municipality level, considering a modified version of the centenarian rate (CR) in two different periods (1995-1999 and 2005-2009). Due to the rareness of such events in small areas, spatio-temporal modelling was used to tackle the random variations in the occurrence of long-lived individuals. This approach allowed us to exploit the spatial proximity to smooth the observed data, as well as controlling for the effects of a set of covariates. As a result, clusters of areas characterised by extreme indexes of longevity could be identified and the temporal evolution of the phenomenon depicted. A persistence of areas of lower and higher occurrences of long-lived subjects was observed across time. In particular, mean and median values higher than the regional ones, showed up in areas belonging to the provinces of Ravenna and Forlì-Cesena, on one side spreading out along the Adriatic coast and, on the other stretching into the Apennine municipalities of Bologna and Modena. Further, a longitudinal perspective was added by carrying out a spatial analysis including the territorial patterns of past mortality. We evaluated the effects of the structure of mortality on the cohort of long-lived subjects in the second period. The major causes of death were considered in order to deepen the analysis of the observed geographical differences. The circulatory diseases seem to mostly affect the presence of long-lived individuals and a prominent effect of altitude and population density also emerges.

Keywords: centenarian rate, hierarchical Bayesian models, longevity, small areas, spatial and spatio-temporal models, Italy.

Introduction

In the last decades, the study of human longevity and its development has caught the attention of researchers from different disciplines. Several studies performed in different Italian regions have shown the presence of specific areas, where the prevalence of the oldest-old[1] people is higher than elsewhere. For instance, a definite, geographical area in Sardinia is characterised by an exceptional male longevity (Poulain et al., 2004) as well as a low female/male centenarian ratio. Further, a significant negative correlation between surname abundance and the longevity index has been detected in Calabria (southern Italy) where some isolated areas of male longevity present a high level of inbreeding (Montesanto et al., 2008). Other longevity areas have been found in other southern regions such as Sicily and Campania (Lipsi, 2009). Finally, in Emilia Romagna some longevity "clusters" have been identified and their persistence has been detected by comparing the results of different spatial scan statistic methods (Miglio et al., 2009).

The explanatory analysis of disparities in the frequencies of the oldest-old population reminds one of discussions mentioning genetics as well as environmental factors such as the topography of certain residential areas. Indeed, several studies in some Italian regions have found that longevity often correspond either to mountainous areas (Poulain et al., 2004) or to smaller elevations and hilly regions (Lipsi, 2009). Moreover, it is well known that the increase in the number of oldest-old is mainly explained by the decline of mortality after the age of 80 (Thatcher, 1992, 2001; Vaupel and Jeune, 1995; Robine and Paccaud, 2005, Caselli and Lipsi, 2006).

One of the scientific approaches aims at mapping the geographical distribution of extreme longevity using spatial analysis techniques, in order to identify areas, or clusters of areas, characterised by particularly high, or low, concentrations of oldest-old population. When a fine-scale geographical analysis is performed and the phenomenon under study is characterised by a low number of units, the territorial distribution of cases may be strongly influenced by random variation due to the sample data. This invalidates both the distribution of the cases at a given time point and its dynamic evolution in terms of comparison between successive time points. A recent solution adopted to

Corresponding author:
Giulia Roli
Department of Statistical Sciences
University of Bologna
via Belle Arti 41, 40126 Bologna, Italy
Tel. +39 051 209 8408; Fax +39 051 232 153
E-mail: g.roli@unibo.it

---

[1] The term " oldest-old" generally refers to persons at the upper segment of the age pyramid.

deal with these problems consists in the use of spatial and spatio-temporal models which involve geographical and time interactions resulting in a smoothing effect among the observations. These methods can be developed either within a frequentist framework (see, for example, Breslow and Clayton, 1993; Langford et al., 1999) or under a fully Bayesian perspective (see, for example, Bernardinelli and Montomoli, 1992; Mollie, 1994; Waller et al., 1997; Knorr-Held, 2000; Lawson, 2009). The development of computational methods, such as the Markov chain Monte Carlo (MCMC) techniques, together with Gibbs sampling or the Metropolis-Hastings algorithm (Carlin and Louis, 1998; Gelman et al., 2003), resulted in a great increase in the use of Bayesian methodology. The application of this approach is also common for the analysis of the geographical distribution of health data as it is feasible for almost any model involving multiple levels, random effects and complex dependence structures. Moreover, this kind of modelling allows embedding different levels of covariates aiming to explain the variability among areas.

Here, we used a Bayesian spatio-temporal model to manage both the geographical structure and the temporal dimension of extreme longevity over the last 15 years in Emilia Romagna. In particular, we considered a modified version of centenarian rate (CR) (Robine and Caselli, 2005) centred on two different five-year periods (1995-1999 and 2005-2009), separately by municipality and gender. The model we employed allowed us to control for variations due to random occurrences, especially in small municipalities which have low demographic dimensions. The main purposes were to depict the temporal evolution and the spatial structure of longevity in this region, identify territorial groups of areas characterised by high or low levels of longevity and investigate the association of the detected geographical distribution of longevity with information on altitude and population density.

A further objective was to investigate how past mortality patterns influence the dimension and the spatial distribution of current proportions of long-lived people. We attempted to evaluate the effects on longevity of mortality structures measured at different ages. Therefore, we considered the levels of mortality (by cause of death) of the last 15 years, which might have affected the survival of the observed cohort of long-lived individuals at younger ages. In particular, we aim at explaining how the current distribution of longevity is related to the past mortality of the same cohort at the municipality level and if there are substantial gender differences. Within this kind of "longitudinal" framework, the longevity outcome is modelled by a Bayesian spatial regression where the area-specific levels of mortality for each group of age in the cohort are included as regressors.

Throughout the analysis, we adopted a gender-specific perspective in order to detect and explain the paths leading to a different spatial distribution of the male and female elder population in the region.

Materials and methods

Area, data and indicators

The north-eastern Italian region of Emilia-Romagna (Fig. 1) shows one of the oldest age structures in Europe (with 22.5% of persons aged 65+ years, and 6.9% of persons aged 80+ years in 2009 in a total population of 4,337,966). This region is characterised

Fig. 1. Political and physical map of Emilia Romagna in the north-eastern part of Italy.

by a pronounced geographical variability: mountains, hills and a wide flat plain, with a consequent heterogeneity in environmental context, population density, social conditions and economic resources. Emilia Romagna is split into nine provinces (Piacenza, Parma, Reggio Emilia, Modena, Bologna, Ferrara, Ravenna, Forlì-Cesena and Rimini) that contain 341 municipalities.

