Scientific Research and Essays Vol. 7(45), pp. 3860 –3876, 19 November, 2012
Available online at http://www.academicjournals.org/SRE
DOI: 10.5897/SRE11.1754
ISSN 1992-2248

Full Length Research Paper

Clustering of mortality rates in Greece's prefectures

Georgia Pistolla*[1], Poulicos Prastacos[2], Nikos Tzanakis[3] and Anastas Philalithis[1]

[1]Department of Social Medicine, Faculty of Medicine, University of Crete, Greece.
[2]Institute of Applied and Computational Mathematics, Foundation for Research and Technology-Hellas (FORTH), Herakleion, Greece.
[3]Department of Clinical Epidemiology, Faculty of Medicine, University of Crete, Greece.

Accepted 15 November, 2012

The purpose of this paper is to present of the geographical variation of mortality rates in the 51 prefectures of Greece per gender in 2001 and 2006. The database consisted of the mortality rates of 65 causes of mortality for all 51 prefectures of Greece in two different time periods: 2001 and 2006. All rates were calculated using the "direct method of standardization" procedure and were selected from Natural Population Research Sector of the Greek Statistical Authority (EL.STAT.). Clustering analysis methods (CAM) was applied to identify the mortality patterns per gender and cause among prefectures and time. The results of the analysis were mapped, indicating a clustered distribution of these rates. It suggests a new way of health planning and supports local policies, since a common pattern can be identified.

Key words: Mortality rates, sub regional analysis, clustering, Greece.

INTRODUCTION

Mortality rates are one of the most representative measures for public health standards. They are affected by numerous human-origin, social, environmental and economic parameters. As every phenomenon that occurs in space and time, they also are characterized by spatial/geographical principles and characteristics. These spatial principles vary from place to place even in the same country. This fact indicates the existence of several causes that affect and lead to these geographical variations. A clustering of areas with similar mortality rates could be a useful procedure that would help to detect the differences and variations between them (Douven and Schilter, 1995; Cassetti et al., 2008).

The analysis of mortality rates is usually carried out using the national and NUTS 2 level, which is the Nomenclature of Territorial Units for Statistics among the European Union (E.U.) (European Commission, Eurostat, 2008-2011). Apart from using a reliable measure and reference level, it is also necessary to analyze mortality rates geographically, in order to study this phenomenon as closer in its real "environment" as it is possible and identify the significant similarities or differences. Detection of common patterns and characteristics through the use of clustering procedures, as well as geographical distribution contributes to improved health services and permits public health authorities to create realistic plans and policies acknowledging the needs even at a local level (Bilancia and Fedespina, 2009; Havard et al., 2009). Finally, the identification of relationships affecting the geographical distribution of the standardized mortality rates is, as far as we know, an open scientific research field of Medical Geography (Havard et al., 2009).

Some major causes of mortality are cardiovascular diseases, cancer and diseases of the respiratory system (WHO, 2008). Cancer caused about six million deaths in 1990 and about 3.4 million in men for the same year. Almost 2.4 million cancer deaths occurred in established market economies and formerly socialist economies of Europe. By 1990, therefore, there were already 50% more cancer deaths in less developed countries than in developed countries (Christopher and Lopez, 1997; Peto et al., 1992). Cardiovascular disease is considered to be the first cause of mortality. In Europe the results of a

*Corresponding author. E-mail: geopist@gmail.com.

relevant cross-sectional study underline the highly variable nature of socioeconomic inequalities in ischemic heart disease mortality in eleven European countries. These inequalities appear to be highly sensitive to social gradients in behavioral risk factors (Peto et al., 1992).

This paper aims to study the mortality rates per cause, gender and prefecture in Greece in 2001 and 2006. Its main research question is whether there is a clustering pattern of mortality rates and identify this pattern. Furthermore, mortality rates' variation per prefecture and gender will be studied in order to identify the possible local particularities that affect mortality and propose intervention plan and policies in Greece.

MATERIALS AND METHODS

Mortality rates at the prefecture level

Data of the number of deaths per cause, gender and prefecture for 2001 and 2006 were obtained from the Hellenic Statistical Authority (HEL.STAT.) (Hellenic Statistical Authority, 2008). The ICD10 classification system of causes of mortality was used, while the different causes were grouped using the E.U. standards. A total of 65 different causes of death were used in the analysis (European Commission, Eurostat, 2008-2011). The direct standardization was carried out per gender, using the G27 population as a reference. This was applied to calculate the standardized rates per 1,000.000 people (SDR) for the whole country and for each prefecture (Pagano and Gauvreau, 1996).

The clustering procedure

As a first step of the analysis, Cluster Analysis Methods (CAM) and techniques were performed for the detection and study of multi-variant and multi-parameter phenomena.

A naive definition of clustering could be the following: "clustering is the process of organizing objects into groups whose members are similar in some way" (Osmar, 2010; Bilancia and Fedespina, 2009). A cluster is a collection of objects which are "similar" among each other and are "dissimilar" to the objects belonging to other clusters. The choice of the norm that will be used to create distance measures is a very important issue of clustering algorithm. Euclidean distance was chosen and performed in this case, due to its reliable function on simple numeric data of same metric units (Osmar, 2010).

These definitions were the criteria used for this part of the analysis in the present study. The grouping of mortality rates in each prefecture per rate, cause and gender was carried out with the CAM, and particularly the k-means cluster, using the program SPSS 18.00 (Dafermos, 2005). Mortality rates were considered to be the dependent variables, and the clustering factors independent variables. The results of our cluster analysis were mapped using the Geographic Information Systems (GIS) software ARC-GIS 9.3 (ESRI, 2000-2004).

K-Means clustering

According to the above mentioned theories and principles, the K-means were applied as the simplest algorithms that can perform clustering procedure (MacQueen, 1967). The number of clusters is fixed a priori, say k. The main idea is to define k centroids, one for each cluster, and these centroids should be placed in a meaningful way, due to the fact that different locations cause different results. This means that the best choice is to place them as far away from each other as it is possible. Then, we must associate each point belonging to a given set of data to the nearest centroid. When no point is pending, the first step is completed and an early grouping (clustering) is fulfilled. At this point we need to re-calculate the k centroids as barycenters of the clusters resulting from the previous step. By the time we have the new k centroids, a new binding has to be done between the same data set points and the nearest new centroid. We continue in the same way, since the k centroids change their location step by step until no more changes are done. In other words, the centroids do not move any more. This algorithm minimizes an objective function, specifically the squared error function (MacQueen, 1967). The objective function is the following:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^{(j)} - c_j \right\|^2 ,$$

where $\left\| x_i^{(j)} - c_j \right\|^2$ is a chosen distance measure between a data point $x_i^{(j)}$ and the cluster centre c_j, is an indicator of the distance of the n data points from their respective cluster centres.

So, the steps of the above algorithm are the following:

1) First place k points into the space represented by the objects that are being clustered. These points represent initial group centroids.
2) Secondly, assign each object to the group that has the closest centroid.
3) When all objects have been assigned, recalculate the positions of the k centroids and finally
4) Repeat Steps 2 and 3 until the centroids no longer move.

This algorithm produces a separation of the objects into groups and minimizes the metric.

The multi-clustered method

Finally, the multi-clustered method was applied. It is a geographical clustering of the prefectures according to a common element. There were 65 causes of mortality, which refers to 65 standardized mortality rates for 2001 and 2006. In order to create the multi-clustered map we found the difference between the mortality rates of 2001 to 2006. Some of these differences were zero, very small and of no significant and only five of them were high and significant. The significant ones were malignant neoplasms, malignant neoplasm of pancreas, diseases of the respiratory system, cerebrovascular disease and other heart diseases (Diagram 1).

RESULTS

Mortality rates in Greece were observed to differ from region to region. The major mortality causes in Greece were the following: diseases of the circulatory system (especially ischaemic heart disease), celebrovascular diseases, malignant neoplasms (mainly malignant neoplasm of the larynx, trachea, bronchus and lung), as presented in Table 1. Mortality rates with zero value refer to causes of mortality that present very low mortality rates

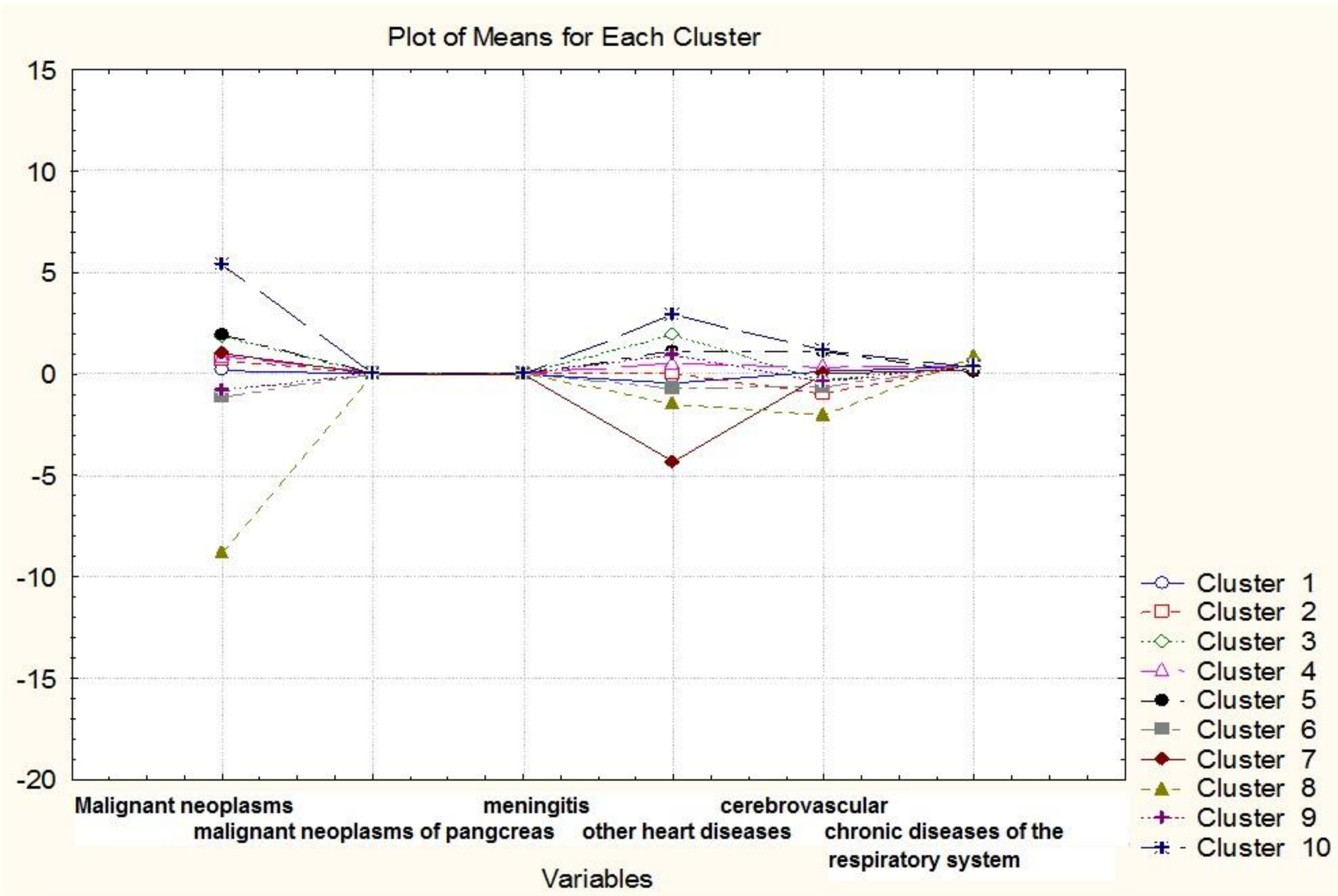


Diagram 1. Multi-clustered plot (five groups).

per 1,000,000 people (for instance, a mortality rate that is equal to 0.00004 per 1,000,000 people).

