Spatial epidemiology of Amyotrophic Lateral Sclerosis in Piedmont and Aosta Valley, Italy: a population-based cluster analysis

Spatial epidemiology of Amyotrophic Lateral Sclerosis in Piedmont and Aosta Valley, Italy: a population-based cluster analysis.

Vasta R.[1] (MD), Calvo A.[1,2] (MD, PhD), Moglia C.[1,2] (MD, PhD), Cammarosano S.[1] (MD, PhD), Manera U.[1] (MD), Canosa A.[1] (MD, PhD), D'Ovidio[1] F. (PhD), Mazzini L.[3] (MD), Chiò A[1,2], (MD, FAAN).

[1] ALS Center, 'Rita Levi Montalcini' Department of Neuroscience, University of Turin, Turin, Italy.
[2] Azienda Ospedaliero Universitaria Città della Salute e della Scienza, Turin, Italy.
[3] ALS Center, Department of Neurology, Azienda Ospedaliero Universitaria Maggiore della Carità, Novara, Italy.

Corresponding Author: Rosario Vasta, ALS Center, 'Rita Levi Montalcini' Department of Neuroscience, University of Turin, email: sarovasta@gmail.com, phone no:+39 0116335439

Type of article: original article.

Abstract word count: 206. **Text word count**: 3376. **Figures:** 1. Tables: 2

Running title: Spatial epidemiology of ALS in Piedmont and Aosta Valley.

Keywords: Amyotrophic Lateral Sclerosis, Motor Neuron Diseases, spatial epidemiology.

Background. The analysis of the spatial distribution of cases could give important cues on putative environmental causes of a disease. We aimed to perform a spatial analysis of a 20-years period ALS cohort from the Piedmont and Aosta Valley ALS register (PARALS).

Methods. Address at the moment of diagnosis was considered for each ALS case. Municipalities' and census divisions' resident populations during the 1995-2014 period were obtained. Cluster analysis was performed adopting both Moran's Index and the Kulldorff's spatial scan statistic.

Results. A total of 2702 ALS patients were identified. Address was retrieved for 2671 (99%) patients. Moran's Index was -0.01 (*p-value* = 0.83) thus revealing no clusters. SatScan identified no statistically significant clusters. When census divisions were considered, Moran's Index resulted 0.13 (*p-value* = 0.45); SatScan revealed one statistically significant small cluster in the province of Alessandria. Here, 0.0099 cases were expected and 3 cases were observed (relative risk 304.60; 95% CI 109.83 - 845.88, *p-value* 0.03).

Discussion. Our study showed a substantial homogeneous distribution of ALS cases in Piedmont and Aosta Valley. The population-based setting and the adoption of proper statistical analyses strengthen the validity of our results. Such finding further suggests the involvement of multiple environmental and genetic factors in the ALS pathogenesis.

Introduction

Amyotrophic Lateral Sclerosis (ALS) is a relentlessly neurodegenerative disease characterized by the degeneration of motor neurons(1).

Beyond a 10% of cases known to be of genetic aetiology, ALS is thought to be caused by the interaction of genetic and environmental factors. Despite all efforts, research on such factors yielded to inconsistent results and etiology of most ALS cases remains unknown so far(2).

Spatial epidemiology could give important clues when searching for putative environmental causes of a disease. Indeed, analysis of the spatial distribution of cases could demonstrate clusters of higher or lower incidence. Subsequently, hypotheses about exposures potentially responsible for such clusters could be made(3).

In 2013, a Bayesian spatial analysis of the 1995-2004 ALS incident cases in Piedmont revealed an excess of risk for ALS in the area of Cuneo, Alessandria and Vercelli (standardized incidence ratio > 1.2)(4). However, Bayesian analysis aims to detect spatial variations of risk but it is not able to estimate spatial clustering(3).

Using a 20-years population-based cohort of ALS incident cases, we aimed to perform a cluster analysis of ALS incident cases in a wider area, including both Piedmont and Aosta Valley regions.

Materials and methods

Study population, study period and study area. Data from the Piedmont and Aosta Valley Register (PARALS) were used. Methods of the PARALS register have been exhaustively described elsewhere(5).

The study was conducted in Piedmont and Aosta Valley. In order to investigate the "modifiable areal unit problem"(6), the analysis was conducted at both municipalities and census divisions levels.

