ASOCIACIÓN ENTRE LA EXPOSICIÓN A LARGO PLAZO A FACTORES AMBIENTALES Y LA OCURRENCIA DE LA ESCLEROSIS LATERAL AMIOTRÓFICA (ELA) EN CATALUÑA, ESPAÑA

MÉTODOS ESTADÍSTICOS EN REAL WORLD DATA Y REAL WORLD EVIDENCE

Dr. Maria Antònia Barceló, Prof. Dr. Marc Saez

Escuela de Estadística, Universidad Nacional de Colombia Medellín, 2 de Septiembre de 2019

Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Spain
CIBER of Epidemiology and Public Health (CIBERESP), Spain

Grup de Recerca en Estadística, Econometria i Salut

UdG

ciberesp
Centro de Investigación Biomédica en Red
Epidemiología y Salud Pública

ASOCIACIÓN ENTRE LA EXPOSICIÓN A LARGO PLAZO A FACTORES AMBIENTALES Y LA OCURRENCIA DE LA ESCLEROSIS LATERAL AMIOTRÓFICA (ELA) EN CATALUÑA, ESPAÑA

Methods in Neuroepidemiology

Neuroepidemiology 2018;51:33–49
DOI: 10.1159/000489664

Received: February 17, 2018
Accepted: April 27, 2018
Published online: May 31, 2018

Spatial Assessment of the Association between Long-Term Exposure to Environmental Factors and the Occurrence of Amyotrophic Lateral Sclerosis in Catalonia, Spain: A Population-Based Nested Case-Control Study

Mònica Povedano[a] Marc Saez[b, c] Juan-Antonio Martínez-Matos[a]
Maria Antònia Barceló[a–c]

[a]Functional Motoneurona Unit (UFMNA), Bellvitge Biomedical Research Institute (IDIBELL), L'Hospitalet de Llobregat, Barcelona, Spain; [b]Research Group on Statistics, Econometrics and Health (GRECS), University of Girona, Girona, Spain; [c]CIBER of Epidemiology and Public Health (CIBERESP), Madrid, Spain

INTRODUCTION

Incidence of ALS: 1.75 per 100,000 persons/years of follow-up.

Prevalence of ALS: 5-5.4 per 100,000 inhabitants.

ALS is a heterogeneous disease and its causes cannot be explained only from the genetic point of view.

INTRODUCTION

Familial ALS only accounts for 10% of all ALS patients, although in approximately 30% of familial ALS, the genetic aetiology is still unknown.

Sporadic ALS has no apparent heritability despite the fact that the genetic aetiology is also not known.

So far, the only recognized risk factors for ALS are advanced age, being male and having a family history of ALS.

INTRODUCTION

Genetic mutations do not entirely explain the heterogeneity because the same mutation can be associated with a large variability of ALS phenotypes.

In fact, exposure to environmental factors in combination with an underlying genetic risk is the most likely explanation for how ALS develops.

INTRODUCTION

The non-genetic factors may include variables related to:

- Lifestyle: smoking, consumption of antioxidants, physical exercise and Body Mass Index.
- Medical conditions: head injury, metabolic diseases, cancer and inflammatory diseases.
- Work and environmental related exposure: BMAA, viral infections, electromagnetic fields, heavy metal toxicity and exposure to pesticides and fertilizers.

INTRODUCTION

However, the systematic evidence of the association between environmental risk factors and the occurrence of ALS, besides being very limited, is inadequate or insufficient.

Comparatively, there is much more evidence on the association with ALS to the exposure to pesticides. In particular, with organochlorine pesticides.

INTRODUCTION

There is also enough evidence of the existence of spatial clusters of ALS, some of which are associated with environmental factors.

Our results suggest that these clusters could be related to some of the environmental variables, in particular agricultural chemicals.

In addition, in high-risk clusters, besides corresponding to agricultural areas, key road infrastructures with a high density of traffic are also located.

OBJECTIVES

Our objective here is twofold:

1) To investigate the association between long-term exposures to environmental factors and the occurrence of amyotrophic lateral sclerosis (ALS) in Catalonia, Spain.

2) To provide evidence of the existence of spatial clusters of ALS related to these environmental factors.

METHODS

DESIGN

We carried out a nested case-control study constructed from a retrospective population-based cohort, covering the entire region of Catalonia, Spain.

This cohort includes all the patients who were assessed at the Motor Disease Functional Unit (UFMNA) of University Hospital of Bellvitge, L'Hospitalet de Llobregat, Spain and met the 'El Escorial' diagnostic criteria for ALS.

METHODS

DESIGN

Cases were subjects diagnosed with ALS between 1st January, 2011 and 31st December, 2016 (n=383, 55.6% of whom were men).

Controls were subjects, alive and free of ALS and other neurodegenerative diseases (including Alzheimer, Parkinson's and parkinsonism), who had had contact with the unit services from 2011 to 2016.

Cases were matched with the controls by sex, age at diagnosis (or visit) with a tolerance range of +/- 5 years and year of the ALS diagnosis or visit.

METHODS

VARIABLES – ENVIRONMENTAL EXPOSURE VARIABLES

- Exposure to pesticides.

- Exposure to air pollutants associated with traffic.

- Other environmental variables.

METHODS

VARIABLES – CONTROL VARIABLES

- Variables associated with the individual: sex, age (at diagnosis or visit), year of ALS diagnosis (for cases) and year of visit (for controls).