In Table 2, results of cluster analysis for the 65 standardized per gender mortality rates and for the 51 prefectures are presented. They create groupings of geographical entities with similar data structure patterns. Low, medium, high and very high clusters refer to a grouping of prefectures that present similar mortality patterns. For instance, 26 prefectures are grouped together for ischemic heart disease (men) in 2001 in a medium cluster, whereas 17 prefectures are grouped in the same cluster in 2006. The number of prefectures that belong to each cluster is different for both year and gender. The fact that there is a continuous "movement" of prefectures from one cluster to another cluster based on the gender and the time period, means that prefectures present different mortality patterns from one another. Therefore, mortality could be considered as a geographical phenomenon with local characteristics and indicators.

Unfortunately, due to the high number of maps (305), they are not presented in this paper, yet they are available in the official site of the University of Crete, Department of Medicine (Health Planning Unit, 2011). In these 305 maps of clusters (CAM), it is easily observed that there are significant differences in both causes and genders per prefecture and cluster of prefectures. For instance, the low ischaemic heart rate of 2001, in the prefecture of Evrytania jumps to a high level in 2006, for both genders. Also, members of clusters change from 2001 to 2006, according to their magnitude change in both directions: low or high in both genders (Maps 1 to 4).

As far as celebrovascular diseases are concerned, another indicative finding is that the group for men involved four prefectures in 2001 while it was only one in 2006. The feminine population increases its rates in the prefecture of Chios in 2006. More prefectures enter the clusters of low rates for the same cause of mortality. The mortality rate in the prefecture of Evrytania appears as similar trends with the previous cause of mortality (Maps 5 to 8). As far as malignant neoplasm of the larynx and trachea/ bronchus/ lung in men both in 2001 and 2006 are concerned, the results showed that low rated prefectures in 2001 moved to the high rate groups in

Table 1. Mortality rates per cause of mortality in Greece in 2001 and 2006.

Mortality cause	National rates 2001	Regional rates Mean / Standard deviation 2001	Range 2001	National rates 2006	Regional rates Mean / Standard deviation 2006	Range 2006
01 Infectious and parasitic diseases (A00-B99)	0.60	0.0136(0.01686)	0.00-0.10	0.81	0.0159(0.01723)	0.00-0.07
02 Tuberculosis (A15-A19.B90)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
03 Meningococcal infection (A39)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
04 AIDS (HIV-disease) (B20-B24)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
05 Viral hepatitis (B15-B19)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
06 Neoplasms (C00-D48)	0.03	0.03(0.00)	0.03-0.03	0.03	0.03(-)	0.03-0.03
07 Malignant neoplasms (C00-C97)	161.59	3.1684(1.2147)	0.27-9.38	166.40	3.2627(1.10615)	0.53-8.94
08 Malignant neoplasm of lip, oral cavity, pharynx (C00-C14	0.10	0.0031(0.00471)	0.00-0.01	0.24	0.0051(0.00547)	0.00-0.02
09 Malignant neoplasm of oesophagus (C15)	0.05	0.0015(0.00359)	0.00-0.01	0.09	0.0020(0.00408)	0.00-0.01
10 Malignant neoplasm of stomach (C16)	2.22	0.0435(0.02288)	0.00-0.11	2.28	0.0447(0.02212)	0.01-0.12
11 Malignant neoplasm of colon (C18)	3.18	0.0663(0.033370)	0.02-0.22	3.78	0.0741(0.03366)	0.02-0.22
12 Malignant neoplasm of rectum and anus (C19-C21)	0.17	0.0041(0.00499)	0.00-0.01	0.20	0.0047(0.00550)	0.00-0.02
13 Malignant neoplasm liver and the intrahepatic bile ducts (C22)	0.16	0.0041(0.00549)	0.00-0.02	0.17	0.0045(0.00504)	0.00-0.01
14 Malignant neoplasm of pancreas (C25)	1.44	0.0288(0.01423)	0.00-0.09	1.78	0.0349(0.01678)	0.02-0.10
15 Malignant neoplasm of larynx and trachea/ bronchus/ lung (C32-C34)	45.48	0.8918(0.34144)	0.09-2.38	49.60	0.9725(0.31668)	0.27-2.44
16 Malignant melanoma of skin (C43)	0.01	0.0002(0.00151)	0.00-0.01	0.10	0.0022(0.01191)	0.00-0.08
17 Malignant neoplasm of breast (C50)	5.70	0.1118(0.06263)	0.00-0.48	6.30	0.1235(0.06118)	0.04-0.48
18 Malignant neoplasm of cervix uteri (C53)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
19 Malignant neoplasm of other parts of uterus (C54-C55)	0.10	0.0024(0.00435)	0.00-0.01	0.13	0.0029(0.00458)	0.00-0.01
20 Malignant neoplasm of ovary (C56)	0.44	0.0105(0.00623)	0.00-0.03	0.50	0.0109(0.00725)	0.00-0.04
21 Malignant neoplasm of prostate (C61)	3.84	0.0753(0.03727)	0.00-0.23	4.51	0.0902(0.03384)	0.05-0.23
22 Malignant neoplasm of kidney (C64)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
23 Malignant neoplasm of bladder (C67)	0.97	0.0190(0.01044)	0.00-0.06	0.96	0.0196(0.00865)	0.00-0.06
24 Malignant neoplasm of lymphatic haematopoietic tissue (C81-C96)	3.69	0.0724(0.03302)	0.00-0.24	3.77	0.0739(0.02980)	0.02-0.23
25 Diseases of the blood(-forming organs). immunological disorders (D50-D89)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
26 Endocrine, nutritional and metabolic diseases (E00-E90)	0.79	0.0180(0.01212)	0.00-0.05	0.52	0.0124(0.00878)	0.00-0.03
27 Diabetes mellitus (E10-E14)	1.37	0.0269(0.01581)	0.00-0.09	2.88	0.0576(0.02737)	0.02-0.15
28 Mental and behavioural disorders (F00-F99)	0.21	0.0064(0.00699)	0.00-0.02	0.10	0.0032(0.00599)	0.00-0.02
29 Alcoholic abuse (including alcoholic psychosis) (F10)	0.00	0.00(0.00)	0.00-0.00	0.01	0.0010(0.00316)	0.00-0.01
30 Drug dependence. toxicomania (F11-F16,F18-F19)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
31 Diseases of the nervous system and the sense organs (G00-H95)	0.91	0.0194(0.01258)	0.00-0.05	1.28	0.0272(0.01556)	0.01-0.07
32 Meningitis (other than 03) (G00-G03)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
33 Diseases of the circulatory system (I00-I99)	534.22	10.4749(4.01102)	1.88-29.84	526.24	10.3184(3.45834)	5.55-27.99
34 Ischaemic heart diseases (I20-I25	208.95	4.0971(1.79455)	0.47-14.28	205.48	4.0290(1.59512)	2.60-13.59

Table 1. Contd.