In order to measure the different spread of longevity in this region, we used a modified version of the CR, as first proposed by Robine and Caselli (2005) and researched by Robine and Paccaud (2005) and Robine et al. (2006)[2]. To avoid inconsistency or lack of data, we considered people aged 95 years and over, denoted P95+, separately by municipality and gender. The proposed indicator of longevity, denoted CR95+, is obtained by dividing this count by the number of 55-64 years old persons (P^{55-64}) who lived in the same area 40 years earlier, according to the following formula:

$$CR^{95+}_{kit} = \frac{P^{95+}_{kit}}{P^{55-64}_{ki(t-40)}} \quad (1)$$

for each time t, gender k and municipality i. In this framework, we assumed that the living individuals in an area 40 years before the time point t under study are those exposed to the "risk" of becoming P^{95+} or, similarly, that P^{95+} observed at time t comes from the cohort of individuals resident in the same area 40 years earlier. In practice, we drew the denominator of CR^{95+} from the nearest available Italian census, with analogous considerations to identify the cohort age group. To depict the temporal dimension of the phenomenon, we calculated the CRs^{95+} centred on two periods: 1995-1999 and 2005-2009[3]. Thus, the corresponding denominators refer to the same cohorts of individuals at the censuses in 1961 and 1971, respectively[4].

In the whole region, we observe an average number of P^{95+} which amounts to 4,700 (0.12% of total population) in the period 1995-1999 and to 10,639 (0.25%) in 2005-2009. Across the municipalities the CR^{95+} index ranges from a minimum of 0 to a maximum of 56.6 (per 1,000 individuals exposed to "risk") in the first period and from 0 to 100 in the second one (see Table 2 for further details).

Additional information on the features of the municipalities was also collected. In detail, we considered the classification of areas with respect to the altitude (mountain, hill, coastal hill and plain) and population density. In particular, 67 municipalities of Emilia Romagna (~20%) are classified as mountainous areas,

99 (~29%) as hills, 10 (~3%) as coastal hills and 165 (~48%) are plains. The population density varies from a minimum of 4.18 individuals per km² in the municipality of Zerba (mountainous zone) to a maximum of 2791.96 of Cesenatico (plains). Altitude and population density are taken into account as proxies of economic, socio-cultural and environmental features. The plains and the areas at the foot of the hill are characterised by industrialization, resulting in crowded and frequently high polluted zones, which include all the biggest towns of the region, but there are also rural areas with a lower density of population. Conversely, high hill and mountainous zones are less crowded and marked by the prevalence of agricultural, forest and handicraft activities and by an ageing of population due to the intense process of depopulation of the occurring over the past several decades.

Altitude and population density were combined to form the following groups[5]: (i) mountainous areas; (ii) hilly areas with a population density below 78 people per km²; (iii) hilly areas with a population density between 78 and 193 per km²; (iv) hilly areas with a population density above 193 km²; (v) coastal hills, where there were no remarkable differences in population density; (vi) plains with a population density below 78 km²; (vii) plains with a population density between 78 and 193 km², and; viii) plains with a population density above 193 km² (Fig. 2).

For the cohort of P^{95+} in the period 2005-2009, we focused on the corresponding patterns of mortality measured when they were aged 80-89 years and 90-99 years, separately. The main aim was to strengthen the

---

[2] The CR has been shown to be an appropriate indicator of longevity as it takes into account the effects of the work-related migration. Indeed, it is well known that in Italy, including the Emilia Romagna region, migration was very common in the recent past, especially for working ages and people who now belong to the oldest-old population. Under this perspective, the CR indicator measures and compares the dimensions of longevity in different areas as the ratio of the observed number of centenarians to the number of the cohort survivors 40 years before the current observation and not to the number of births in the corresponding generation as is usually done. As a result, the CR removes the unknown influence of the migration process, which is assumed to be negligible only after the 60 years of age. Moreover, it is independent from the size of birth cohorts, infant mortality, past migrations, and policies of naturalization.

[3] A multiple-year aggregation of data was introduced to avoid random fluctuations due to specific years or cohorts.

[4] Since these censuses are referred to the second half of October 1961 and 1971, we conventionally consider the data as a proxy of population on 1st January 1962 and 1972, respectively. Therefore, for each period we have $CR^{95+}_{ki} = \frac{P^{95+}_{ki}}{P^{60-69}_{ki}}$.

[5] We split the ordered data for the population density into three equal-sised subsets, i.e. terciles or 3-quantiles.
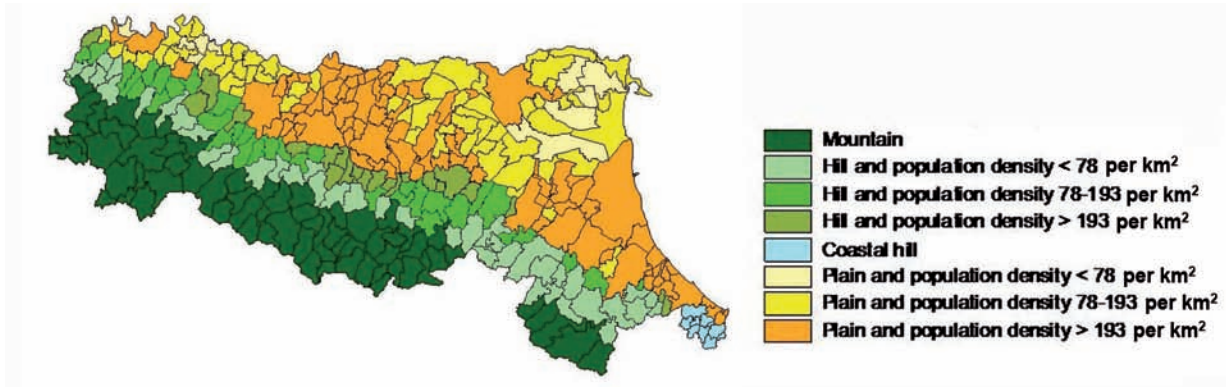
Fig. 2. Altitude and population density (number of people per km²) distribution in Emilia Romagna.

ties between the count of present long-lived population and the mortality patterns that characterised the same population in the recent past (Thatcher, 1992, 2001; Vaupel and Jeune, 1995; Robine and Paccaud, 2005; Caselli and Lipsi, 2006). Then, we turn the attention to the two main causes of death. In particular, for each municipality we consider death counts by gender and cause of death according to the two age groups of the cohort. These data are averaged over five calendar years, corresponding to 1990-1994 for the age group 80-89 years and 2000-2004 for those aged 90-99 years. Moreover, the data are grouped into two categories of causes of death according to the international classification of diseases (ICD), revision IX (ISTAT, 1997): malignant neoplasms and diseases of the circulatory system. These two disease groups jointly represent the most common causes of death for people aged 80 years and above (approximately 70% of the cases excluding injuries and poisonings). We further consider a broad class including all the causes of death, except for injuries and poisonings.