Piedmont covers an area of about 25.400 km^2 divided into 1206 communities. According to the official 2001 Italian Census(7), it includes 32.988 census divisions, with a total population of 4.214.677 residents (2.034.161 males and 2.180.516 females). Aosta Valley covers an area of about 3.200 km^2 divided into 74 communities. It includes 1860 census divisions, with a total population of 119.548 residents (58.563 males and 60.985 females)(7).

Geocoding. Municipalities' and census divisions' resident populations were obtained from the National Institute of Statistics (Istituto Nazionale di Statistica, ISTAT)(7).

Address at the moment of the diagnosis was considered for each ALS case. When census divisions were considered, addresses were geocoded using the *MMQGIS* plugin of QGIS(8).

Cluster analysis. A cluster is defined as a geographically and/or temporally bounded group of occurrences of sufficient size and concentration to be unlikely to have occurred

by chance(9). Cluster analysis was assessed using both Moran's Index (Moran's I)(10) and the spatial Kulldorff's spatial scan statistic(11).

Moran's I studies spatial correlation by quantifying the similarity between neighbor areas (3). Briefly, contiguous divisions are identified constructing a spatial weights matrix by which adjacent areas are given a greater weight than those that are distant. Queen contiguity method (i.e. areas were considered contiguous if they share either a border or a corner) was adopted. After that, the age and sex-standardized incidence ratio (SIR) of each division is correlated with the mean SIR of its neighbor divisions. A Moran's I of 0 indicates the null hypothesis of no clustering; a value of +1 indicates positive spatial correlation (clustering) while a value of -1 indicates negative spatial correlation (dispersion).

Kulldorff's analysis scans for clusters of any size by imposing for each point circular windows of varying radii, from zero to an arbitrary pre-defined maximum of population included in the window. We set this limit at 20% of the population at risk. For each circular window, the null hypothesis is that the incidence rate is the same as in all other windows. The alternative hypothesis is that inside the window the incidence rate is higher or lower than in the remainder part of the study area. The statistic test adopted is the likelihood ratio. *P-value* was set at 0.05%. The Kulldorff's analysis was age- and sex-adjusted.

The Monte Carlo simulation was used to assess the significance of both Moran's I and Kulldorff's analysis(3); Bonferroni correction was adopted for Moran's I, whether SatScan software already takes into account it.

Statistical analysis. R 3.4 software(12) was used to perform basic statistics and data preparation for both Moran's I and SatScan analysis.

Moran's I was calculated using the *moran.test* function of the R's *spdep* package. A neighboring matrix was determined using the *poly2nb* and *nb2listw* functions of the R's *spdep* package. The spatial Kulldorff's spatial scan statistic was implemented using the SatScan software, version 9.4.4(13).

Ethical approval. The study was approved by the Città della Salute e della Scienza of Torino ethical committee. Databases were treated according to the Italian privacy regulations.

Results

Descriptive statistics. During the 20-years study period, a total of 2702 Piedmont and Aosta Valley residents have been diagnosed with ALS. Demographical and clinical characteristics of ALS cases included in the register have been described elsewhere(5).

Address was retrieved for 2671 (98.9%) patients, 1447 males and 1224 females (male:female ratio = 1.2). Age at onset was 65.7±11.2 years (65.2±11.1 years for males; 66.2±11.2 years for females). The presentation was bulbar in 962 patients (36%) and spinal in 1709 patients (64%).

The mean annual crude incidence rate for the 1995-2014 period was 3.0 cases per 100.000 person-years.

Cluster analysis. Moran's I resulted -0.02 (*p-value* = 0.83), thus revealing no clusters. SatScan revealed 6 clusters (1 of higher risk and 5 of lower risk). None of these reached the statistical significance (Table 1).

When census divisions were considered, Moran's I resulted slightly positive (I = 0.13) but not significant (*p-value* = 0.45), whether SatScan revealed 16 clusters (11 of higher risk and 5 of lower risk), only one of which reached the statistical significance (Table 2). It includes only one census division, in the community of Acqui Terme, in the province of Alessandria. Here, 0.0099 cases were expected and 3 cases were observed (RR 304.60; 95% CI 109.83 - 845.88, *p-value* = 0.03) (Figure 1). Two of these 3 cases were spouses.

Discussion

The study of the spatial distribution of ALS cases could be a useful tool to pinpoint potential environmental risk factors of the disease. We analyzed the spatial epidemiology of ALS cases in Piedmont and Aosta Valley, over a 20-years period. Using two dedicated spatial statistics(10,11), a substantial homogenous distribution of cases throughout the study area was found.