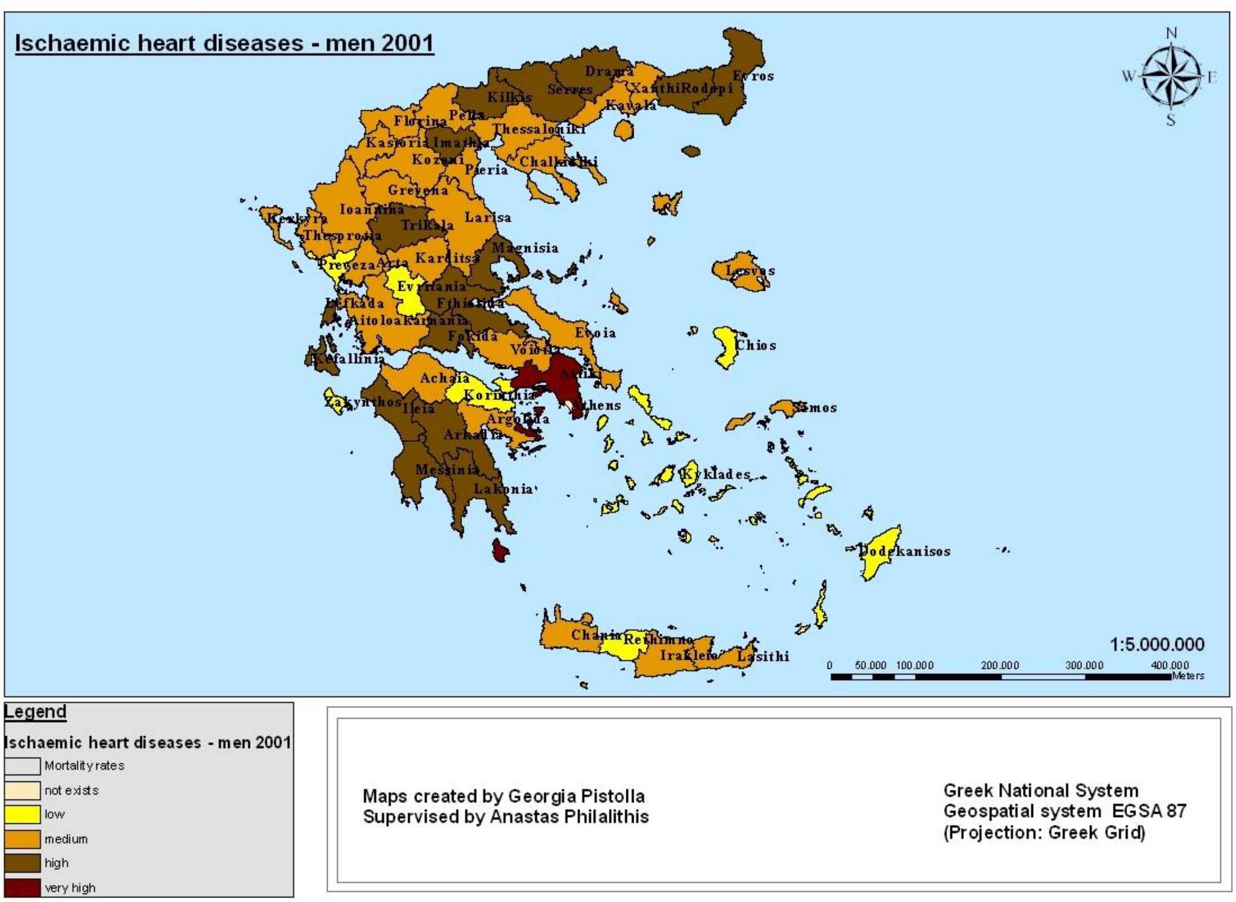
35 Other heart diseases (I30-I33,I39-I52)	81.15	1.6230(0.85801)	0.27-5.93	85.44	1.6753(0.78865)	0.56-4.44
36 Cerebrovascular diseases (I60-I69)	248.31	4.8688(1.61301)	0.51-11.24	224.10	4.3941(1.28144)	0.85-9.59
37 Diseases of the respiratory system (J00-J99)	46.89	0.9194(0.50177)	0.09-3.30	57.69	1.1538(0.57135)	0.54-4.25
38 Influenza (J10-J11)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
39 Pneumonia (J12-J18)	1.69	0.0352(0.02783)	0.00-0.13	1.97	0.0410(0.02699)	0.01-0.15
40 Chronic lower respiratory diseases (J40-J47)	20.14	0.3948(0.19061)	0.03-0.99	22.35	0.4382(0.20967)	0.18-1.08
41 Asthma (J45-J46)	0.00	0.00(0.00)	0.00-0.00	0.39	0.0244(0.07071)	0.00-0.26
42 Diseases of the digestive system (K00-K93)	6.72	0.1318(0.07301)	0.01-0.44	6.38	0.1302(0.06146)	0.03-0.41
43 Ulcer of stomach. duodenum and jejunum (K25-K28)	0.05	0.0011(0.00321)	0.00-0.01	0.05	0.0011(0.00312)	0.00-0.01
44 Chronic liver disease (K70, K73-K74)	1.32	0.0264(0.01290)	0.00-0.07	1.80	0.0353(0.01567)	0.01-0.09
45 Diseases of the skin and subcutaneous tissue (L00-L99)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
46 Diseases of the musculoskeletal system/connective tissue (M00-M99)	0.07	0.0017(0.00377)	0.00-0.01	0.08	0.0019(0.00450)	0.00-0.02
47 Rheumatoid arthritis and osteoarthrosis (M05-M06, M15-M19)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
48 Diseases of the genitourinary system (N00-N99)	0.05	0.0028(0.00575)	0.00-0.02	0.09	0.0035(0.00485)	0.00-0.01
49 Diseases of kidney and ureter (N00-N29)	1.64	0.0322(0.01880)	0.00-0.11	1.97	0.0386(0.01662)	0.01-0.10
50 Complications of pregnancy. childbirth and puerperium (O00-O99)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
51 Certain conditions originating in the perinatal period (P00-P96)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
52 Congenital malformations and chromosomal abnormalities (Q00-Q99)	0.01	0.0050(0.00707)	0.00-0.01	0.00	0.00(0.00)	0.00-0.00
53 Congenital malformations of the nervous system (Q00-Q07)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
54 Congenital malformations of the circulatory system (Q20-Q28)	0.00	0.00(0.00)	0.00-0.00	0.00	0.00(0.00)	0.00-0.00
55 Symptoms. Signs, abnormal findings, ill-defined causes (R00-R99)	0.18	0.0075(0.00737)	0.00-0.02	0.21	0.0095(0.00653)	0.00-0.02
56 Sudden infant death syndrome (R95)	0.29	0.0242(0.03502)	0.00-0.12	0.00	0.00(0.00)	0.00-0.00
57 Unknown and unspecified causes (R96-R99)	2.13	0.0484(0.04297)	0.00-0.19	2.36	0.0482(0.04716)	0.00-0.22
58 External causes of injury and poisoning (V01-Y89)	7.75	0.1520(0.09687)	0.01-0.70	6.43	0.1312(0.07135)	0.04-0.49
59 Accidents (V01-X59)	8.45	0.1690(0.07203)	0.00-0.40	7.30	0.1431(0.06263)	0.01-0.43
60 Transport accidents (V01-V99)	0.15	0.0300(0.03082)	0.00-0.08	0.00	0.00(0.00)	0.00-0.00
61 Accidental falls (W00-W19)	1.77	0.0361(0.02008)	0.00-0.14	1.84	0.0361(0.01674)	0.01-0.12
62 Accidental poisoning (X40-X49)	0.06	0.0013(0.00397)	0.00-0.02	0.08	0.0018(0.00442)	0.00-0.02
63 Suicide and intentional self-harm (X60-X84)	1.22	0.0249(0.01769)	0.00-0.07	1.78	0.0356(0.01918)	0.01-0.10
64 Homicide. assault (X85-Y09)	0.10	0.0023(0.01508)	0.00-0.10	0.00	0.00(0.00)	0.00-0.00
65 Events of undetermined intent (Y10-Y34)	0.50	0.0098(0.00424)	0.00-0.02	0.42	0.0084(0.00422)	0.00-0.02

2006. For instance, the prefecture of Ioannina, which had almost a zero rate in 2001 had a very high rate in 2006. The prefecture of Evrytania has the same “behavior” in both year - periods (Maps 9 to 12). Furthermore, the results of the accidents’ mortality rates are worth mentioned. Accidents for men presented a significant increase from 2001 to 2006. On the other hand, the same mortality rates for women remained at the same levels (high rates

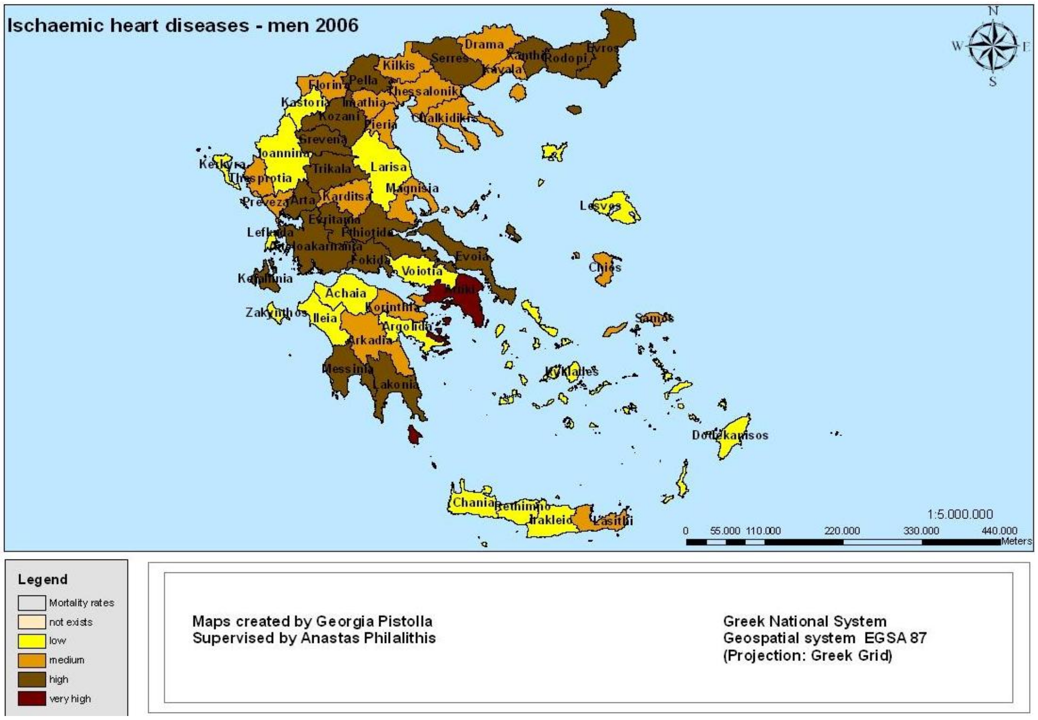
Table 2. Number of prefectures that belong to each cluster, by CAM* , in Greece in 2001 and 2006.