The structure of mortality depicted by the death counts can be appropriately summarised by the relative risk of death. An estimate of this measure is represented by the standardised mortality ratio (SMR), defined as the ratio between the observed O and the expected E death counts with respect to any specific territorial unit (Kahn and Sempos, 1989). The SMR allowed us to control for the influence of the age structure within the age groups, making it possible to compare different mortality levels among the municipalities. In our case, the SMRs were computed for each gender k, municipality i, age group $a=x-x + 9$ and cause of death d as follows:

$$SMR_{kiad} = \frac{O_{kiad}}{E_{kiad}} \qquad (2)$$

In particular, the expected death counts for each area, E_{kiad}, were calculated, separately by gender and

cause of death, through the average mortality rate, m_{kad}, for the overall region in the age group a and multiplied by the population in area i with ages in a (*internal standardization*).

In order to facilitate the interpretation of the results, the SMRs were grouped into three levels of mortality based on their quintile[6] distribution for each cause of death, gender and age group. The first and the last quintile subsets identify municipalities with extreme (low or high) mortality levels. Areas with values of SMR close to the overall regional rate were included into the residual group.

The sources of the data all refer to official statistics published by the Italian institutional agency for statistical data collection (ISTAT) and by the regional authorised agency of Emilia Romagna.

Spatio-temporal development

In order to draw the spatial distribution of longevity in Emilia Romagna, we investigated the space-time pattern of "risk" of becoming P^{95+} among the 341 municipalities separately by gender adopting a hierarchical Bayesian approach which exploits the adjacency and interaction of the geographical areas. We further associated a temporal dimension to the phenomenon by considering the evolution of P^{95+} in the two periods 1995-1999 and 2005-2009.

In detail, we employed a hierarchical Bayesian regression for areal data. With respect to the i-th municipality area, we assumed the observable P^{95+} at each time period t (1995-1999 and 2005-2009), denoted y_{it}, are Poisson distributed with the means p_{it} θ_{it}. In this formulation, p_{it} represents the potential P^{95+} and θ_{it} is the estimate of the CR^{95+} in the required location and period. A Poisson distribution is chosen in order to deal with the small numbers of people aged

---

[6] i.e. 5-quantiles (see note 5).

95 years and above compared to the number of 55-64 years old persons living in the same area 40 years earlier. In the model specification, we followed the conventional log-linear formulation on the rate θ_{it} and allowed for the possibility of different components that additively contribute to explain the spatio-temporal distribution of these rates. In particular, we assumed the presence of a temporal effect, as well as a spatial component, that may be specified as follows:

$$log(\theta_{it}) = \alpha_t + u_i + v_i \qquad (3)$$

where α_t represents the time-varying intercept and u_i and v_i are the correlated and uncorrelated spatial heterogeneitics, respectively, which are both assumed to be constant in time.

In a fully Bayesian setting, we specified a prior distribution for each parameter involved in (3). The intercept α_t includes unspecified features of period t without assuming a temporal structure *a priori* due to the only two time observations considered. For the correlated spatial component u_i, we assumed a Gaussian conditionally autoregressive (CAR) model (Besag, 1974; Banerjee et al., 2004) of the form:

$$p(u_1, ..., u_{341} \mid \tau_u) \propto \exp\left\{-\frac{\tau_u}{2} \sum_{i \neq j} w_{ij} (u_i - u_j)^2\right\} \quad (4)$$

where τ_u represents the precision parameter and w_{ij} are the adjacency weights across the areas. Although improper, the CAR prior leads to a posterior distribution which is proper, allowing the Bayesian inferences still to proceed.

The random effects v_i which capture the region-wide heterogeneity are supposed to follow an ordinary exchangeable normal prior:

$$v_i \overset{iid}{\sim} N\left(0, \frac{1}{\tau_u}\right) \qquad (5)$$

where τ_v is the precision term.

Then, we needed to specify (hyper-) prior distributions for the precision parameters τ_u and τ_v. In particular, we considered vague and proper distributions which are also "fair", i.e. they yield the proportion of the variability due to the spatial homogeneity to be ~1/2 *a priori*. As suggested by Best et al. (1999), we used the following Gamma priors:

$$\tau_u \sim Gamma(0.1; 0.1) \qquad (6)$$

and

$$\tau_v \sim Gamma(0.001; 0.001) \qquad (7)$$

In order to further explain the differences in the presence of P^{95+} across the areas, some covariates are included into equation (3). In particular, we controlled for some areal features and considered the eight categories of altitude and population density introduced above.

The model with covariates for the logarithm of the rate θ_{it} can be analytically expressed as:

$$log(\theta_{it}) = \alpha_t + \beta x_i + u_i + v_i \qquad (3)$$

where x_i identifies the altitude and population density group of the i-th municipality and β is the corresponding effect on the log rate.

The regression effect of the mortality patterns

In order to study how the past mortality patterns influenced the dimension and the spatial distribution of current long-lived population, an additional analysis was carried out following a "longitudinal" perspective. We attempted to investigate the effect of the past structure of mortality for people aged 80-89 years (in 1990-1994) and 90-99 years (in 2000-2004) on the same cohort of individuals who reached 95 years of age or above in the more recent period of 2005-2009.

For this purpose, a spatial model was employed separately by gender. The outcome Y now represents the count of people aged 95 years and above (P^{95+}) related only to the period 2005 to 2009. For each municipality i, the y_i were again assumed to follow a Poisson distribution of parameter $p_i \theta_i$, with analogous meanings to those described in the previous section. The log-linear model on the rate θ_i involved the spatial structured (u_i) and unstructured (v_i) components, besides controlling for the effects of altitude and population density (x_i). In addition, we included the groups of SMRs for all the causes of death (excluding injuries and poisonings) as regressors and, in a separate model both malignant neoplasms and diseases of the circulatory system that characterised the survival of the current P^{95+} when they were 90-99 and 80-89 years old. These quantities are denoted by z_{iad}, where a is the age group of the cohort and d the specific cause of death. Thus, the model can be formalised as follows:

$$log(\theta_i) = \alpha_0 + \beta x_i + \sum_{a,d} \gamma_{ad} z_{iad} + u_i + v_i \qquad (9)$$

where the estimation of parameters γ_{ad} now represents our main objective.

In addition, we allowed for a correlation among causes of death and among the repeated measure of

mortality over time by modelling the priors on the effects γ_{ad} through a non-informative, multivariate, normal distribution with an unknown population mean vector and a variance covariance matrix Σ. The inverse of Σ corresponds to a precision matrix Ω, which is assumed to follow a vague Wishart distribution (Spiegelhalter et al., 2003). For the other parameters, we assigned prior distributions analogous to those specified in the previous section.