When census divisions were considered, a small cluster was identified in the community of Acqui Terme. Whilst *p-value* resulted 0.03, the exiguous number of observed cases makes chance a reasonable explanation. Moreover, this cluster includes a conjugal ALS case for which both common environmental and genetic factors have been previously excluded(14).

These results suggest that either an environmental factor is homogenously distributed or that multiple factors, including genetic, may explain the homogenous distribution of cases.

Our results did not confirm those from a previous study conducted in Piedmont and from a previous national study(4). Indeed, using the 1995-2004 PARALS cohort and adopting a Bayesian approach, a higher incidence (SIR>1.2) was found out in the province of Alessandria, Vercelli and Cuneo. Nonetheless, Bayesian spatial analysis aims to study the distribution of risk and is not able to detect clusters(3). Uccelli et al. analyzed the Italian national ALS mortality rate during the 1980-2001 period. Adopting the spatial scan statistics, 16 statistically significant ALS clusters were identified throughout the national area, three of which in Piedmont (namely in the municipalities of Trino, Tronzano Vercellese and Briga Novarese)(15). None of these clusters has been revealed by our

analysis. Nonetheless, it should be noted that patients are likely to move after ALS diagnosis, to a nursing home or back to family, thus address at death could not be a suitable indicator of previous exposures. Indeed, a subsequent study using both hospital discharge charts and death certificates did not confirmed an excess of ALS incidence in the Briga Novarese community(16).

Our study contrasts with many worldwide studies who did find spatial clusters(4,15,17–34). In Sicily, Italy, a higher ALS incidence (RR=2.75, 95% CI 1.64-4.89) was found in the eastern flank of the volcano Mount Etna with respect to the western, and a role of volcanogenic metals has been hypothesized to justify such distribution(17). In New Hampshire, the analysis of electronic records over an 18-years period revealed a higher risk (odds ratio = 2.32) in an area within 0.5 mile around Lake Mascoma. Chronic exposure to cyanobacteria has been proposed as a possible cause (24). The cyanobacterial hypothesis has been further investigated in the Herault district, in Southern France. Over the 1994-2009 period, authors identified a significant ALS cluster (RR = 2.24, *p-value* =0.0024) surrounding the Thau lagoon (25). In Jefferson County, using SatScan analysis and several administrative data for case ascertainment, a small cluster (RR = 6.4, *p-value* =0.04) was found around a lead smelter (26). Analyzing Finnish 1985-1995 death certificate and adopting the spatial scan statistics, two neighboring clusters were identified in the south of Finland for deathplace (RR = 1.79, *p-value* = 0.00001, and RR = 1.32, *p-value* = 0.013), one of which was similarly found analyzing patients' birthplace (RR = 1.53, *p-value* = 0.00001). Based on the population history in which major bottlenecks occurred, authors concluded that genetics factors were likely to explain such clusters(31). In a 2010 study, using national mortality data of a ten-years period, three

significant ALS clusters were identified in the center of Japan (two of males patients, RR = 1.56, *p-value* <0.001, and RR = 1.39, *p-value* = 0.03, and one of females patients, RR = 1.50, *p-value* = 0.02) but no explanation for this finding were prompted(30). In New Jersey, analyzing the 2009-2011 data from a national surveillance project, ALS risk was found to vary across the country (with RR ranging from 0.75 to 1.42) but no significant clusters were identified by SatScan analysis(32). In the Limousin region, France, the analysis of 10-years data identified three clusters; paper paste and water treatment plants were found to be statistically correlated to such distribution(28). In South-East England, 1990-2006 register data and SatScan analysis revealed a significant cluster (RR = 1.70, *p-value* = 0.012) in Greater London, and several smaller clusters in the urban areas around London. Occupational and educational factors have been proposed to have shed light on in further studies(29). Rooney et al. used both SatScan and Flexscan softwares to analyze the 18-years data from the Irish ALS register(27). Two lower ALS risk areas were identified in the south and in the southwest of Ireland (RR = 0.0, *p-value* = 0.029, and RR = 0.53, *p-value* = 0.012, respectively). No obvious explanation were proposed, despite authors postulated that genetic admixture in these regions could be the basis for such finding.

The population-based setting, the large cohort size and the adoption of two proper spatial statistics strongly support the validity of our results. Indeed, the use of more than one approach is recommended in order to overcome the limits of the each spatial statistics(3). Furthermore, in order to exclude that a different administrative partition of the study area would have led to different results (a problem commonly known as “modifiable areal unit

problem") we performed the analysis at both municipalities and census divisions level, finding the same result(6).