Group cases	Men 2001		Men 2006		Women 2001		Women 2006	
	Cluster[a]	Cases	Cluster[a]	Cases	Cluster[a]	Cases	Cluster	Cases
Ischaemic heart disease	Low	8	Low	16	Low	1	Low	9
	Medium	26	Medium	17	Medium	30	Medium	36
	High	16	High	17	High	19	High	5
	Very high	1	Very high	1	Very high	1	Very high	1
Celebrovascular diseases	Low	9	Low	2	Low	2	Low	7
	Medium	17	Medium	22	Medium	26	Medium	20
	High	19	High	26	High	22	High	23
	Very high	6	Very high	1	Very high	1	Very high	1
Neoplasm of larynx and trachea/ bronchus/ lung	Low	32	Low	4	Low	7	Low	11
	Medium	15	Medium	31	Medium	35	Medium	34
	High	1	High	15	High	6	High	5
	Very high	1	Very high	1	Very high	1	Very high	1
Transport accidents	Low	2	Low	11	Low	12	Low	2
	Medium	18	Medium	22	Medium	25	Medium	19
	High	27	High	17	High	9	High	20
	Very high	2	Very high	1	Very high	2	Very high	5

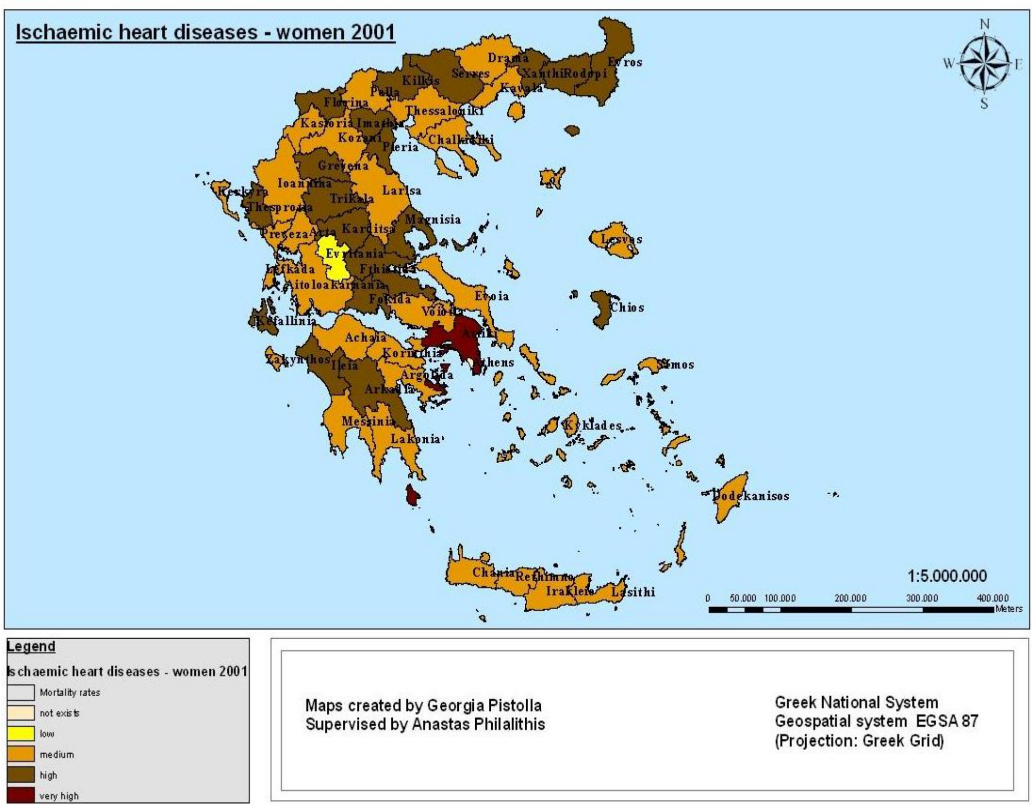
*CAM: Cluster Analysis Method; [a] Low, medium, high and very high clusters refer to a group of Prefectures that present similar patterns, for instance high mortality rates in 2001 or medium mortality rates in 2006.



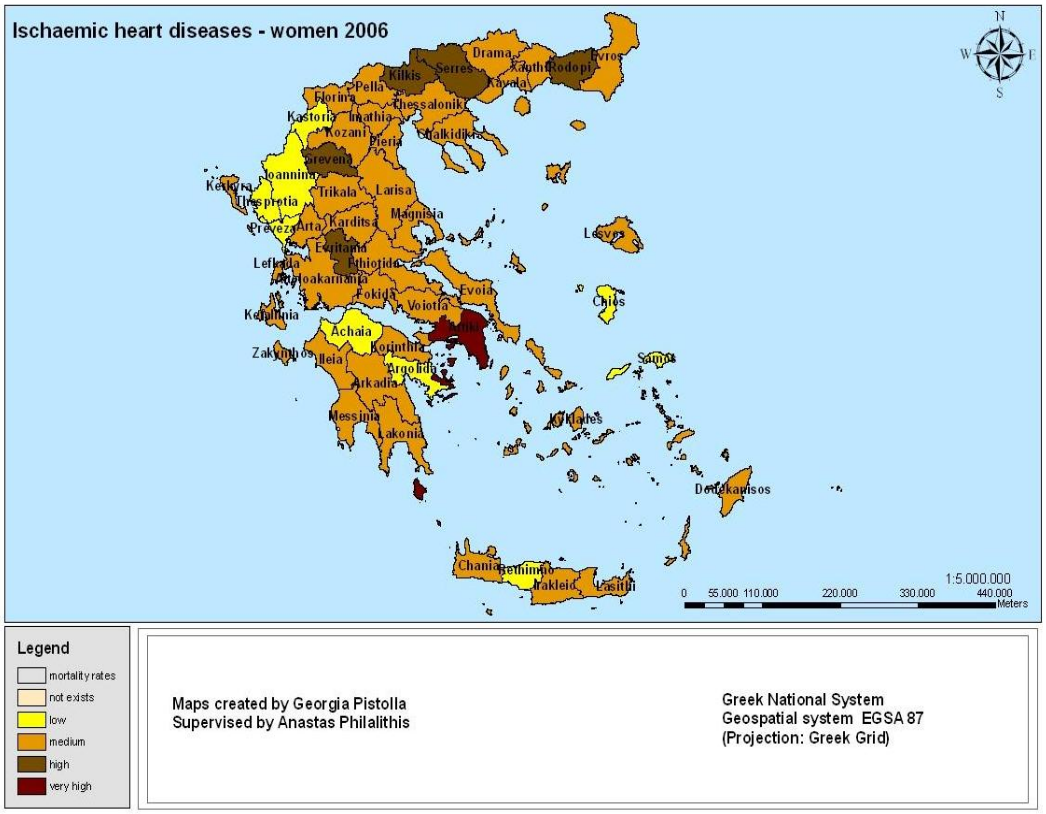
Map 1. Map of Greece showing ischaemic heart diseases as cause of death per prefecture among men in 2001.



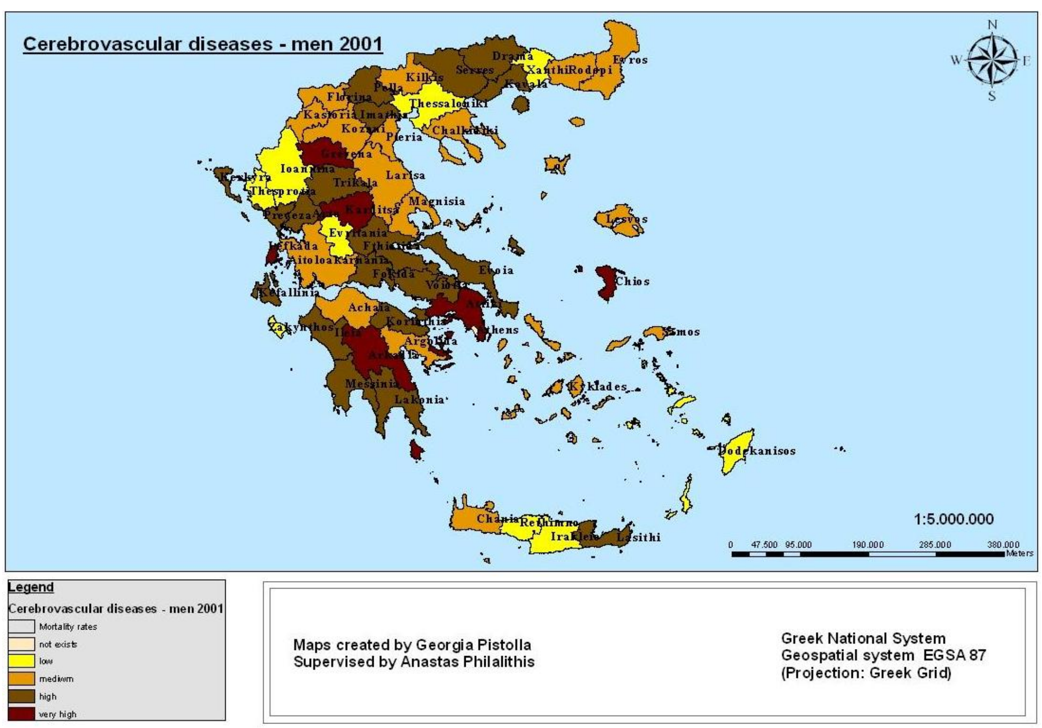
Map 2. Map of Greece showing ischaemic heart diseases as cause of death per prefecture among men in 2006.



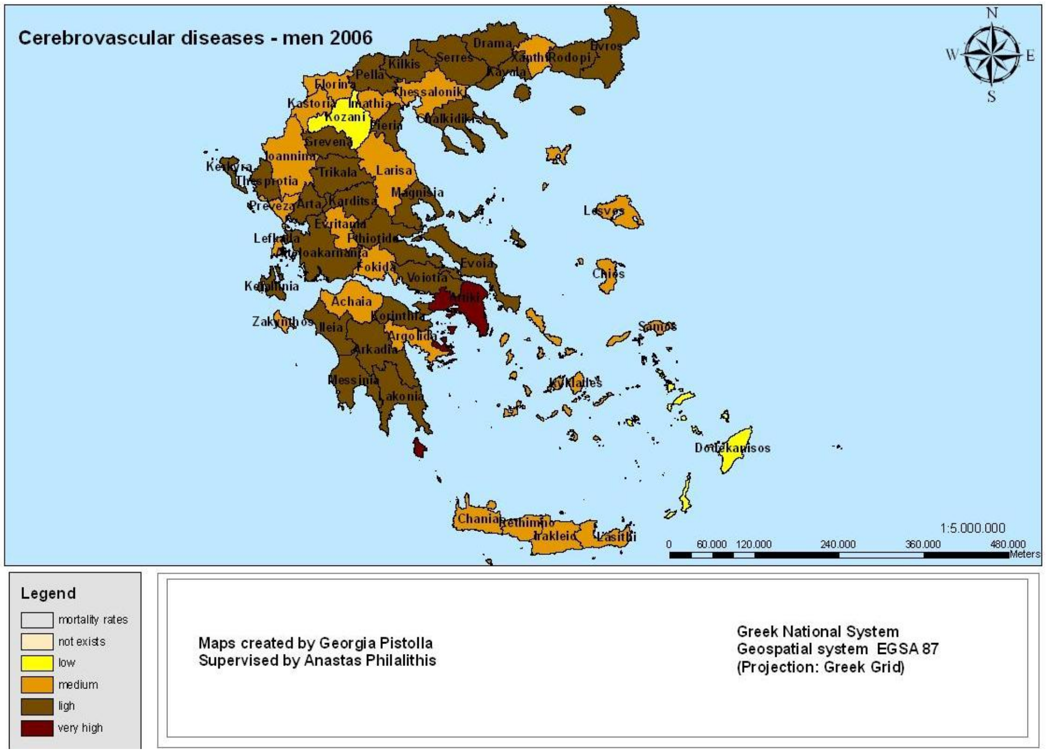
Map 3. Map of Greece showing ischaemic heart diseases as cause of death per prefecture among women in 2001.