Results

Spatio-temporal development

The application of the hierarchical, spatio-temporal model introduced above offers the opportunity to investigate different aspects of the spatial distribution of longevity pattern, as well as controlling for variations due to random occurrences. The last feature is pursued through the so-called "smoothing effect", which in practice can be highlighted by comparing the observed with the estimated values of the CRs^{95+} for each area. Figure 3 plots these numbers in the first and second period for males in the 341 municipalities of Emilia Romagna (the same findings can be obtained for the female group). The observed CRs^{95+} showed a great variability with peaks of extremely high and extremely low numbers, while the CRs^{95+} estimated by our model were closer to each other (i.e. to the region-

al average). Note that all the results we report for the CRs^{95+} refer to a number of 1,000 individuals who can reach the age of 95 years and above.

The smoothing effect, which finally reflects the efficiency of such a model, is also evident when comparing the maps of the observed and estimated values of the CRs^{95+} (Figs. 5-8). For both genders, the model yields more homogeneous estimates according to the geographical proximity of areas. The observed CRs^{95+} are instead denoted by a set of spots, identifying locations with strange different values from the nearby areas. The same conclusions can be alternatively reached by comparing the ranges of observed and estimated values of the CRs^{95+} (Table 2).
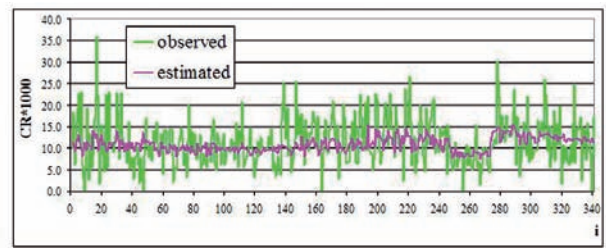
(a) 1995-1999.
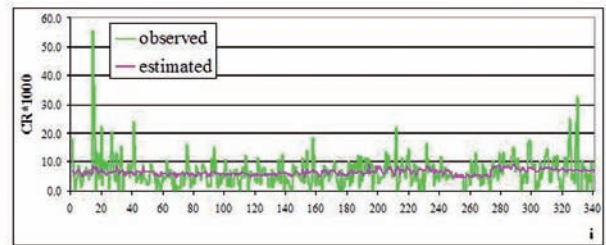
(b) 2005-2009.

Fig. 3. CRs^{95+} in men, observed and estimated by the model.

Table 1. Average CR^{95+} (x 1,000) for the overall region.

Period	Men	Women
1995-1999	6.39	21.47
2005-2009	10.68	37.54
Difference	4.29	16.07

Table 2. Decile distribution of the CRs^{95+} (x 1,000).

	Observed				Estimated			
	1995-1999		2005-2009		1995-1999[a]		2005-2009	
	Men	Women	Men	Women	Men	Women	Men	Women
Min	0.0	0.0	0.0	7.4	4.5	13.4	7.5	23.4
1st decile	0.9	4.7	5.3	23.1	5.2	18.2	8.8	31.7
2nd decile	2.4	10.2	6.7	29.6	5.5	19.5	9.3	34.0
3rd decile	3.8	13.4	8.3	33.7	5.8	20.2	9.8	35.3
4th decile	4.7	15.8	9.4	35.9	6.0	20.7	10.2	36.0
5th decile	5.6	18.1	10.5	38.2	6.4	21.4	10.8	37.3
6th decile	6.4	20.4	11.8	41.5	6.7	22.1	11.4	38.4
7th decile	7.9	23.0	13.7	44.6	7.0	23.0	11.8	40.1
8th decile	9.6	26.0	15.7	48.4	7.3	23.9	12.4	41.7
9th decile	12.0	31.0	19.7	55.2	7.9	25.9	13.3	45.1
Max	55.6	56.6	36.0	100.0	9.1	32.4	15.3	56.5

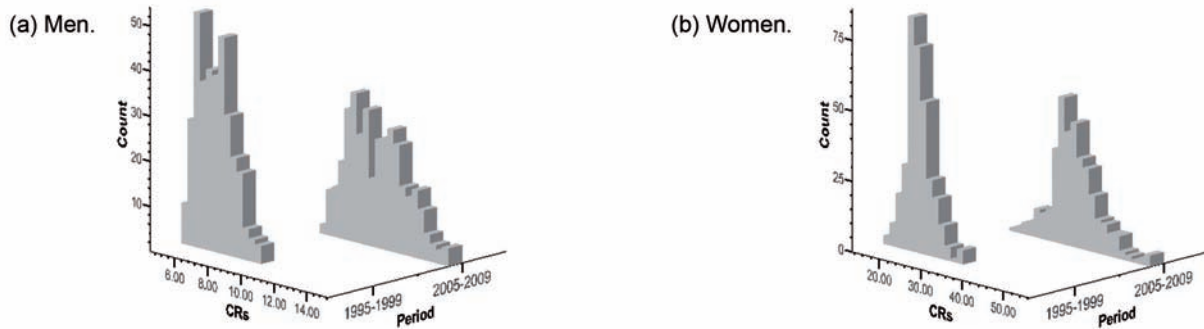[a] Starting distribution used to classify the areas in the maps.

Fig. 4. Frequency distributions of the estimated CRs^{95+}.

For all municipalities of Emilia Romagna we obtained the distributions of the male and female CRs^{95+} estimated by the model for the two periods (Fig. 4). We observed an increase of the phenomenon across time for both genders, but more evident for women. Indeed, an overall regional rise of about 4.29 points for men and 16.07 for women in the values of the CR^{95+} (per 1,000 subjects) was found (Table 1).

In order to evaluate the different territorial contributions to the CRs^{95+}, separately by gender, we first considered ranking the municipalities according to the decile[7] distribution of the estimated values of the CR^{95+} in the first period, separately by gender (Table 2). The spatial distribution of both the observed and estimated CRs^{95+} in 1995-1999 was mapped by classifying the municipalities into seven groups: the first and the last three decile subsets were considered separately and the central ones were put together (Figs. 5a, 6a, 7a and 8a). For the second period, the amount of the regional rise in the CRs^{95+} across the two periods was added to each class of the former classification. As a result, we controlled for the variation of each area in

the value of the CR^{95+} with respect to the overall increase (Figs. 5b, 6b, 7b and 8b).

Both a persistence of areas of lower and higher occurrences of P^{95+} across time and a rise in the values of the CR^{95+} were noted. In particular, mean and median values higher than the regional ones showed up in the municipalities belonging to the provinces of Ravenna and Forli-Cesena, on one side spreading out along the Adriatic coast and, on the other stretching into the Apennine municipalities of Bologna and Modena. Especially for men, some areas of the province of Piacenza still stood out showing high CR^{95+} values. Conversely, the municipalities of the province of Ferrara are characterised by a lower longevity[8]. The widening of areas characterised by the lowest and the highest values of the CR^{95+} in the sec-

---

[7] i.e.10-quantiles (see note 5).