- Variables associated with the subject's family: indicator of family, family history of the disease.

- Contextual variables: deprivation index.

Since the original cohort consisted of a non-random sample, we included in the model as an offset the expected numbers of ALS cases in each census tract of each municipality of the study area.

METHODS

Estimating and representing the smoothed standardized incidence rates

$$\log\left(\frac{\text{Prob}(Y_i = 1)}{1 - \text{Prob}(Y_i = 1)}\right) = \eta_i$$

METHODS

Estimating and representing the smoothed standardized incidence rates

To estimate the smoothed standardized incidence rates we included several random effects in the linear predictor of the logistic model but no observable explanatory variables (although we did include the expected cases in each census tract as an offset).

- Family and individual heterogeneity.

- Spatial dependence.

- Temporal trends.

METHODS

WHY SMOOTH SMALL AREA DISEASE RATES?

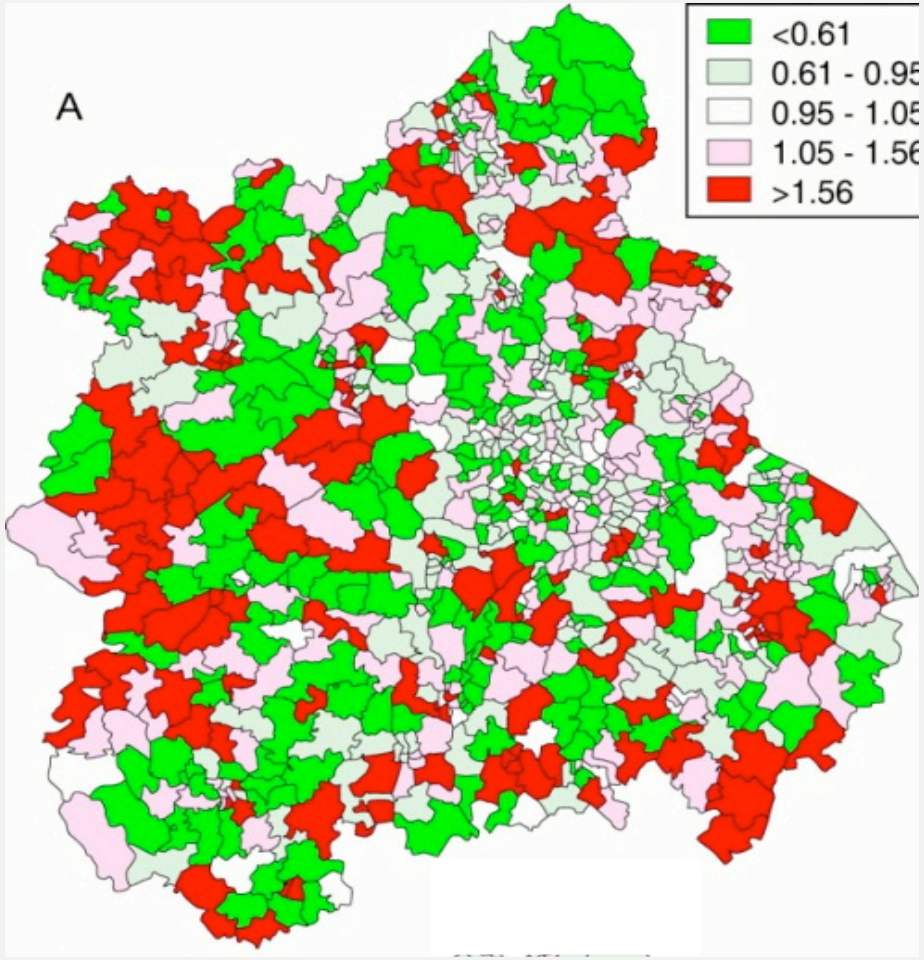
- Typically dealing with rare events in small areas A_i

  Y_i is the observed count of disease in area A_i

  E_i is the expected count based on population size, adjusted for age, sex, other strata ….,

- Relative risk usually estimated by $SIR_i = Y_i / E_i$
- Standard practice is to map SIRs

METHODS



Map of SMR of adult leukaemia in West Midlands Region, England 1974-1986
(Olsen, Martuzzi and Elliott, *BMJ* 1996;313:863-866)

WHY SMOOTH SMALL AREA DISEASE RATES?

- SIR represents estimate of 'true' (underlying) risk in an area, R_i, i.e. $R_i = SIR_i$
- Statistical uncertainty about estimate based on assuming Binomial (or Poisson) sampling variation for data