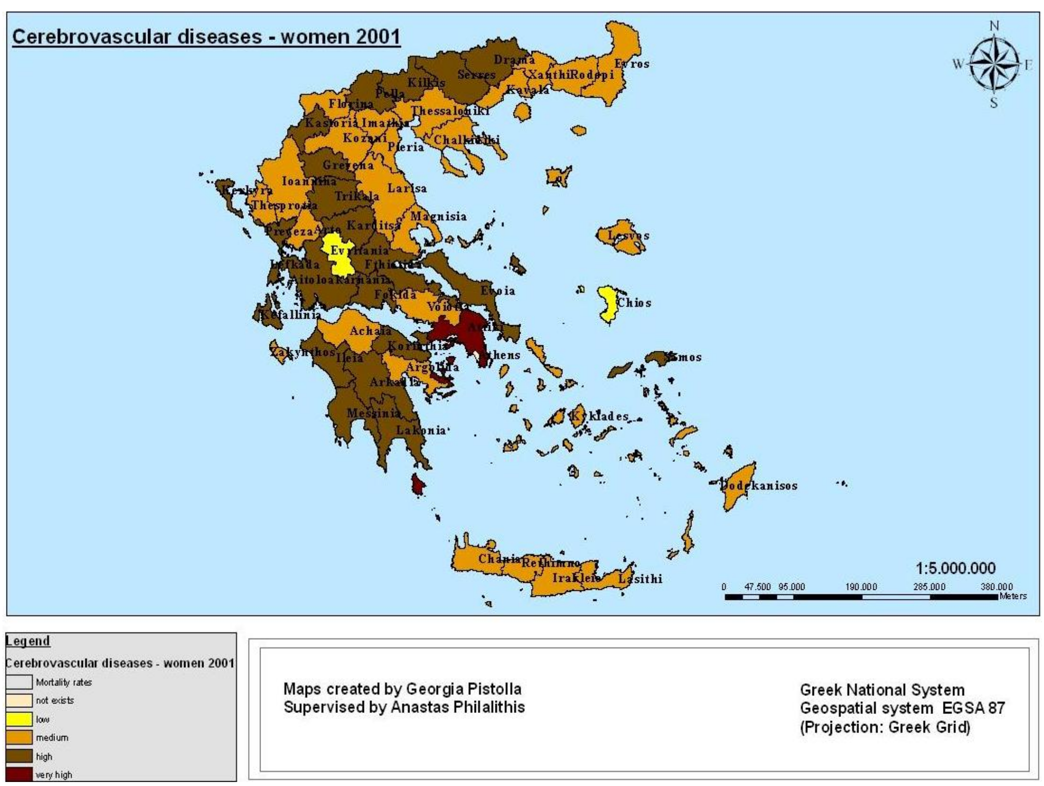
Map 4. Map of Greece showing ischaemic heart diseases as cause of death per prefecture among women in 2006.



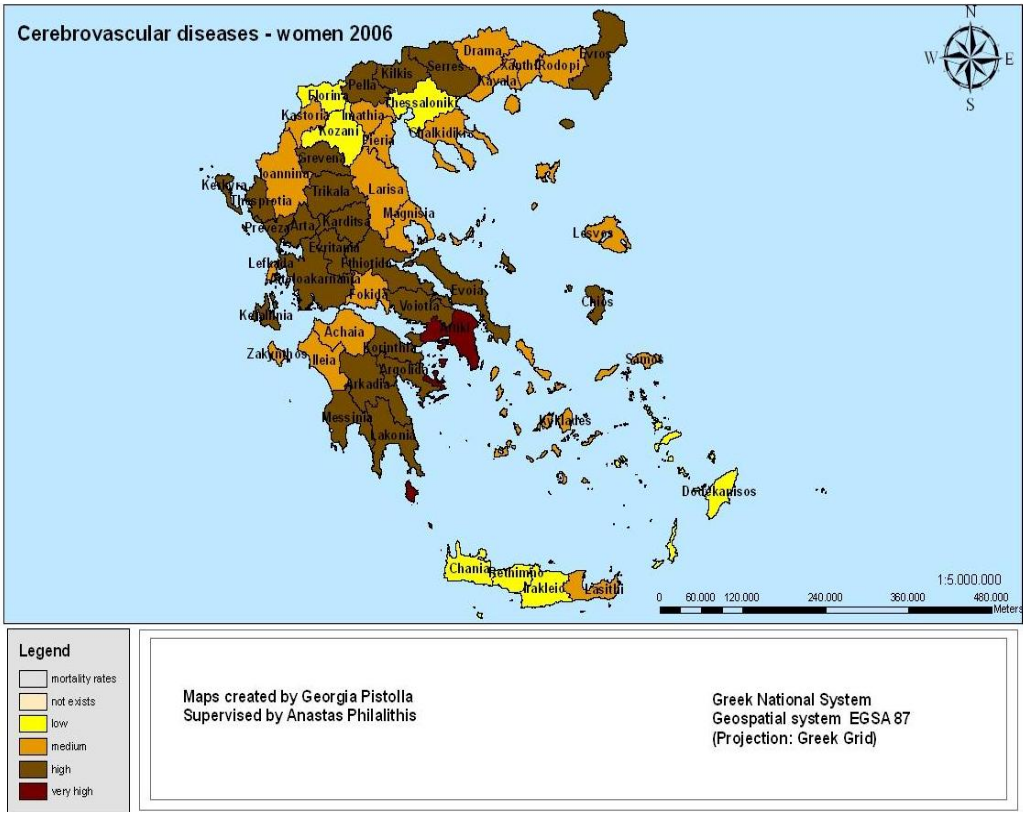
Map 5. Map of Greece showing cerebrovascular diseases as cause of death per prefecture among men in 2001.



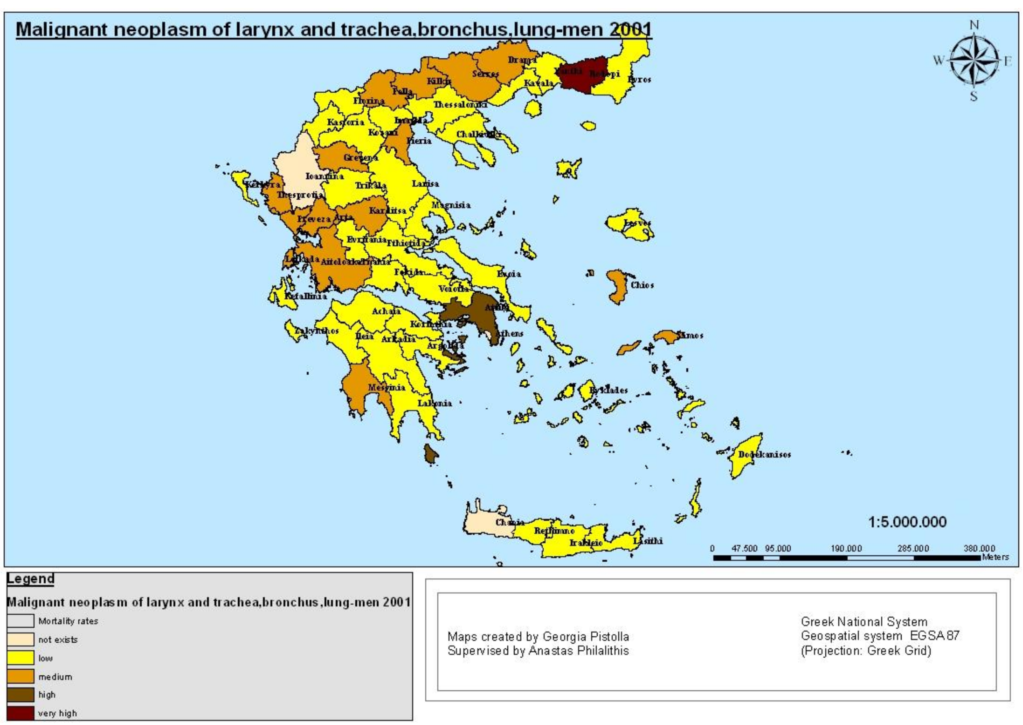
Map 6. Map of Greece showing cerebrovascular diseases as cause of death per prefecture among men in 2006.



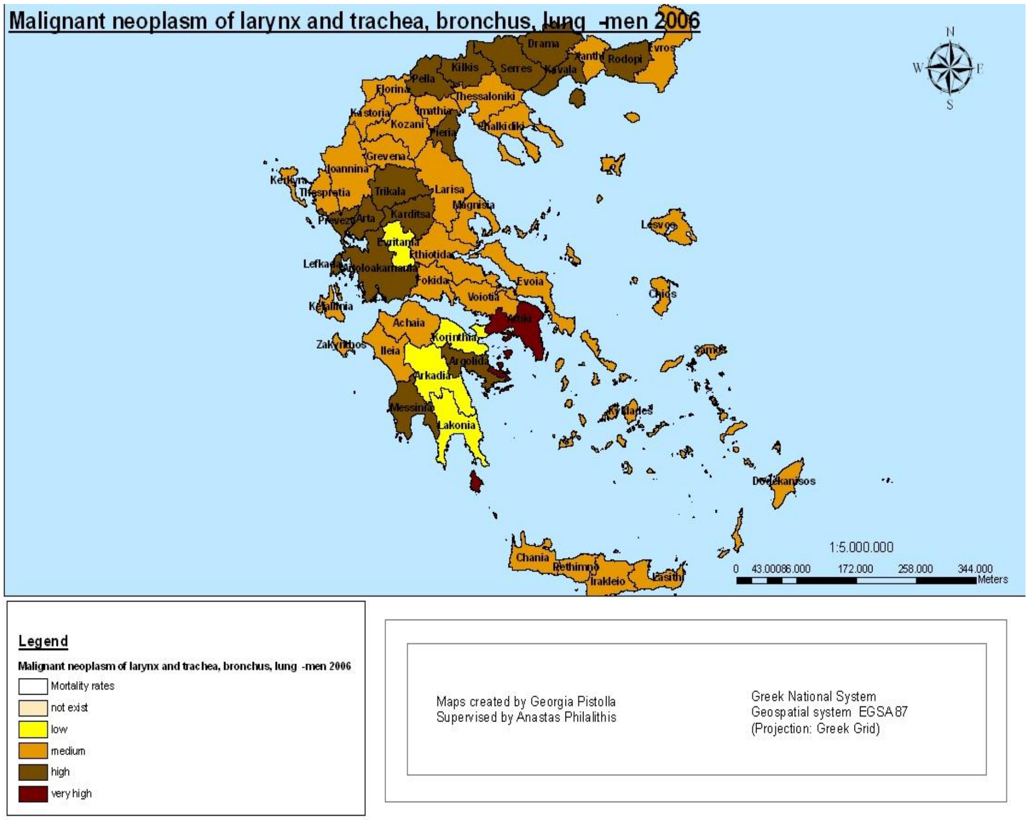
Map 7. Map of Greece showing cerebrovascular diseases as cause of death per prefecture among women in 2001.