[8] We cannot exclude that some fluctuations of the CRs^{95+} could depend on migration movements. For example, Ferrara has been an emigration area for a long time. This could lower the CR^{95+} index if we assume that emigrants have a better health profile than those who stay.
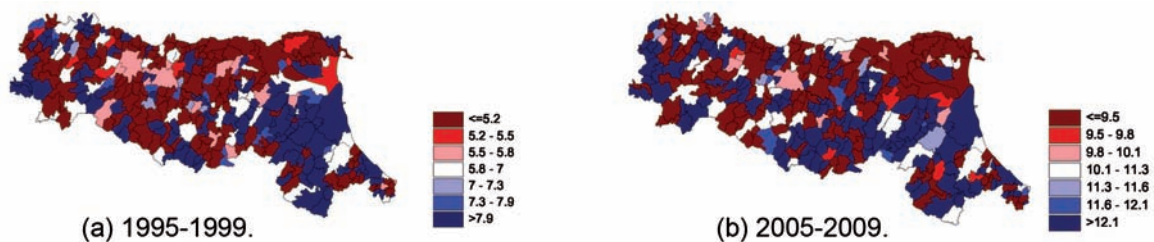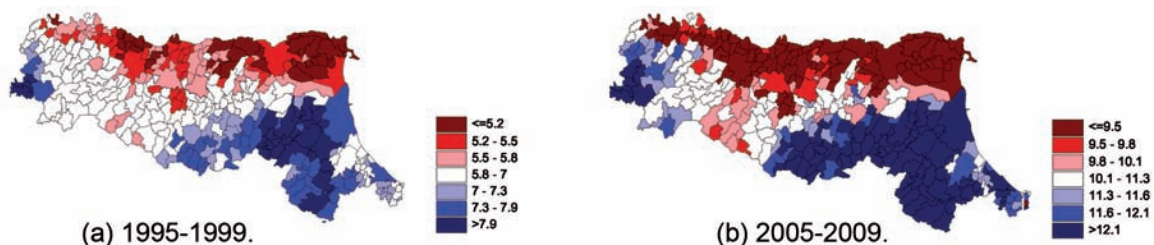

Fig. 5. The observed CRs^{95+} in men.

Fig. 6. The CRs^{95+} in men estimated by the Bayesian spatio-temporal model.
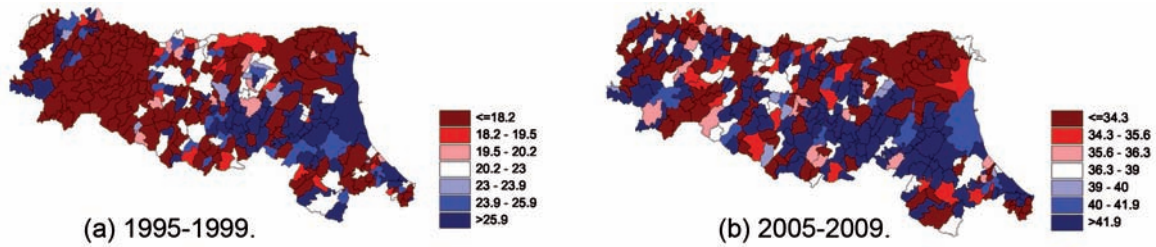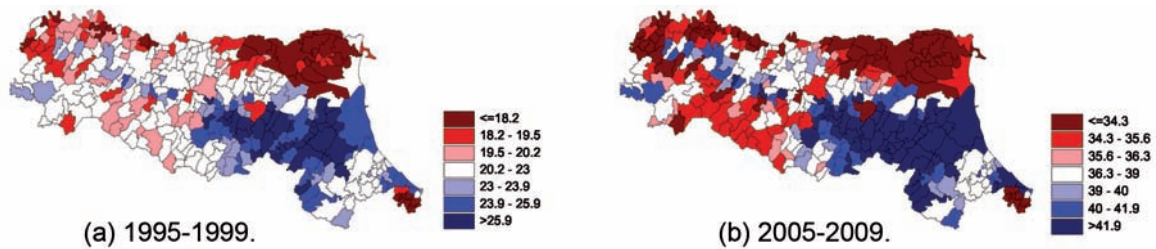
Fig. 7. The observed CRs^{95+} in women.

Fig. 8. The CRs^{95+} in women estimated by the Bayesian spatio-temporal model.

ond period showed that the smallest and largest increases of longevity concern the same municipalities identified in the first period and the adjacent ones. For women, low levels of longevity also appeared for some groups of municipalities in the Apennine areas of Parma and Reggio Emilia, in the north-west of the region (province of Piacenza) and in the coastal hills of the province of Rimini. Male longevity is low in the whole northern areas of the region, in a group of municipalities of Reggio Emilia bordering on the province of Modena and to a lesser extent in the area of the Apennine areas of Parma and Reggio Emilia.

The spatio-temporal model adopted here further allowed us to split the spatial variability into two parts. The first concerned the geographical structure modelled by the component u which detects the extra-Poisson variability in the log rate "that varies locally, so that nearby regions will have more similar rates" (Banerjee et al., 2004). It was revealed by yielding homogeneous clusters of adjacent areas. The second part of variability is due to the peculiarities of individual areas and results from municipalities with observed values that significantly differ from nearby areas. This is reflected

by the component v in the model specification which captures the overdispersion in the whole region.

The correct sentence is: "The contribution of the clustering and heterogeneity effects, in terms of total territorial variability, can first be evaluated through the estimates of the posterior variance of the two components u and v, given by VÂR (U|Y) and VÂR (V|Y), respectively (Table 3) and of the posterior proportion of the variability in the random effects that is due to clustering, i.e.

$$\delta = \frac{sd(u)}{sd(u) + sd(v)} \qquad (10)$$

where $sd(\cdot)$ is the empirical marginal standard deviation function. For both genders, the structured spatial variability seems to prevail. As a consequence, the global model representations for the CRs^{95+} are more influenced by the territorial clusters with relatively similar longevity "risks", rather than by the heterogeneity effect (80% and 78% for males and females, respectively). Once the covariates are included into the model specification, they mainly contribute to explain the homogeneity across the municipalities, as the esti-

Table 3. Posterior estimates of variances for clustering and heterogeneity components and δ.