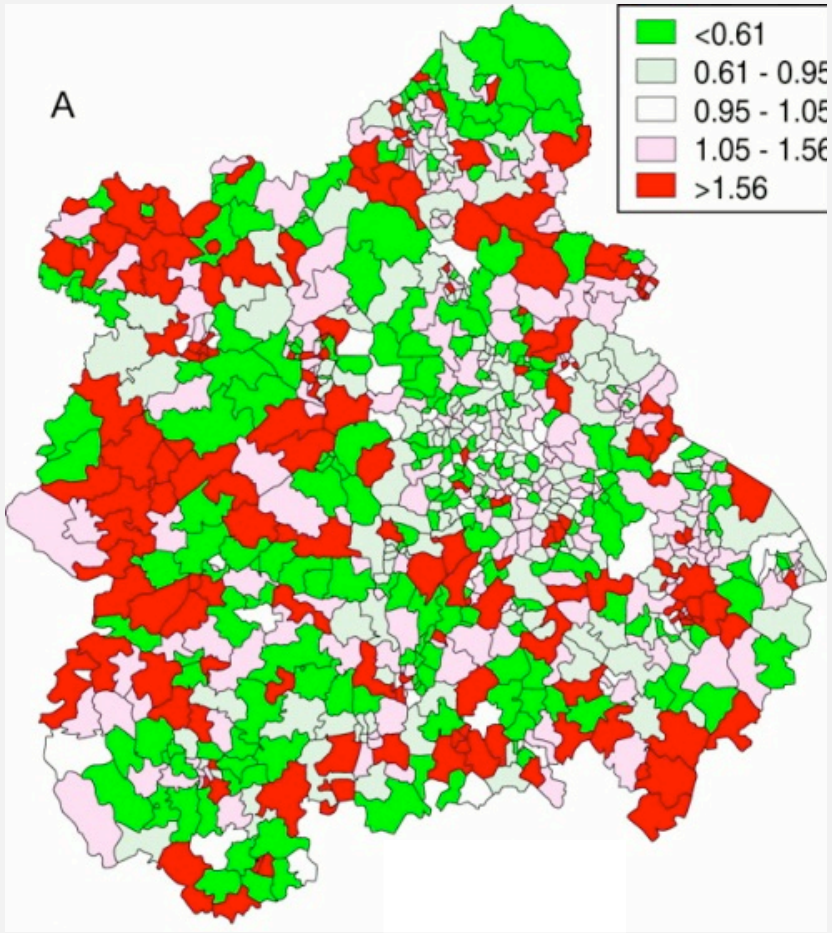
$$Y_i \sim \text{Binomial}(R_i\, E_i)$$

- $SE(R_i) = SE(SIR_i) \propto 1 / E_i$
  - SIR_i very imprecise for rare diseases and small populations
  - precision can vary widely between areas

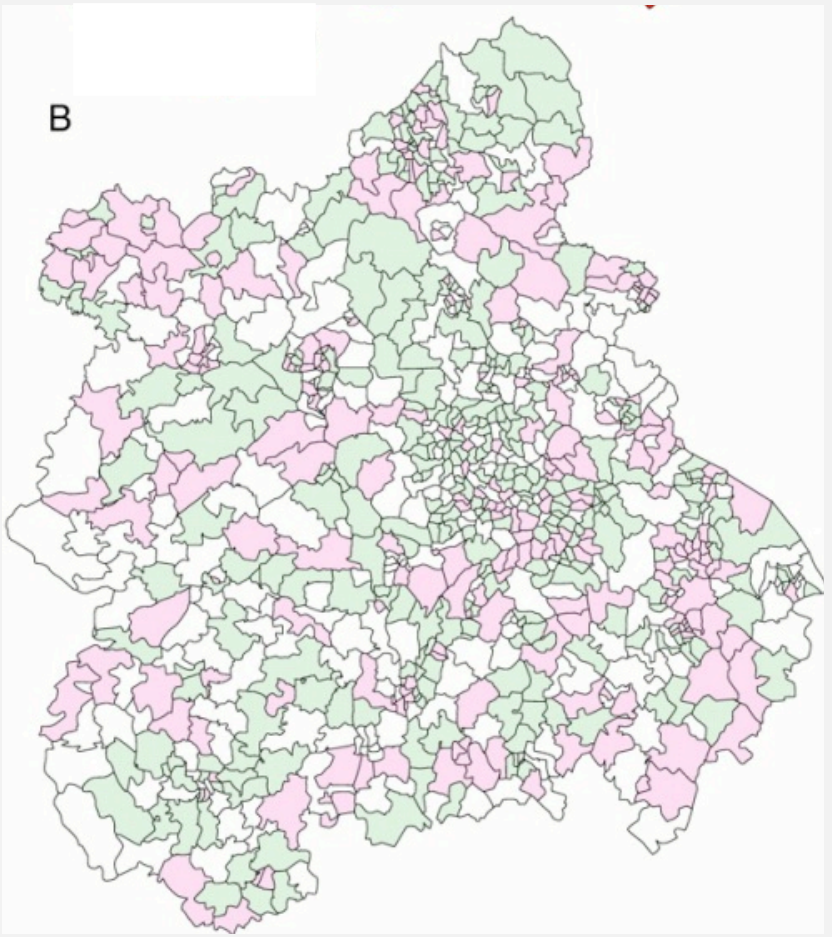
Map of SMR of adult leukaemia in West Midlands Region, England 1974-1986

(Olsen, Martuzzi and Elliott, *BMJ* 1996;313:863-866)