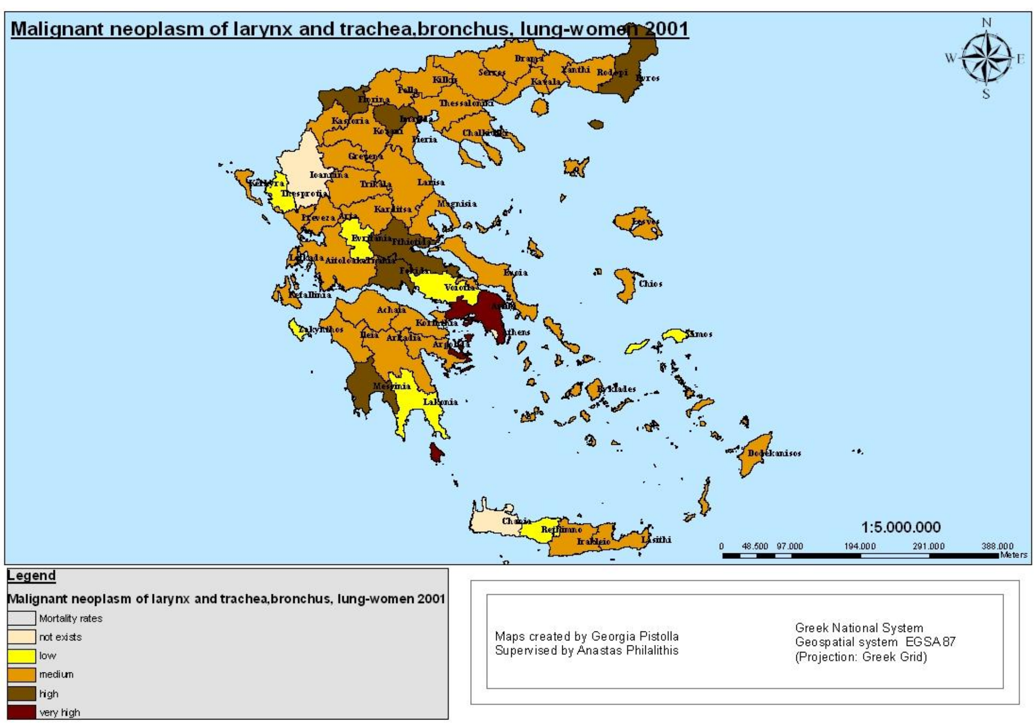
Map 8. Map of Greece showing cerebrovascular diseases as cause of death per prefecture among women in 2006.



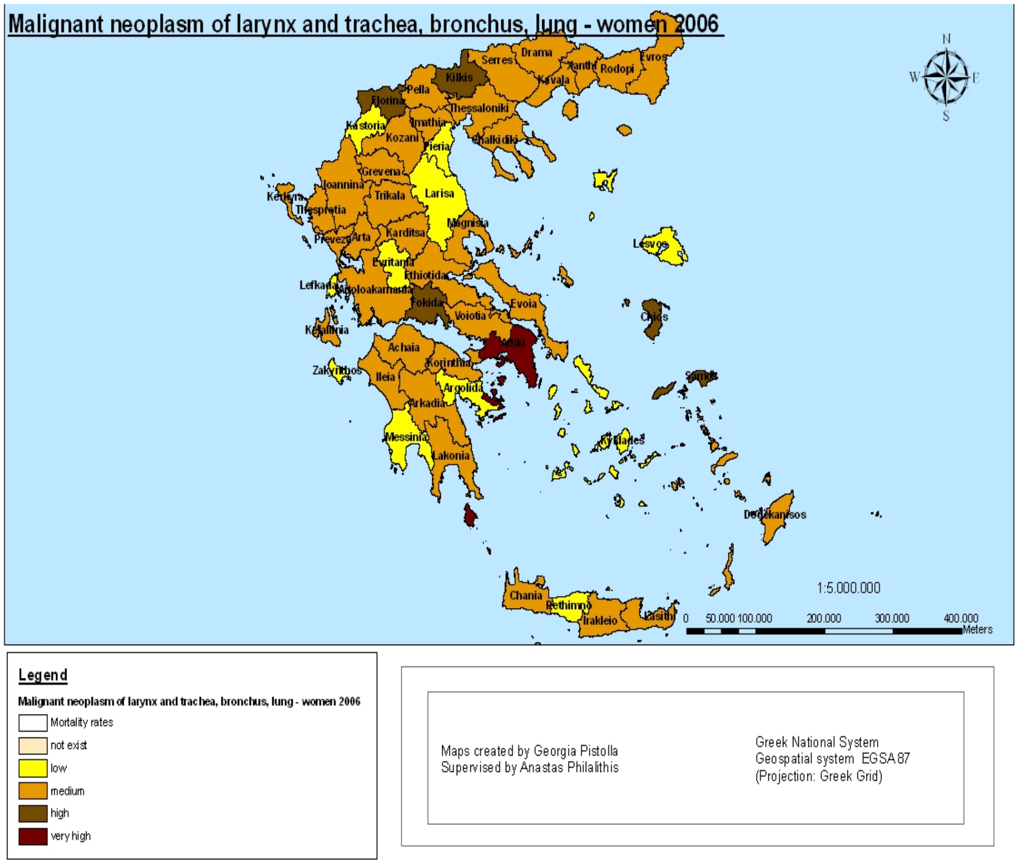
Map 9. Map of Greece showing malignant neoplasms of larynx and trachea, bronchus, lung as cause of death per prefecture among men in 2001.



Map 10. Map of Greece showing malignant neoplasms of larynx and trachea, bronchus, lung as cause of death per prefecture among men in 2006.



Map 11. Map of Greece showing malignant neoplasms of larynx and trachea, bronchus, lung as cause of death per prefecture among women in 2001.



Map 12. Map of Greece showing malignant neoplasms of larynx and trachea, bronchus, lung as cause of death per prefecture among women in 2006.

for both time - periods). Finally, in 2006 the rates appeared to be almost zero in most of the perfectures (Maps 13 to 16).

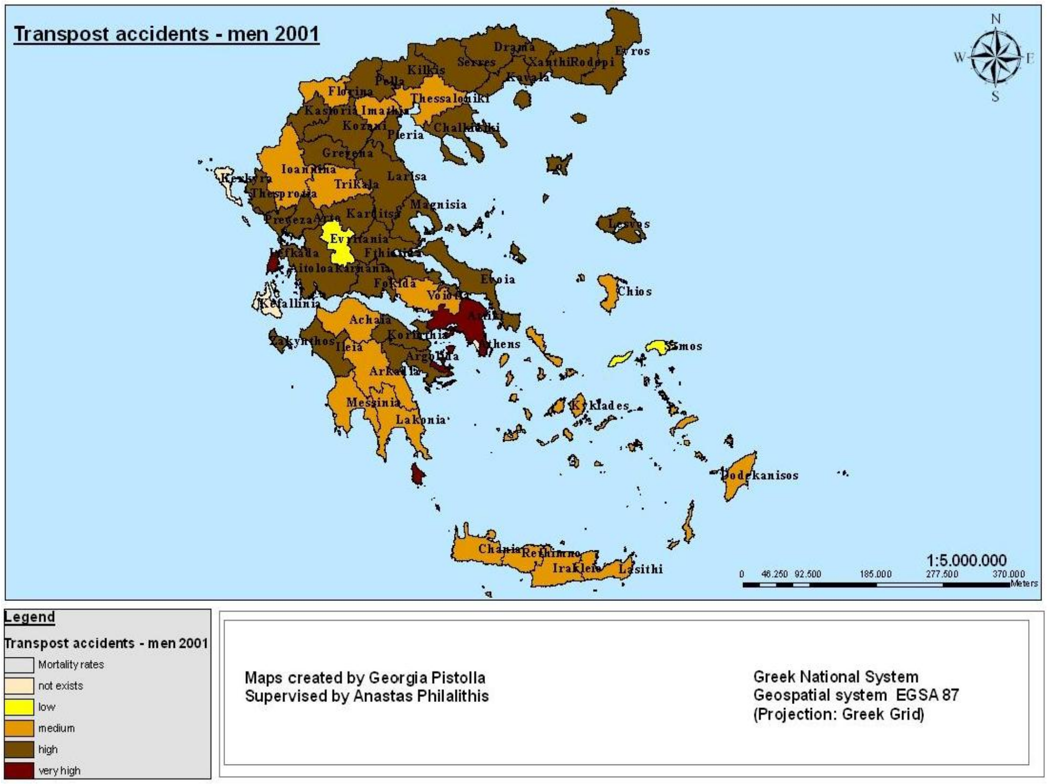
As far as variations between genders are concerned, they appeared to be grouped in the same clusters for both men and women, in most causes of mortality in 2001 and 2006. The only significant difference was observed in malignant neoplasms of lip, oral cavity and pharynx among men in 2001 and in transport accidents among women in 2001. These variations were not very high but they still were significant and caused the creation of another clustering which differentiated men from women. Similar trends also exist in the prefectures of Achaia and Chalkidiki but at a lower degree. Five causes of mortality were estimated to present significant differences between the mortality rates from 2001 to 2006. These were malignant neoplasms, malignant neoplasm of pancreas, other heart diseases, diseases of the respiratory system and cerebrovascular ones. These five causes of mortality were turned into one new clustered map per prefecture that follows as multi-clustered map as was mentioned in the “Materials and Methods” of this paper (Diagram 1 and Map 17). Each color symbolizes all five causes of mortality whereas it indicates a different group of prefectures which are similar as far as mortality patterns are concerned.