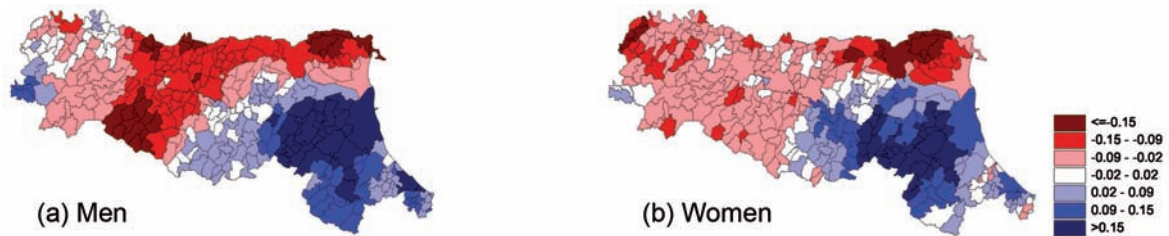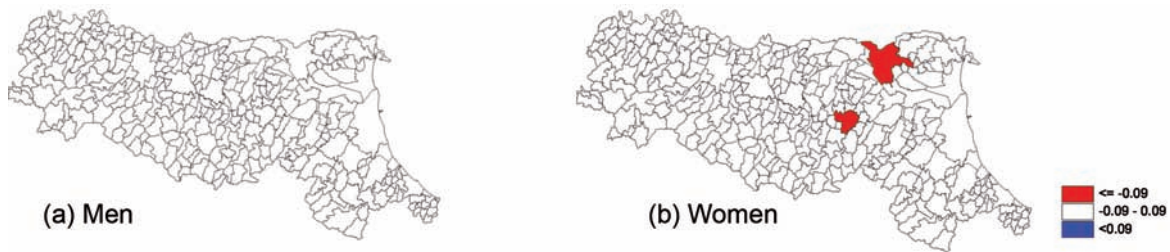|	Empty model			With covariates		
	VÂR (U|Y)	VÂR (V|Y)	δ	VÂR (U|Y)	VÂR (V|Y)	δ
Men	0.062	0.003	0.799	0.054	0.003	0.783
Women	0.053	0.003	0.766	0.037	0.005	0.670

Fig. 9. Cluster maps.

Fig. 10. Heterogeneity maps.

mate of the corresponding posterior variance decreases. In order to compare different contributions, distinct maps for each component, *u* and *v*, can be further provided (Figs. 9 and 10). These results refer to the spatio-temporal model with covariates (8).

Especially for men, two large clusters with a positive contribution on the values of the CR^{95+} were clearly identified. These substantially correspond to the same areas described before and are characterised by high CR^{95+} values. On the other hand, the Ferrara province was again confirmed the "worst" area in terms of homogeneity contribution to the values of the CR^{95+}, together with the Apennine area of the provinces of Parma and Reggio Emilia.

Conversely, the heterogeneity maps identify areas with a positive (blue) or negative (red) notable component not due to territorial adjacency. The overall contribution of the heterogeneity to the values of the CR^{95+} was small and, as a result, for men no areas appeared to have noticeable values. For women, two municipalities, Ferrara and Bologna, were found to be characterised by a strong negative contribution of the heterogeneity component. The former is part of a larger area identified by low CR^{95+} values. The latter already emerged as an area with a low longevity level with respect to the adjacent areas.

Table 4 reports the estimates of the time-dependent intercepts, while Table 5 shows the effects of the covariates included into the analysis and corresponding 90% credibility intervals (CrI).

The fitted rates for the intercepts in the two periods, $exp(\alpha_1)$ and $exp(\alpha_2)$, represent the level of probability of observing P^{95+} in the mountainous areas of Emilia Romagna, which is the reference group. An overall increase in these values was observed over time for

Table 4. Intercepts.

	Men		Women	
	Fitted rate	90% CrI	Fitted rate	90% CrI
1995-1999	0.007	0.006 - 0.008	0.022	0.020 - 0.023
2005-2009	0.012	0.011 - 0.013	0.038	0.035 - 0.040

Table 5. Covariates.

	Men		Women	
	Fitted rate ratios	90% CrI	Fitted rate ratios	90% CrI
Hill and density < 78	0.961	0.827-1.114	0.956	0.872-1.047
Hill and density 78-193	0.978	0.846-1.127	1.132	1.037-1.233
Hill and density > 193	0.852	0.741-0.973	1.027	0.936-1.126
Coastal hill	0.895	0.620-1.295	0.751	0.590-0.952
Plain and density < 78	0.772	0.611-0.972	0.743	0.640-0.862
Plain and density 78-193	0.794	0.690-0.912	0.942	0.859-1.033
Plain and density > 193	0.906	0.801-1.021	1.015	0.935-1.103

both genders. The effects of some covariates included into the analysis (Table 5) were shown, for males, by rate ratios lower than 1 in all municipalities that do not belong to the mountainous area, with significant values for hilly zones with high population density and plains with low or medium density. Conversely, in the female group, people living in hilly areas with medium population densities had a significantly higher rate of reaching 95 years of age and above than the residents of the mountainous areas. Significant values were also provided for coastal hill and plain municipalities with low densities. These are both areal characteristics which seem to negatively affect the presence of oldest-old individuals.

Table 6. Effect of past mortality on P^{95+} in 2005-2009.

	Men	
	Fitted rate ratios	90% CrI
All causes[a] (80-89)		
Medium mortality	1	-
Low mortality	1.005	0.904 - 1.116
High mortality	0.864	0.783 - 0.953
All causes[a] (90-99)		
Medium mortality	1	-
Low mortality	1.304	1.192 - 1.425
High mortality	0.719	0.638 - 0.811

[a] Excluding injuries and poisoning.

The regression effect of mortality patterns

Through the longitudinal spatial analysis we would assess the relationship between the proportions of P^{95+} and the levels of mortality above 80 years of age across the areas by evaluating the effects of both the "recent" (age group 90-99 years) and "past" (80-89 years) structures of mortality on the same cohorts of individuals. In particular, we supposed and allowed for the possibility of different contributions of the mortality levels in the two classes of age on male rather than on female longevity distribution. We considered only the period 2005 to 2009 when the larger territorial differences were observed for both genders.

The results on the fitted rate ratios and corresponding 90% CrI are shown in Tables 6 and 8. In the former, we observed that:

(i)   the lower the ranking in mortality levels, the higher the values of the CR^{95+} (with stronger associations for males), and vice versa;

(ii) the negative effects of high levels in "past" mortality (80-89 years) are substantial and significant for men, but not for women; and

(iii) for both genders, the mortality at 90-99 years of age seems to exert a clear influence on the current proportions of P^{95+} in the study areas.

Table 7. Effect of the past mortality on P^{95+} in 2005-2009: groups of areas.