Nonetheless, some limits should be taken into account for this study too. Given a one-year median diagnostic delay(35), address at diagnosis can be approximated to address at disease onset. Nonetheless, putative environmental factors are thought to act several years before the disease onset(2) and thus residential history would have been a more reliable index of patients' exposures. On the other hand, considering ALS pathogenesis as a multistep process(36), address at diagnosis could give cues on later pathogenic events. Moreover, residence does not prompt the entire subject's environmental exposure and job exposures should be considered too.

In conclusion, using the 20-years population-based data from the PARALS register and a proper statistical analysis, we found a substantial homogenous distribution of ALS cases in Piedmont and Aosta Valley. Such result constitutes a further proof of the involvement of multiple factors, both environmental and genetic, in ALS pathogenesis.

Disclosure of conflicts of interest. The authors declare no financial or other conflicts of interest.

References

1. Talbot K. Motor neuron disease: the bare essentials. Pract Neurol. 2009 Oct;9(5):303–9.

2. Al-Chalabi A, Hardiman O. The epidemiology of ALS: a conspiracy of genes, environment and time. Nat Rev Neurol. 2013 Nov;9(11):617–28.

3. Pfeiffer D, Robisons T, Stevenson M, Stevens K, Rogers D, Clements A. Spatial analysis in epidemiology. Oxford University Press; 2008.

4. Migliaretti G, Berchialla P, Dalmasso P, Cavallo F, Chiò A. Amyotrophic lateral sclerosis in Piedmont (Italy): a Bayesian spatial analysis of the incident cases. Amyotroph Lateral Scler Front Degener. 2013 Jan;14(1):58–65.

5. Chiò A, Mora G, Moglia C, Manera U, Canosa A, Cammarosano S, et al. Secular Trends of Amyotrophic Lateral Sclerosis: The Piemonte and Valle d'Aosta Register. JAMA Neurol. 2017 Jul 10;

6. Openshaw S. The Modifiable Areal Unit Problem. Norwich, UK: Geo Books; 1984.

7. Istituto Nazionale di Statistica [Internet]. Available from: http://demo.istat.it/

8. Benvenuto in QGIS! [Internet]. [cited 2017 Jun 6]. Available from: http://www.qgis.org/it/site/

9. Knox, E.G. P. Detection clusters. In: Methodologies of Enquiry into Disease Clustering., Elliot P. London: Small area health statistics unit; 1989.

10. Moran P a. P. Notes on continuous stochastic phenomena. Biometrika. 1950 Jun;37(1–2):17–23.

11. Kulldorff, M. A spatial scan statistic. Commun Stat Theory Methods. 1997;26:1481–96.

12. R: The R Project for Statistical Computing [Internet]. [cited 2017 Jun 6]. Available from: https://www.r-project.org/

13. SaTScan - Software for the spatial, temporal, and space-time scan statistics [Internet]. [cited 2017 Jun 6]. Available from: https://www.satscan.org/

14. Bersano E, Servo S, Cantello R, Mazzini L. A rare case of conjugal amyotrophic lateral sclerosis. J Neurol. 2014 Jun;261(6):1216–7.

15. Uccelli R, Binazzi A, Altavista P, Belli S, Comba P, Mastrantonio M, et al. Geographic distribution of amyotrophic lateral sclerosis through motor neuron disease mortality data. Eur J Epidemiol. 2007;22(11):781–90.

16. Tesauro M, Consonni M, Filippini T, Mazzini L, Pisano F, Chiò A, et al. Incidence of amyotrophic lateral sclerosis in the province of Novara, Italy, and possible role of environmental pollution. Amyotroph Lateral Scler Front Degener. 2017 May;18(3–4):284–90.

17. Nicoletti A, Vasta R, Venti V, Mostile G, Lo Fermo S, Patti F, et al. The epidemiology of amyotrophic lateral sclerosis in the Mount Etna region: a possible pathogenic role of volcanogenic metals. Eur J Neurol. 2016 May;23(5):964–72.

18. Caller TA, Chipman JW, Field NC, Stommel EW. Spatial analysis of amyotrophic lateral sclerosis in Northern New England, USA, 1997-2009. Muscle Nerve. 2013 Aug;48(2):235–41.