DISCUSSION

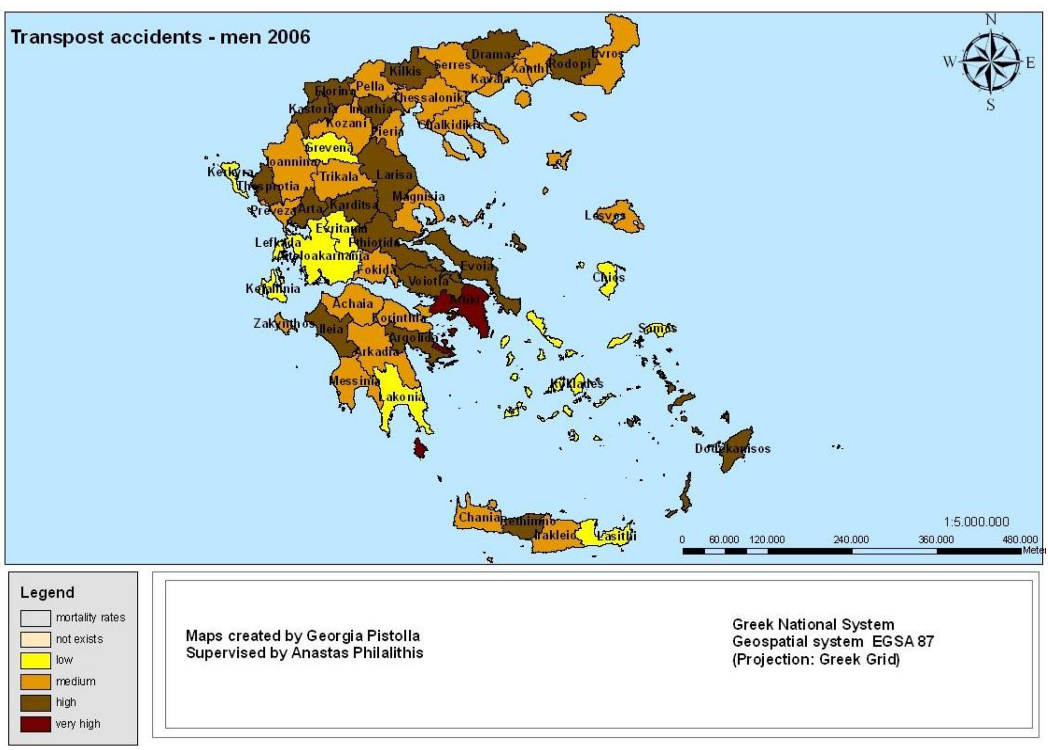
Main findings

In the present study the observed variations for the same cause of mortality, gender and year are significant and considered to be one of the most important findings. Prefectures “move” from one cluster to another per gender or year, either to a cluster with higher mortality rate or to a cluster with a lower rate. Such an example is the prefecture of Evritania that in 2001 was clustered in the low group, for all causes, while in 2006 it “moved” to the high or very high rates. Additionally, differences were also observed in the groups between genders for the same year and for the same cause of mortality (eg. accidents). Women present higher rates for most of causes of mortality and are grouped in the highest clusters.

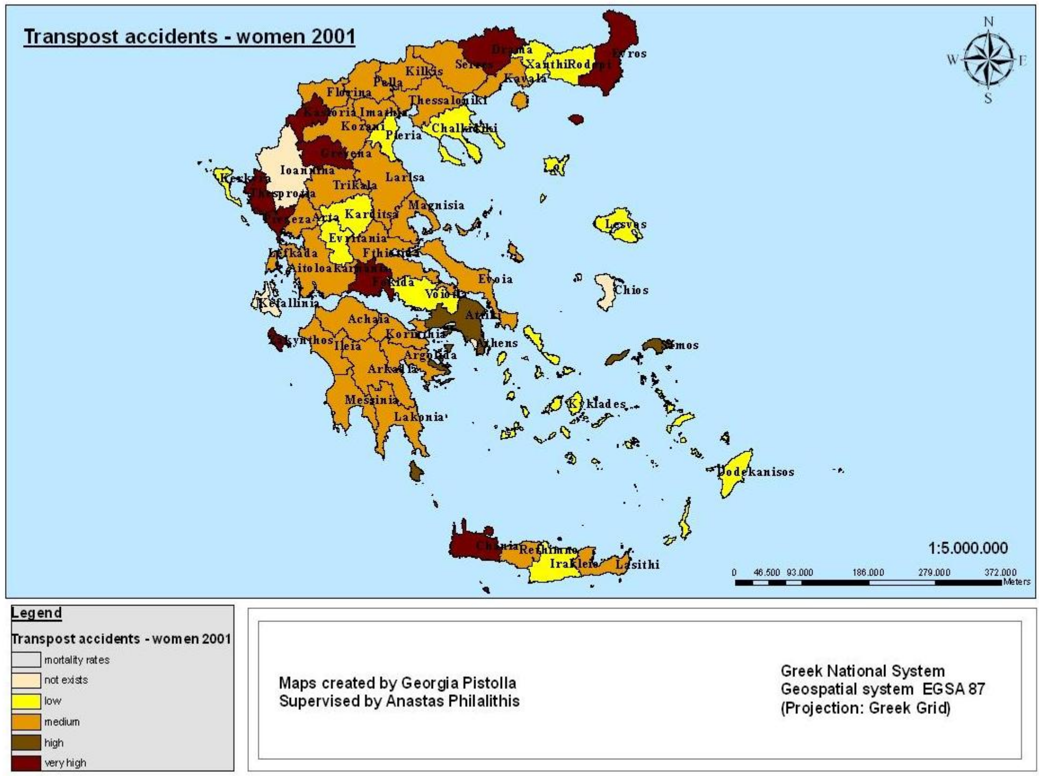
Another important finding of the present study was derived from the multi-cluster analysis. The main time



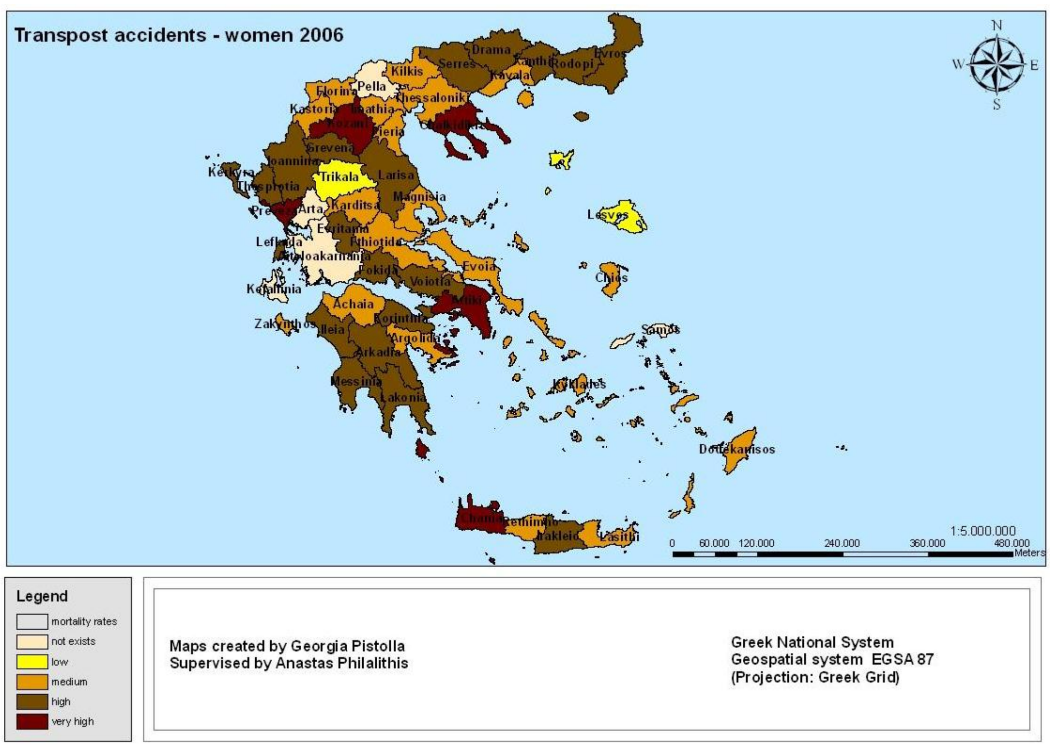
Map 13. Map of Greece showing transport accidents as cause of death per prefecture among men in 2001.



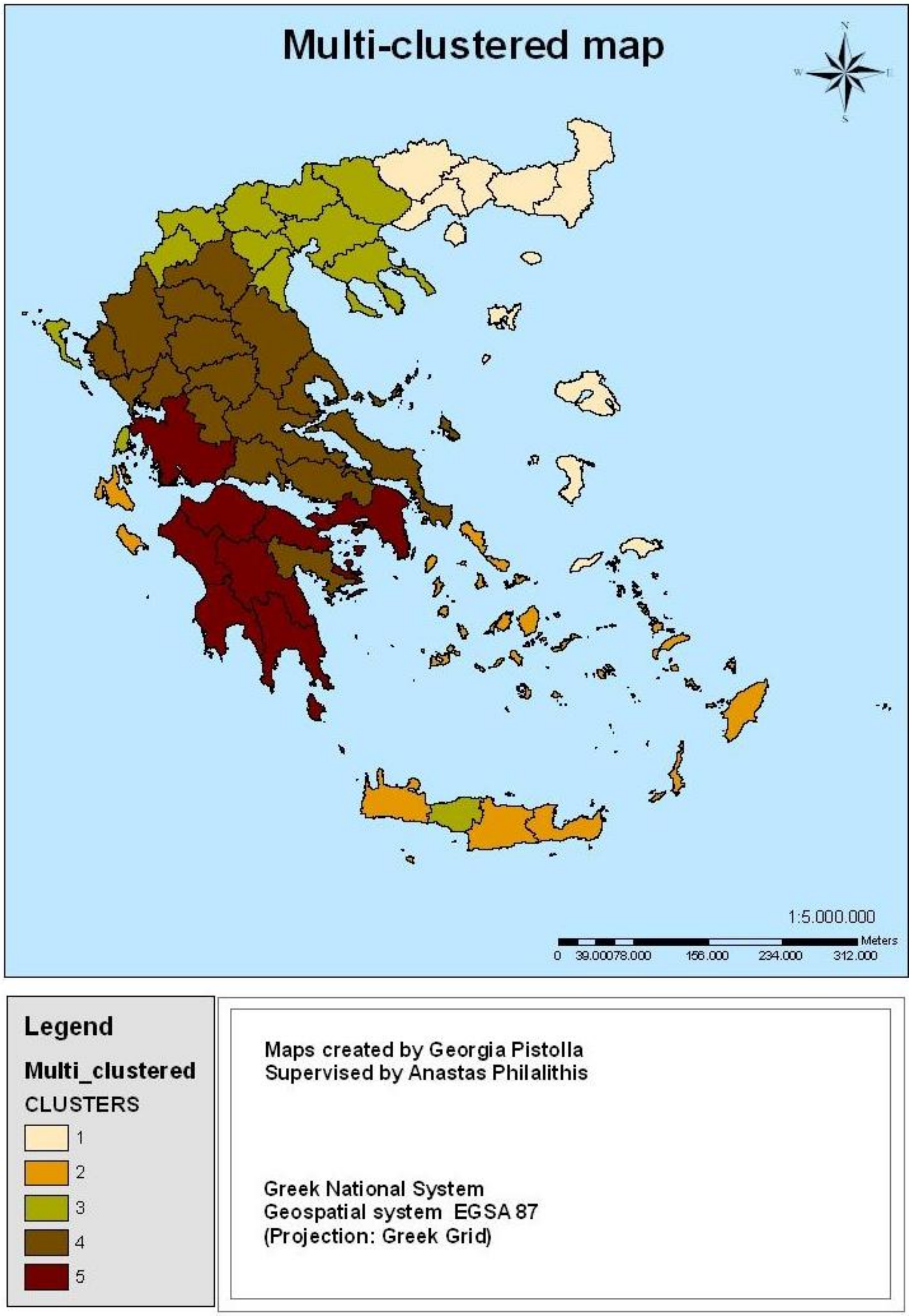
Map 14. Map of Greece showing transport accidents as cause of death per prefecture among men in 2006.