	Fitted rate ratios	
	Men	Women
Mountain and medium mortality at ages 80-89 and 90-99 years	1.000	1.000
Hill and density <78 people per km² and medium mortality at ages 80-89 and 90-99 years	1.046	0.989
Hill and density 78-193 people per km² and medium mortality at ages 80-89 and 90-99 years	1.076	1.207
Hill and density >193 people per km² and medium mortality at ages 80-89 and 90-99 years	0.824	0.982
Coastal hill and medium mortality at ages 80-89 and 90-99 years	0.864	0.667
Plain and density <78 people per km² and medium mortality at ages 80-89 and 90-99 years	0.682	0.772
Plain and density 78-193 people per km² and medium mortality at ages 80-89 and 90-99 years	0.767	0.908
Plain and density >193 people per km² and medium mortality at ages 80-89 and 90-99 years	0.889	0.974
Mountain and low mortality at ages 80-89 and 90-99 years	1.311	1.243
Hill and density <78 people per km² and low mortality at ages 80-89 and 90-99 years	1.371	1.230
Hill and density 78-193 people per km² and low mortality at ages 80-89 and 90-99 years	1.410	1.501
Hill and density >193 people per km² and low mortality at ages 80-89 and 90-99 years	1.080	1.220
Coastal hill and low mortality at ages 80-89 and 90-99 years	1.132	0.829
Plain and density <78 people per km² and low mortality at ages 80-89 and 90-99 years	0.894	0.960
Plain and density 78-193 people per km² and low mortality at ages 80-89 and 90-99 years	1.005	1.128
Plain and density >193 people per km² and low mortality at ages 80-89 and 90-99 years	1.165	1.211
Mountain and high mortality at ages 80-89 and 90-99 years	0.622	0.857
Hill and density <78 people per km² and high mortality at ages 80-89 and 90-99 years	0.650	0.847
Hill and density 78-193 people per km² and high mortality at ages 80-89 and 90-99 years	0.669	1.034
Hill and density >193 people per km² and high mortality at ages 80-89 and 90-99 years	0.512	0.841
Coastal hill and high mortality at ages 80-89 and 90-99 years	0.537	0.571
Plain and density <78 people per km² and high mortality at ages 80-89 and 90-99 years	0.424	0.662
Plain and density 78-193 people per km² and high mortality at ages 80-89 and 90-99 years	0.477	0.778
Plain and density >193 people per km² and high mortality at ages 80-89 and 90-99 years	0.553	0.834

All the causes of deaths (excluding injuries and poisonings) are considered.

In order to investigate the effects of altitude and density of population, we further considered the contribution to the fixed effects of the model (i.e., excluding the homogeneity and heterogeneity components) of some groups of areas characterised by altitude and density groups and levels of mortality at ages 80-89 and 90-99 years (Table 7). For both genders the highest value of fitted rate ratio was experienced in hilly areas with a population density between 78 and 193 people per km² and low mortality at both periods of age studied (the reference group is mountainous areas with medium mortality at both classes of age). For females, a fitted rate ratio greater than 1 for this group of municipalities was found to holds on even when the levels of mortality are high. As a result, altitude and population density represent important features to explain the proportion of P^{95+} in these areas. Conversely, the lowest values of the CR^{95+} were observed, for men, in plain municipalities with a population density lower than 78 people per km² and, for women, in the coastal hill areas. It is noteworthy that these areas keep on having an estimated rate ratio lower than 1 even when their mortality is low, suggesting again the crucial role of altitude and population density with respect to the proportion of P^{95+}.

When the specific contribution of the major causes of deaths was evaluated (Table 8), it was confirmed that female mortality at 80-89 years of age does not exert any significant influence on the current count of P^{95+} when it is due to both circulatory disease and malignant neoplasms. Areas with high proportions of P^{95+} were shown to be significantly associated with low levels of mortality due to malignant neoplasms at 80-89 years of age for males and due to circulatory diseases at age 90-99 years for both sexes. Conversely, a negative and significant effect of high mortality levels due to malignant neoplasms was noticeable at age 80-89 years for men and to both the causes at age 90-99 years regardless of gender.

Discussion

Statistical theory, several simulation studies and a large number of applications all support the use of hierarchical Bayesian modelling for spatial and spatio-temporal data as a powerful method which results in more consistent estimates on quantities of interest and inferences (Lawson, 2009). Indeed, the consequent smoothing effect, both in the spatial and the temporal sense, allows controlling for the variation in the population size across the geographical areas (Clayton and Kaldor, 1987) and for extra-variation among the units (Breslow, 1984). Moreover, the opportunity of including different levels of covariates permits investigating some crucial effects for the study at hand and explaining parts of the variability.

A critical aspect of these methods is certainly represented by the sensitivity regarding the choice of the hyperpriors (Bernardinelli et al., 1995), which can sometimes be avoided by the use of an empirical-Bayes approach (Carlin and Louis, 1998) rather making assumptions, reasonable for the problem at hand (Besag, 1974; Banerjee et al., 2004). In this application, a crucial point is represented by the choice of the Gamma priors for the precision parameters τ_u and τ_v, which control the strength of the smoothing and heterogeneity effect. These components are shown to be quite sensitive with respect to prior specification, but

Table 8. Effect of the past mortality on P^{95+} in 2005-2009 by cause.

	Men		Women	
	Fitted rate ratios	90% CrI	Fitted rate ratios	90% CrI
Circulatory diseases (age group 80-89 years)				
Medium mortality	1	-	1	-
Low mortality	1.091	0.996-1.195	1.005	0.947-1.067
High mortality	0.980	0.878-1.094	0.979	0.920-1.042
Circulatory diseases (age group 90-99 years)				
Medium mortality	1	-	1	-
Low mortality	1.197	1.081-1.326	1.068	1.009-1.130
High mortality	0.797	0.714-0.914	0.848	0.795-0.906
Malignant neoplasms (age group 80-89 years)				
Medium mortality	1	-	1	-
Low mortality	1.124	1.012-1.248	1.000	0.937-1.066
High mortality	0.828	0.752-0.909	0.997	0.943-1.054
Malignant neoplasms (age group 90-99 years)				
Medium mortality	1	-	1	-
Low mortality	1.070	0.963-1.190	1.036	0.974-1.100
High mortality	0.896	0.808-0.991	0.926	0.871-0.983

the effect is mainly only reflected by the variances of the parameters (Divino et al., 2009). Therefore, the posterior estimates are only weakly affected.

The use of a hierarchical Bayesian modelling for spatial and spatio-temporal data for the analysis of the longevity pattern in the municipalities of Emilia Romagna represents an innovative application that confirms the usefulness of these methods also in socio-demographic research. In the spatio-temporal application, the estimated values of the adopted longevity index allow the quantification of the rise of this phenomenon over time. Moreover, some groups of neighbouring areas, where people are likely to reach 95 years of age or more, can be identified and consistently observed over time. The picture detected shows, on the one hand, the presence of quite large areas which are homogeneous with respect to the CR^{95+} values suggesting that some environmental factors are likely to influence the spatial distribution of longevity. On the other hand, the detection of isolated areas characterised by high levels of heterogeneity with respect to neighbour municipalities, calls for in-depth analysis to reveal the specific reasons causing the particularly high or low levels of longevity. As far as the effects of altitude and population density are concerned, the plains are generally shown to correlate with a lower count of the oldest-old, which was also detected in Campania and in zones of Sicily (Lipsi, 2009). Moreover, hilly zones seem to be favourable for longevity, at least for females.