19. Miranda ML, Alicia Overstreet Galeano M, Tassone E, Allen KD, Horner RD. Spatial analysis of the etiology of amyotrophic lateral sclerosis among 1991 Gulf War veterans. Neurotoxicology. 2008 Nov;29(6):964–70.

20. Schwartz GG, Rundquist BC, Simon IJ, Swartz SE. Geographic distributions of motor neuron disease mortality and well water use in U.S. counties. Amyotroph Lateral Scler Front Degener. 2017 May;18(3–4):279–83.

21. Noonan CW, White MC, Thurman D, Wong L-Y. Temporal and geographic variation in United States motor neuron disease mortality, 1969-1998. Neurology. 2005 Apr 12;64(7):1215–21.

22. Sienko DG, Davis JP, Taylor JA, Brooks BR. Amyotrophic lateral sclerosis. A case-control study following detection of a cluster in a small Wisconsin community. Arch Neurol. 1990 Jan;47(1):38–41.

23. Forbes RB, Colville S, Parratt J, Swingler RJ. The incidence of motor nueron disease in Scotland. J Neurol. 2007 Jul;254(7):866–9.

24. Caller TA, Doolin JW, Haney JF, Murby AJ, West KG, Farrar HE, et al. A cluster of amyotrophic lateral sclerosis in New Hampshire: a possible role for toxic cyanobacteria blooms. Amyotroph Lateral Scler Off Publ World Fed Neurol Res Group Mot Neuron Dis. 2009;10 Suppl 2:101–8.

25. Masseret E, Banack S, Boumédiène F, Abadie E, Brient L, Pernet F, et al. Dietary BMAA exposure in an amyotrophic lateral sclerosis cluster from southern France. PloS One. 2013;8(12):e83406.

26. Turabelidze G, Zhu B-P, Schootman M, Malone JL, Horowitz S, Weidinger J, et al. An epidemiologic investigation of amyotrophic lateral sclerosis in Jefferson County, Missouri, 1998-2002. Neurotoxicology. 2008 Jan;29(1):81–6.

27. Rooney J, Vajda A, Heverin M, Elamin M, Crampsie A, McLaughlin R, et al. Spatial cluster analysis of population amyotrophic lateral sclerosis risk in Ireland. Neurology. 2015 Apr 14;84(15):1537–44.

28. Boumédiène F, Druet-Cabanac M, Marin B, Preux P-M, Allée P, Couratier P. Contribution of geolocalisation to neuroepidemiological studies: incidence of ALS and environmental factors in Limousin, France. J Neurol Sci. 2011 Oct 15;309(1–2):115–22.

29. Scott KM, Abhinav K, Stanton BR, Johnston C, Turner MR, Ampong M-A, et al. Geographical clustering of amyotrophic lateral sclerosis in South-East England: a population study. Neuroepidemiology. 2009;32(2):81–8.

30. Doi Y, Yokoyama T, Tango T, Takahashi K, Fujimoto K, Nakano I. Temporal trends and geographic clusters of mortality from amyotrophic lateral sclerosis in Japan, 1995-2004. J Neurol Sci. 2010 Nov 15;298(1–2):78–84.

31. Sabel CE, Boyle PJ, Löytönen M, Gatrell AC, Jokelainen M, Flowerdew R, et al. Spatial clustering of amyotrophic lateral sclerosis in Finland at place of birth and place of death. Am J Epidemiol. 2003 May 15;157(10):898–905.

32. Henry KA, Fagliano J, Jordan HM, Rechtman L, Kaye WE. Geographic Variation of Amyotrophic Lateral Sclerosis Incidence in New Jersey, 2009-2011. Am J Epidemiol. 2015 Sep 15;182(6):512–9.

33. Betemps EJ, Buncher CR. Birthplace as a risk factor in motor neurone disease and Parkinson's disease. Int J Epidemiol. 1993 Oct;22(5):898–904.

34. Mitchell JD, Gibson HN, Gatrell A. Amyotrophic lateral sclerosis in Lancashire and South Cumbria, England, 1976-1986. A geographical study. Arch Neurol. 1990 Aug;47(8):875–80.

35. Hardiman O, Al-Chalabi A, Brayne C, Beghi E, van den Berg LH, Chio A, et al. The changing picture of amyotrophic lateral sclerosis: lessons from European registers. J Neurol Neurosurg Psychiatry. 2017 11;

36. Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. - PubMed - NCBI [Internet]. [cited 2017 Jun 5]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/25300936

Figure 1. The community of Acquiterme including the high ALS incidence cluster (dark grey) detected by SatScan analysis.