Map 15. Map of Greece showing transport accidents as cause of death per prefecture among women in 2001.



Map 16. Map of Greece showing transport accidents as cause of death per prefecture among women in 2006.



Map 17. Multi-clustered map (five groups).

variations from 2001 to 2006 are observed in five causes of mortality (malignant neoplasms, malignant neoplasm of pancreas, other heart diseases, diseases of the respiratory system and cerebrovascular diseases) which have been mentioned in the "Results" of this research, while the clusters grouped together prefectures from different geographical regions of the country. This highlights the importance of geographical analyses in the study of mortality rates. Overall, the results presented in this paper show that in six years (from 2001 to 2006) significant changes occurred in the standardised mortality rates and confirms the fact that we should analyse the data as closer in their real “environment” as possible (Pistolla et al., 2011).

Other studies in the literature

To the best of our knowledge, it is the first study that has been carried out for all causes of mortality for both genders at a geographical level. However, there is a wide variety of studies that select one or two significant causes of mortality and follow similar methodological patterns, or research questions. Similarly, to the present findings, Bilancia and Fedespina (2009) support that cancer epidemiology is a clustered phenomenon in the province of Lecce in Italy and that high risk areas as well as low risk areas exist in the same province. They finally suggest the use of spatial clustering methods for such phenomena. Another study of Boscoe et al. (2004) presents the use of spatial applications in epidemiology and cancer research, highlighting the need of local strategies. Finally, several studies in mortality or burden of disease have already used clustering methods such as CAM (Merlo et al., 2004; Buntix et al., 2003).

Interpretation

Greece is a country which is characterized by high diversity in cultural, social, economical, environmental, health and other standards. There is a total of 51 prefectures in the country that their population ranges from 19,518 to 1,089,311. Although the population of Greece is not equally distributed among the various prefectures, it is a fact that prefectures are the appropriate geographical unit for carrying out this analysis for the following reasons: First of all, it allows the researchers to study the regional differences and variations more accurately rather than performing the same analysis at the NUTS 2 level which there were a total of 13 NUTS regions in Greece (European Commission, Eurostat, 2008-2011). Secondly, the possibility for errors is significantly reduced compared to performing an analysis per municipality, since the data are more accurate. Finally, the number of observations is considered to be a reliable sample size for an analysis with statistical techniques such as the CAM.

Furthermore, several differences among variables were detected. Variations were obvious, either between the two time periods or gender, and for some mostly among the prefectures. For instance, one of the most significant causes of mortality, ischemic heart disease, presented similar patterns in mortality patterns through time and prefectures for men whereas women presented strong variations among prefectures even for the same year.

Perspectives

The decision makers responsible for health policies and planning should be influenced by such findings and construct policies and preventive measures based on the results. There is a high need of formulating different policies for each cause of mortality per prefecture, according to their clustering. Specifically, prefectures of high or very high clusters should be considered to be areas of high risk and direct need of intensive health care and health management. Even, differences between genders and causes of mortality influence in a different way the public health standards. It is therefore concluded that generalised and universal health policies may not be the most effective in our time if they are not community oriented.

Strengths and limitations

The strengths of this study could be divided in two parts; first, the methodology and use of CAM. Cluster methodology was applied to study and compare the mortality rates by gender and cause of death in the prefectures of Greece. For the years 2001 and 2006 it was obvious that similarities in mortality patterns were not related to the geographical proximity of the prefectures. CAM was selected as the best methods to contribute to the exhibition of common “behavior” patterns of mortality (Osmar, 2010). It may be useful in explaining, how health services are delivered at a local and individual level. Therefore, it may be valuable in health services planning, in a way that has local relevance. Clustering algorithms have already been applied in many fields, for instance marketing for finding groups of customers with similar behavior, given a large database of customer data concerning their properties and past buying records. Also, it is applied in biology for the classification of plants and animals, according to their features. Furthermore, they are useful for book ordering in libraries, for identification of groups of motor insurance policy holders with a high average claim cost or identifying frauds in iinsurance. Other fields of application are city planning (to identify groups of houses according to their house type, value and geographical location). Also, they are used for the observation of earthquake epicenters and identification of dangerous zones and finally, in internet pages (www.) for document classification and clustering web-log data to discover groups of similar access patterns (Osmar, 2010; Bilancia and Fedespina, 2009).

Contrary to the above strengths of the study, there are also several limitations that could be addressed in a next study. Specifically, different type of rates could also be tested, such as prevalence or incidence of a cause of a disease. Another, interesting point is that these findings indicate the existence of several factors that affect and differentiate the mortality patterns among geographical units and time. Such factors could be demographic (e.g. age), socioeconomic, environmental or physiological and it would be really useful if correlated with these mortality rates spatially. This could finally be an incentive for further study. Finally, time trends could be analyzed for a longer time period, such as the last four decades, in order to capture the dynamics of each cause of mortality.

Conclusion

The present study captured the time and spatial dynamics of mortality of all causes in Greece for two different time periods (2001 and 2006). It used reliable methodology which minimizes any potential bias of the data and managed to make assumptions about the future health planning. Policies and strategies that should be a major priority are related to the major cause of mortality and particularly the prefectures that they present the clusters with the highest values. Hence, this study supports that local health strategies should be applied and that the use of CAM analysis could be a useful tool.

REFERENCES

Bilancia M, Fedespina A (2009). Geographical clustering of lung cancer in the province of Lecce, Italy: 1992-2001. Int. J. Health Geogr. 8:40.

Boscoe FP, Ward MH, Reynolds P (2004). Current practices in spatial analysis of cancer data: Data characteristics and data sources for geographic studies of cancer. Int. J. Health Geogr. pp. 20-28.

Buntinx F, Geys H, Lousbergh D, Broeders G, Cloes E, Dhollander DL, Op De Beeck, Vanden Brande J, Van Waes A , Molenberghs G (2003). Geographical differences in cancer incidence in the Belgian province of Limburg. Eur J. Cancer 39:2060-2065.

Cassetti T, La Rosa F, Rossi L, D'Alò D, Stracci F (2008). Cancer incidence in men: A cluster analysis of spatial patterns. BMC Cancer 8:344.

Christopher JLM , Lopez AD (1997). Mortality by cause for eight regions of the world: Global Burden of Disease Study. The Lancet 349:1269-1276.

Dafermos V (2005). Social Statistics with SPSS. Ziti,Thessaloniki (in Greek).

Douven W, Schilter HJ (1995). Spatial Analysis in Health. Research in de Lepper pp. 117-133.

ESRI (2000-2004). ARCGIS 9-USING ARC MAP.

European Commission, Eurostat (2008-2011). NUTS 2 level. Available at: http://ec.europa.eu/eurostat/ramon/nuts/codelist_en.cfm?list=nuts (Accessed: 10 August 2008).

European Commission, Eurostat (2008-2011). E.U Standards for classification. Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home (Accessed: 10 August 2008).

Havard S, Deguen S, Zmirou-Navier D, Schillinger C, Bard D (2009). Traffic-Related Air Pollution and Socioeconomic Status: A Spatial Autocorrelation Study to Assess Environmental Equity on a Small-Area Scale. Epidemiology 20(2):223-230.

Health Planning Unit, Faculty of Medicine, University of Crete. Maps of Mortality Mapping Project. February 2011. Available at: http://healthplanning.uoc.gr/mps/maps.zip.

Hellenic Statistical Authority (2008). Number of deaths. Available at: http://www.statistics.gr/portal/page/portal/ESYE (Accessed: 10/10/2010).

MacQueen JB (1967). Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, University of California Press 1:281-297.

Merlo J, Chaix B, Yang M, Lynch J, Råstam L (2004). A brief conceptual tutorial of multilevel analysis in social epidemiology: Linking the statistical concept of clustering to the idea of contextual phenomenon. J. Epidemiol. Commun. Health 59:443-449.

Osmar RZ (2010). Principles of Knowledge Discovery in Databases, Data Clustering: Chapter 8. Available at: http://www.cs.ualberta.ca/~zaiane/courses/cmput690/slides/Chapter8/index.html (Accessed: 12 August 2010).

Pagano M, Gauvreau K (1996). Principles of Biostatistics. Harvard School of Public Health. Buhbury Press.

Peto R, Boreham J, Lopez AD, Heath MTC (1992). Mortality from tobacco in developed countries: indirect estimation from national vital statistics. The Lancet 339:1268-1278.

Pistolla G, Prastakos P, Vassilaki M, Philalithis A (2010). Spatial-Mathematic methods for analysis of indicators of mortality. IJAEST 1(2):135-146.

WHO (2008). World Health Organization database, Available at: http://www.who.int/library/databases/en/, Accessed in: 28/01/2012.