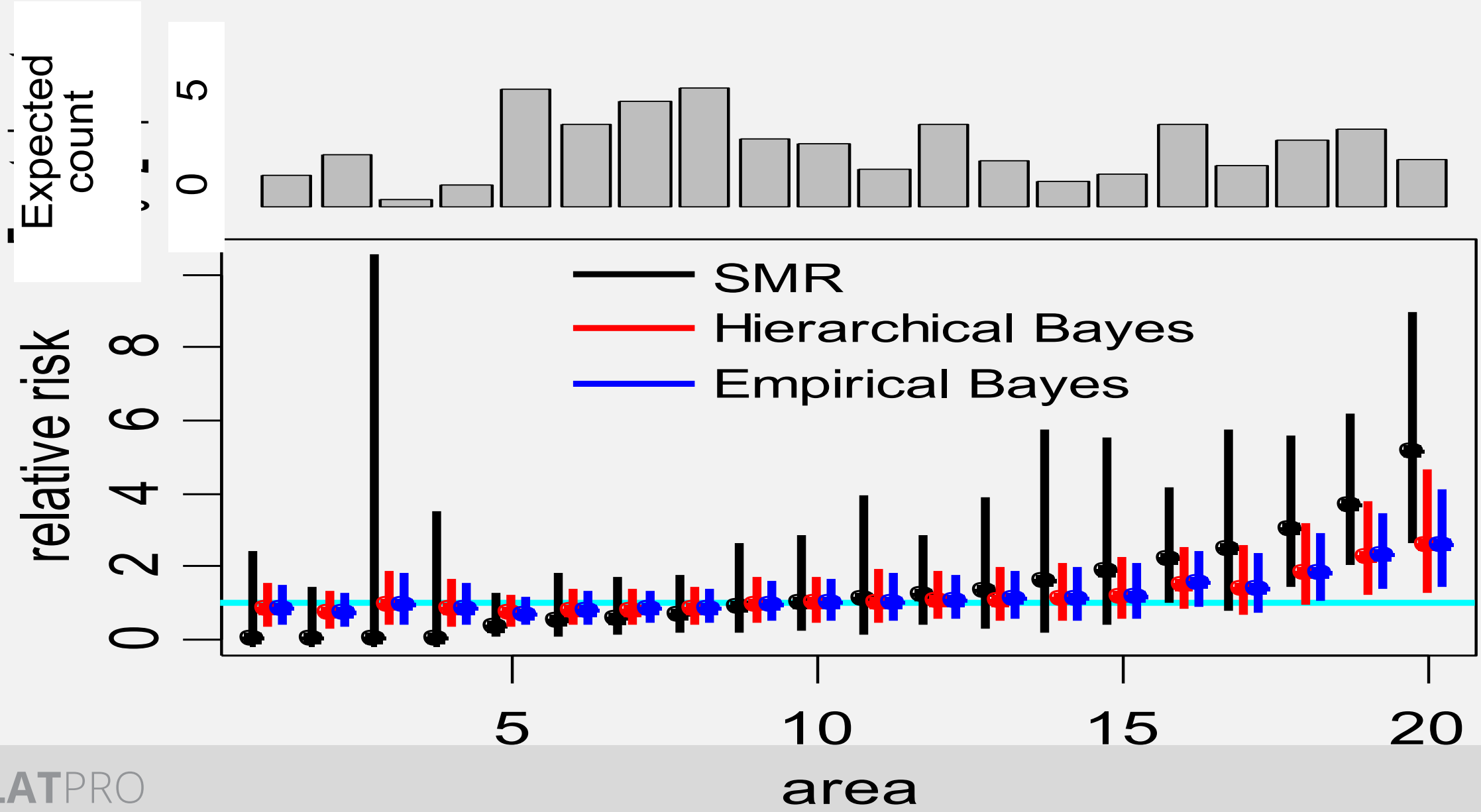
(A) unsmoothed SMR



(B) smoothed by Bayesian methods



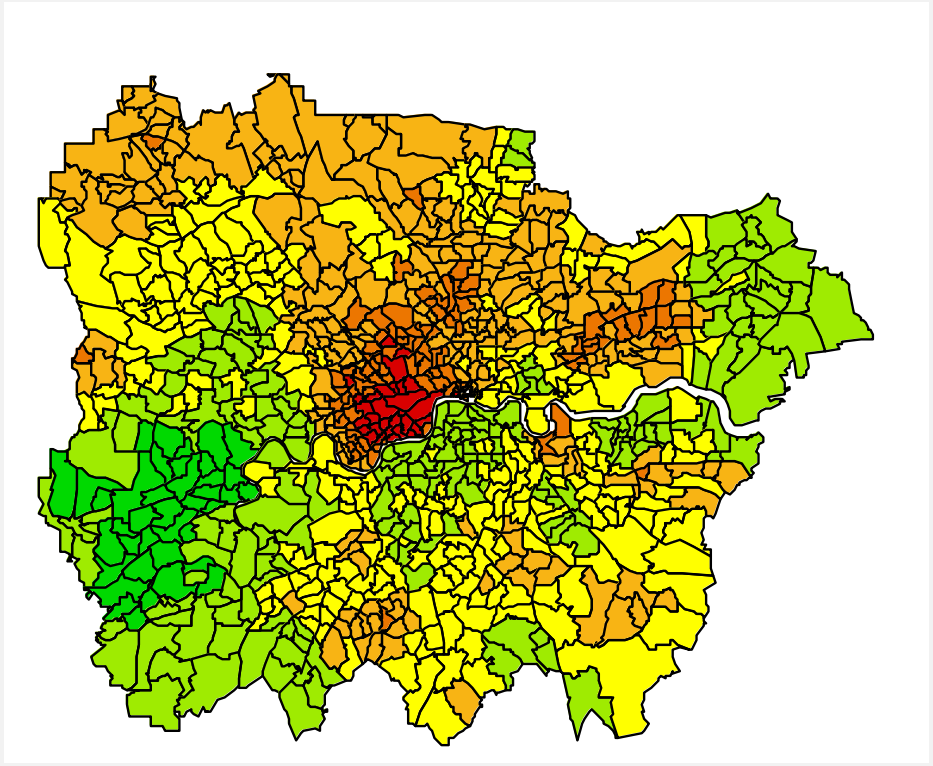
Comparison of estimation methods

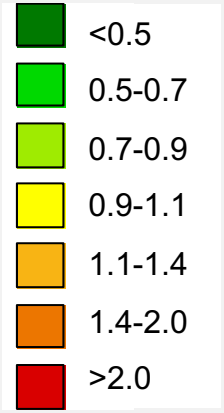


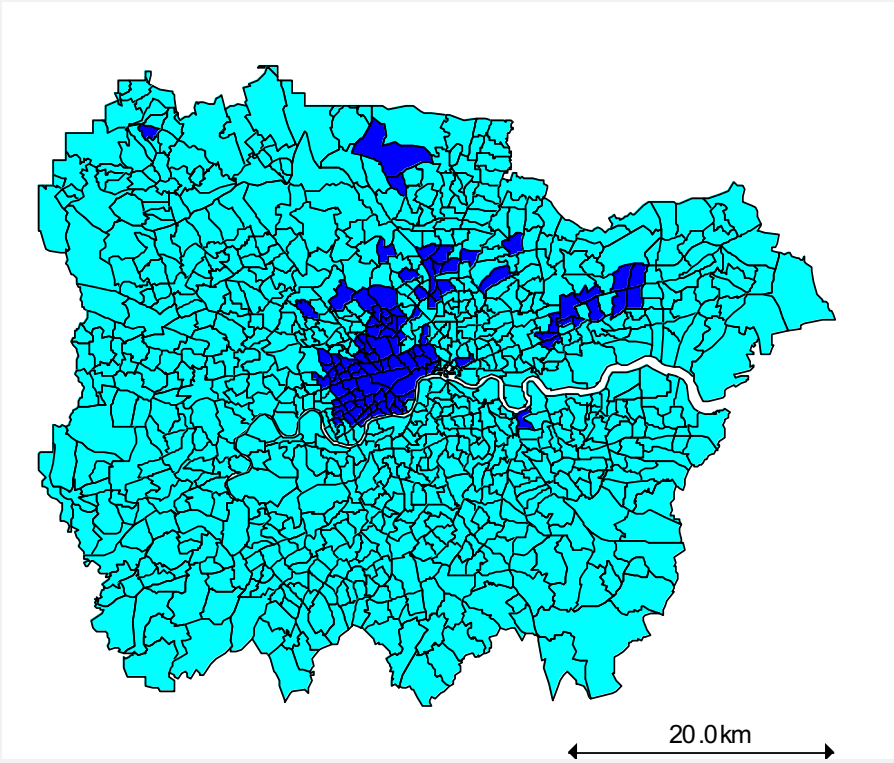
WHY SMOOTH SMALL AREA DISEASE RATES USING HIERARCHICAL BAYESIAN METHODS?

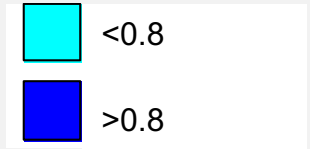
- SIR_i in each area is estimated independently
  - ignores possible spatial correlation between disease risk in nearby areas due to possible dependence on spatially varying risk factors
  - leads to problems of multiple significance testing