The kind of modelling used here offers the opportunity of adopting a cohort perspective in order to study several aspects in depth, which can affect the survival of people aged 95 and above, and explain the territorial differences in the period 2005-2009. Several studies (Thatcher, 1992, 2001; Vaupel and Jeune, 1995; Robine and Paccaud, 2005; Caselli and Lipsi, 2006) assert that the decline in mortality seen after the age of 80 years mainly explains the relevant increase in the number of the oldest-old. Therefore, a negative association between the proportions of P^{95+} and the levels of mortality above 80 years of age across the areas should be expected. Under this perspective, we jointly investigated the effects of some crucial patterns of mortality by sex and age group, as well as a set of areal features and correlations. As a result, different contributions of the levels of mortality on longevity for males and females appeared. As far as the major causes of death are concerned, the circulatory diseases in the most recent age group seem to mostly affect the occurrence of P^{95+}. In addition, a prominent, positive effect of altitude and population density on longevity emerged.

The covariates included into the analysis appear to mostly explain the reason why people do not reach the age of 95 years and above. Therefore, it becomes crucial for further development along the lines of this study to find additional covariates associated with the territorial areas (e.g. indexes of deprivation and pollution, environmental indicator, presence of incinerators), together with a more in-depth analysis of areas characterised by the highest levels of the CR^{95+}.

References

Banerjee S, Carlin BP, Gelfand AE, 2004. Hierarchical Modeling and Analysis for Spatial Data. Chapman and Hall, London, UK, 452 pp.

Bernardinelli L, Clayton DG, Montomoli C, 1995. Bayesian estimates of disease maps: how important are priors? Stat Med 14, 2411-2431.

Bernardinelli L, Montomoli C, 1992. Empirical Bayes versus fully Bayesian analysis of geographical variation in disease risk. Stat Med 11, 983-1007.

Besag J, 1974. Spatial interaction and the statistical analysis of lattice systems (with discussion). J Roy Stat Soc B Met 36, 192-236.

Best NG, Waller LA, Thomas A, Conlon EM, Arnold RA, 1999. Bayesian models for spatially correlated diseases and exposure data. In: Bayesian Statistics 6 (JM Bernardo et al. eds). Oxford University Press, Oxford, UK, pp. 131-156.

Breslow NE,1984. Extra-Poisson variation in log-linear models. Appl Stat-J Roy St C 33, 33-44.

Breslow NE, Clayton DG, 1993. Approximate inference in generalized linear mixed models. J Am Stat Assoc 88, 9-25.

Carlin B, Louis T, 1998. Bayes and empirical Bayes methods for data analysis. Chapman and Hall/CRC, New York, USA, 448 pp.

Caselli G, Lipsi RM, 2006. Survival differences among the oldest old in Sardinia: who, what, where and why? Demogr Res 14, 267-294.

Clayton DG, Kaldor J, 1987. Empirical Bayes estimates of age-standardised relative risks for use in disease mapping. Biometrics 43, 671-681.

Divino F, Egidi V, Salvatore MA, 2009. Geographical mortality patterns in Italy: a Bayesian analysis. Demogr Res 20,435-466.

Gelman A, Carlin JB, Stern HS, Rubin DB, 2003. Bayesian Data Analysis. 2nd edn. Chapman and Hall/CRC, New York, USA, 668 pp.

ISTAT, 1997. Classificazione delle malattie, traumatismi e cause di morte, IX Revisione 1975. Roma: Istituto Nazionale di Statistica, Metodi e Norme, Serie C, No. 10, Voll. 1 e 2, V ristampa, 483 pp.

Kahn HA, Sempos CT, 1989. Statistical Methods in Epidemiology. Oxford University Press, Oxford, UK, 292 pp.

Knorr-Held L, 2000. Bayesian modelling of inseparable space-time variation in disease risk. Stat Med 19, 2555-2567.

Langford IH, Leyland AH., Rasbash J, Goldstein H, 1999.Multilevel modelling of the geographical distributions of diseases. J Roy Stat Soc C-App 48, 253-268.

Lawson AB, 2009. Bayesian disease mapping, CRC press, 364 pp.

Lipsi RM, 2009. Longevity in small areas and their socioeconomic, demographic and environmental characteristics: a hierarchical Bayesian approach. Poster presented at XXVI IUSSP International Population Conference, Marrakech.

Miglio R, Marino M, Rettaroli R, Samoggia A, 2009. Spatial analysis of longevity in a northern Italian region, paper presented at the PAA Annual Conference, Detroit, USA.

Mollie A, 1994. Bayesian Mapping of Disease. In: Gilks W, Richardson S, Spiegelhalter DJ (eds.). Markov Chain Monte Carlo in Practice. Chapman and Hall, London, UK, 359-379.

Montesanto A, Passarino G, Senatore A, Carotenuto L, De Benedictis G, 2008. Spatial analysis and surname analysis: complementary tools for shedding light on human longevity patterns. Ann Hum Genet 72, 253-260.

Poulain M, Pes G, Grasland C, Carru C, Ferrucci L, Baggio G, Franceschi C, Deiana L, 2004. Identification of a geographic area characterised by extreme longevity in the Sardinia Island:

the AKEA study. Exp Gerontol 39, 1423-1429.

Robine JM, Caselli G, 2005. An unprecented increase in the number of centenarians. Genus LXI, 57-82.

Robine JM, Caselli G, Rasulo D, Cournil A, 2006. Differentials in the femininity ratio among centenarians: variations between northern and southern Italy from 1870. Pop Stud-J Demog 60, 99-113.

Robine JM, Paccaud F, 2005. Nonagenarians and centenarians in Switzerland, 1860-2001: a demographic analysis. J Epidemiol Commun H 59, 31-37.

Spiegelhalter D, Thomas A, Best N, Lunn D, 2003. WinBUGS User Manual, Version 1.4.

Thatcher AR, 1992. Trends in numbers and mortality at high-ages in England and Wales. Pop Stud-J Demog 46, 411-426.

Thatcher AR, 2001 The demography of centenarians in England and Wales. Population-An English Selection. INED 13, 1-139.

Vaupel JW, Jeune B, 1995. The emergence and proliferation of centenarians. In Jeune B., Vaupel JW (eds). Exceptional longevity: from prehistory to the present. Odense University Press, 109-116.

Waller LA, Carlin BP, Xla H, Gelfand AE, 1997. Hierarchical spatio-temporal mapping of disease rates. J Am Stat Assoc 92, 607-617.