Childhood leukaemia incidence in London, 1986-1998

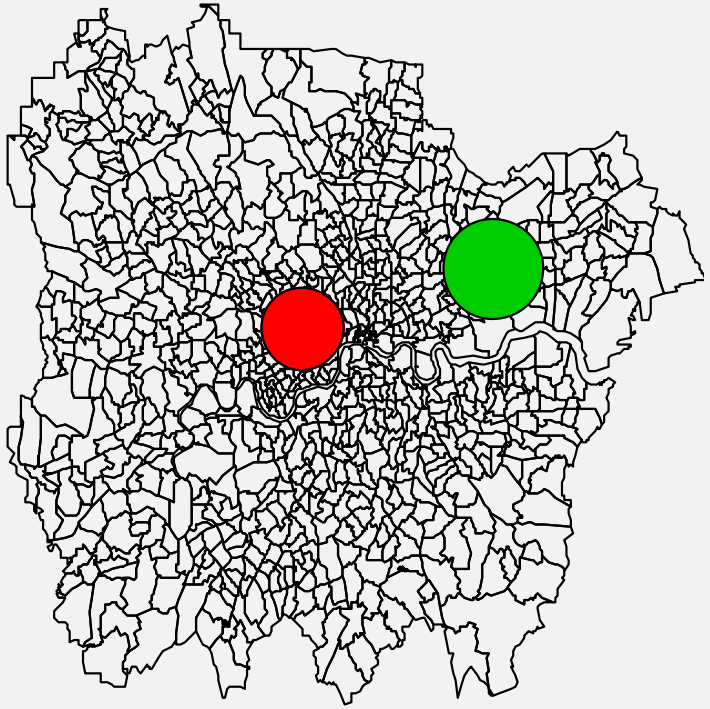


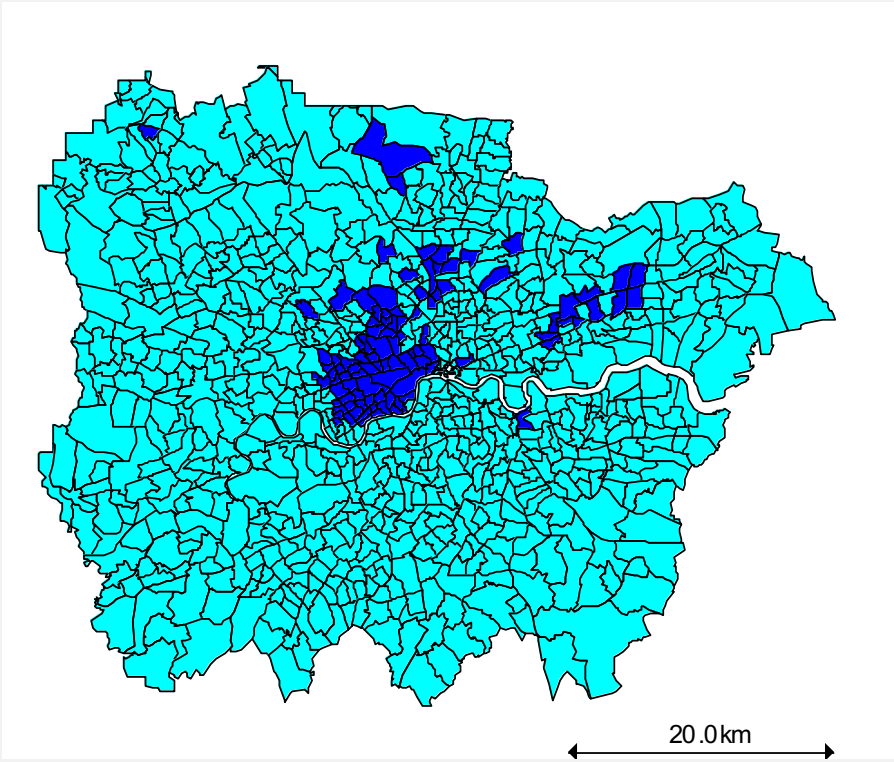


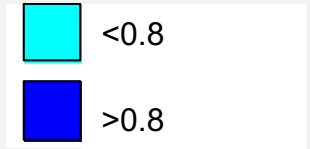




Childhood leukaemia incidence in London, 1986-1998







- Most likely cluster; p<0.001
- 2nd most likely cluster; p = 0.2

METHODS

Estimating the probability of being a case conditioned on the explanatory variables

Our hypothesis is that most of the geographical patterns for ALS, if any, could be explained by environmental variables.

$$\log\left(\frac{\text{Prob}(Y_i = 1)}{1 - \text{Prob}(Y_i = 1)}\right) = \eta_i$$

We included in the linear predictor of each subject in the logistic model, those variables that might explain the probability of being a case, i.e. the environmental variables.

RESULTS

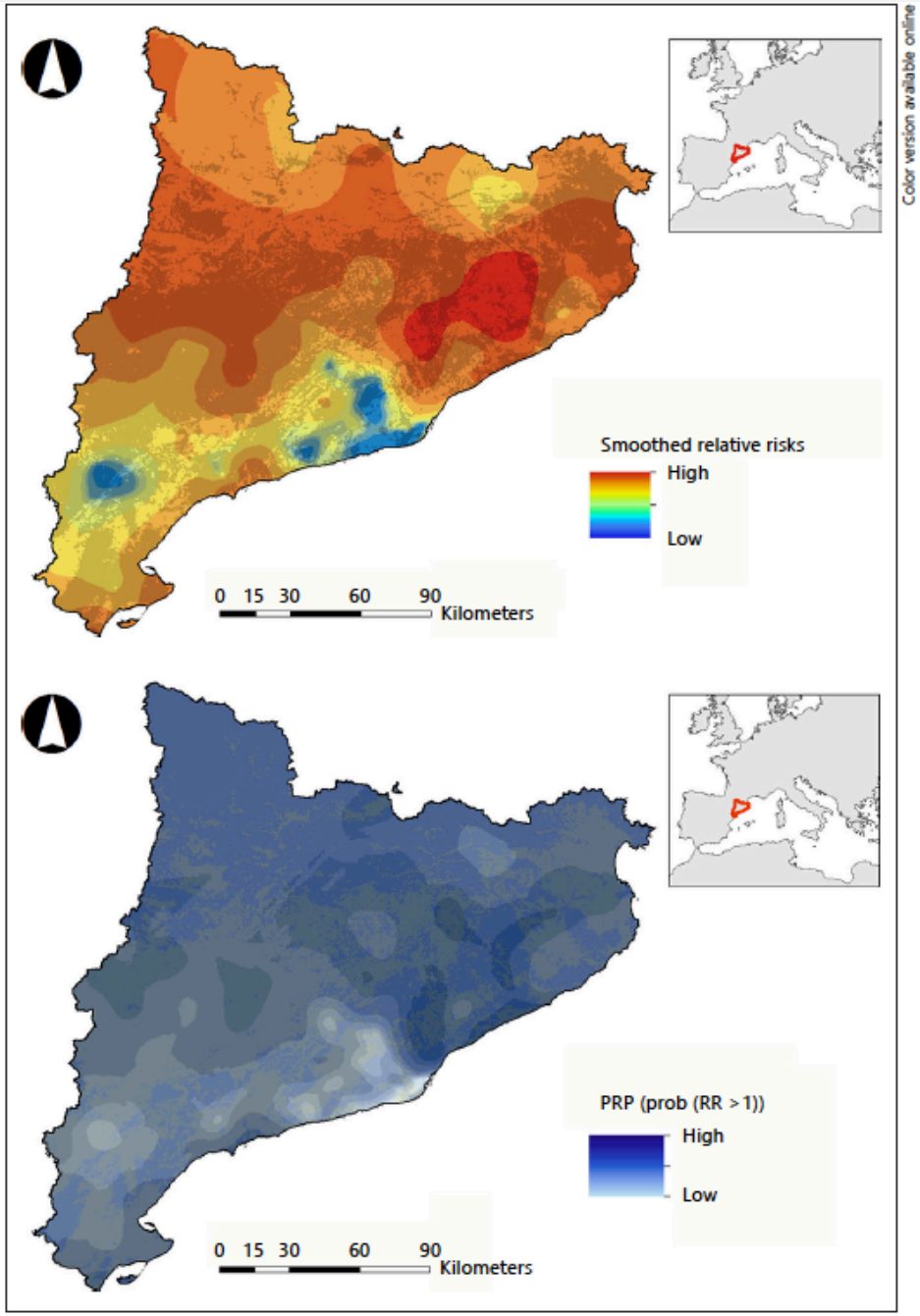


Fig. 1. a Map of the smoothed relative risks over the study region (Catalonia). **b** Map over Catalonia of the posterior probability that the smoothed relative risks were greater than unity (PRP). Model with heterogeneity and spatial adjustment only (besides the expected cases in the census tract as an offset), without explanatory variables [1].

RESULTS

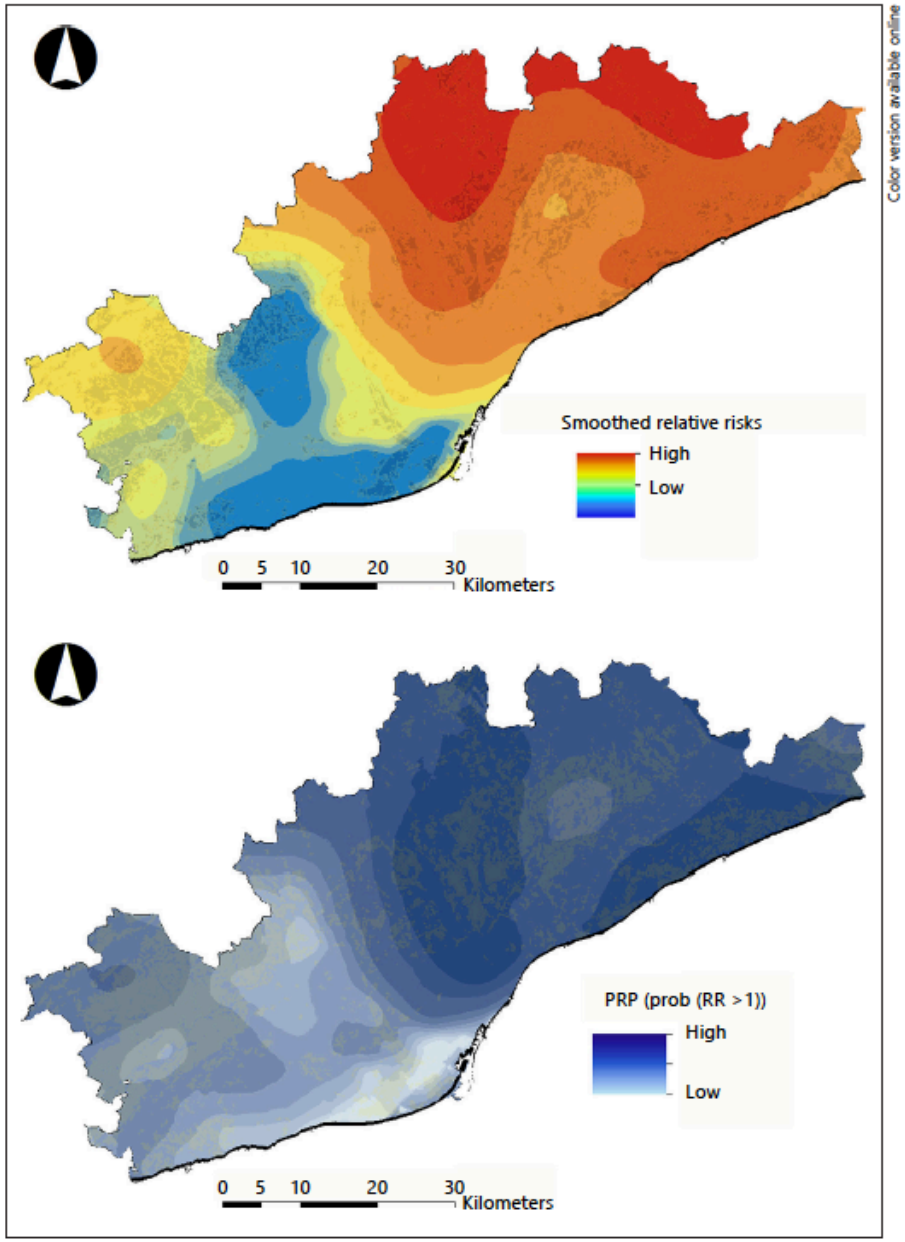


Fig. 2. a Map of the smoothed relative risks over the Metropolitan Area of Barcelona. **b** Map over the Metropolitan Area of Barcelona of the posterior probability that the smoothed relative risks were greater than unity (PRP). Model with heterogeneity and spatial adjustment only (besides the expected cases in the census tract as an offset), without explanatory variables [1].

Grup de Recerca en Estadística, Econometria i Salut

UdG

ciberesp
Centro de Investigación Biomédica en Red
Epidemiología y Salud Pública

RESULTS

Table 2. Association between environmental variables and occurrence of ALS, Catalonia 2011–2016

Variables	OR (95% credibility interval)	Prob ([log(OR)]) >0
Distance agricultural areas (>300 m), m		
<100	5.483 (1.279–25.23)	0.9893
100–199	1.559 (0.809–3.012)	0.9075
200–299	0.910 (0.492–1.672)	0.6197
Distance streets (>100 m), m		
<25	1.615 (0.692–3.778)	0.8658
25–100	1.364 (0.885–2.104)	0.9199
Distance local and county (>200 m), m		
<50	0.514 (0.211–1.198)	0.7666
50–100	1.394 (0.736–2.636)	0.8459
101–200	1.196 (0.709–2.014)	0.7492
Distance dual carriageways and higways (>300 m), m		
<50	0.754 (0.208–2.670)	0.6659
50–100	1.285 (0.588–2.800)	0.7357
101–200	0.665 (0.368–1.191)	0.8152
200–300	0.963 (0.537–1.720)	0.5516
Distance petrol stations (>150 m)	0.502 (0236–1.044)	0.8674
Distance green areas (quintile 1)		
Quintile 2	0.910 (0.481–1.715)	0.6145
Quintile 3	1.283 (0.686–2.403)	0.7814
Quintile 4	0.834 (0.437–1.593)	0.7058
Quintile 5	1.381 (0.662–2.910)	0.8034
Daytime environmental noise (quintile 1)		
Quintile 2	4.531 (0.290–106.8)	0.8447
Quintile 3	0.528 (0.009–36.06)	0.6255
Quintile 4	0.085 (0.001–6.692)	0.8693
Quintile 5	0.386 (0.004–37.10)	0.6635
Evening-time environmental noise (quintile 1)		
Quintile 2	0.443 (0.018–7.500)	0.6931
Quintile 3	4.956 (0.055–380.1)	0.7632
Quintile 4	29.14 (0.294–2,587)	0.8261
Quintile 5	13.12 (0.104–1,499)	0.8538
Night-time environmental noise (quintile 1)		
Quintile 2	1.437 (0.438–4.539)	0.7305
Quintile 3	0.886 (0.199–3.865)	0.5631
Quintile 4	0.817 (0.149–4.381)	0.5930
Quintile 5	0.439 (0.065–2.926)	0.8031
NO_2 (quartile 1)		
Quartile 2	1.872 (1.487–2.023)	0.9973
Quartile 3	2.047 (1.698–2.898)	0.9973
Quartile 4	2.703 (1.265–3.255)	0.9984
NO (quartile 1)		
Quartile 2	0.515 (0.140–1.958)	0.7216
Quartile 3	0.329 (0.074–1.482)	0.8394
Quartile 4	1.321 (0.530–3.340)	0.9271
Benzopirene	1.122 (0.359–4.140)	0.8539
Cadmium	1.332 (0.729–13.71)	0.9262
Ozone	0.973 (0.949–1.061)	0.9267

Adjusted by sex, age, year of diagnosis, family history of disease, indicator of family in the database, contextual deprivation index.

Prob (abs [log(OR)] >0) higher than 0.95. Prob (abs [log(OR)] >0) higher than 0.90.

DISCUSSION

We found a certain geographical pattern for the risk of ALS occurrence. In addition, three clusters can be observed. Two of them with a high risk of occurrence of ALS - one in the centre of the study region (north-south direction) and another in the east (direction southwest-northeast), and the other with a moderate-high risk in the west.

DISCUSSION

Our results could be in line with the findings of the studies that attribute some role to the exposure to agricultural chemicals

In this sense, all the clusters we identified correspond to areas of intensive agriculture.

In our case, the high-risk clusters, besides corresponding to agricultural areas, also correspond to significant road infrastructures that carry a high density of traffic.

DISCUSSION

The results of the multivariate model suggest that these clusters could be related to some of the environmental variables.

Living near an agricultural area increased the risk of ALS occurrence.

Air pollution resulting from traffic could also be related to the occurrence of ALS.

We also found a statistically significant association between exposure to ozone and the occurrence of ALS.

And with benzopyrene and cadmium levels in the air.

DISCUSSION

Summing up, exposure to pesticides, which we approximated by proximity to agricultural areas, as well as to certain pollutants (particularly those whose source is traffic, most likely diesel vehicles, i.e. exposure to nitrogen oxides), could be independent predictors of the occurrence of ALS.

But, in addition, both the high-risk clusters that we have found, as well as the significance of the interactions in the multivariate models, allow us to hypothesize that exposure to high levels of air pollutants as a result of traffic could increase the risk associated with living close to agricultural areas.

THANK YOU

For Your Precious Time and Attention…

antonia.barcelo@udg.edu

marc.saez@udg.edu