Geographical pattern of brain cancer incidence in the Navarre and Basque Country regions of Spain

G López-Abente, M Pollán, E Ardanaz, M Errezola

Background: The study of the distribution of brain cancer mortality in Spain shows a grouping of highest risk provinces corresponding to the autonomous regions of Navarre and the Basque Country.

Aim: To explore the possible existence of geographical patterns in these areas.

Methods: Municipal maps of brain cancer incidence were drawn up and the influence of land use related variables on the distribution of the disease duly analysed. Autoregressive conditional models were used to plot smoothed municipal maps. The influence of explanatory land use variables, ascertained by remote sensing, was assessed.

Results: The maps revealed that certain towns situated in the "Media" and "Cantábrico-Baja Montaña" districts of Navarre were areas of highest risk. Among the towns in question, those in the "Media" district lie very close to the city of Pamplona. However, the pattern of brain cancer incidence in Navarre and the Basque Country could not be conclusively said to be determined by any specific type of land cover and/or crop.

Conclusions: Results suggest a possible increase of risk linked to areas devoted to a high percentage of non-irrigated arable land.

The study of the distribution of brain cancer mortality in Spain shows a grouping of highest risk provinces corresponding to the autonomous regions of Navarre and the Basque Country, with a statistically significant autocorrelation pattern which cannot be explained by random chance.¹ Navarre registers the highest mortality rates in Spain for both sexes, namely 8.45 and 4.69 per 100 000 population in men and women respectively, for the period 1978–92, with a male:female rates ratio of 1.8:1. Similarly, Navarre also registers the highest child and adolescent brain cancer mortality.²

In geographical studies, the choice of a large sized administrative spatial unit (such as a province) tends to dilute the pattern, there being a tendency for differences in risk between smaller intraprovincial areas to be mutually offset. The study of smaller sized, and thus more homogeneous areas (for example, towns) can be useful as a technique for detecting underlying environmental problems. However, the choice of the town as the unit of analysis for the study of geographical patterns poses the problem of low numbers of cases. As a consequence, the use of classic indicators (adjusted rates, standard mortality ratios (SMR), rate ratios) may yield unstable results, and so the usefulness of smoothed estimators (empirical Bayes, full Bayes) in such a situation has been acknowledged.³,⁴

The aetiology of most central nervous system tumours is unknown. Ionising radiations and certain hereditary syndromes⁵ are clearly established risk factors. An exposure that has shown a more “consistent” relation with brain cancer is that deriving from farm work.⁶ This may reflect the effect of different, as yet undiscovered risk factors, especially exposure to pesticides, such as N-nitroso compounds (nitrosoureas in particular), which can produce brain tumours in experimental animals, even at low doses.⁷ Consequently, characterising the types of crop and their phytosanitary treatment may well be of great interest.

Remote sensing renders it possible to obtain information that can be used in environmental epidemiology. To date it has been employed by biologists in ecology, but in epidemiology its use has largely been limited to the control of communicable disease vectors.⁸ This study uses satellite based information to classify crop types, on the assumption that many of the chemical agents employed in agriculture (phytosanitary products) are linked to a specific type of crop.

This study aims to: (1) furnish a risk map depicting the mortality pattern at a municipal level; and (2) analyse the influence of land use related variables on the geographical pattern.

MATERIALS AND METHODS

The study covered all cases of brain cancer (ICD: 191) registered during the period 1986–93 in the Navarre and Basque Country regional population based cancer registries. Standardised incidence ratios (SIR) were calculated for each of the 512 towns, with the overall incidence for both autonomous regions by age group and sex taken as reference. Expected cases were computed using person-years calculated on the basis of the 1991 census of municipal populations by sex and five year age groups.

Municipal classification by crop type was based on data drawn from the CORINE (Coordination of Information on the Environment) Landcover project. The CORINE Landcover project produces high resolution maps using remote sensing techniques processing the LANDSAT satellite imagery. These are vector maps compiled at a scale of 1:100 000 for the whole of Spain and divided into different pages. These pages show land use and, in the case of agricultural pages, include 21 different crop types; the information refers to 1987. Expected cases were computed using person-years calculated on the basis of the 1991 census of municipal populations by sex and five year age groups.

Municipal classification by crop type was based on data drawn from the CORINE (Coordination of Information on the Environment) Landcover project. The CORINE Landcover project produces high resolution maps using remote sensing techniques processing the LANDSAT satellite imagery. These are vector maps compiled at a scale of 1:100 000 for the whole of Spain and divided into different pages. These pages show land use and, in the case of agricultural pages, include 21 different crop types; the information refers to 1987. Expected cases were computed using person-years calculated on the basis of the 1991 census of municipal populations by sex and five year age groups.

The maps revealed that certain towns situated in the “Media” and “Cantábrico-Baja Montaña” districts of Navarre were areas of highest risk. Among the towns in question, those in the “Media” district lie very close to the city of Pamplona. However, the pattern of brain cancer incidence in Navarre and the Basque Country could not be conclusively said to be determined by any specific type of land cover and/or crop.

Conclusions: Results suggest a possible increase of risk linked to areas devoted to a high percentage of non-irrigated arable land.

Abbreviations: CAR, conditional autoregressive; DIC, deviance information criterion; GLMM, generalised linear mixed model; SIR, standardised incidence ratio
dump sites; 6, non-irrigated arable land; 7, permanently irrigated land; 8, vineyards; 9, non-irrigated orchards; 10, citrus farming; 11, permanently irrigated orchards; 12, olive groves; 13, meadows; 14, complex cultivation patterns; 17, evergreen forest; 18, deciduous forest; 19, coniferous forest; 20, mixed forest; 21, pasture; 22, sclerophyllous vegetation; 23, transitional woodland shrub; and 24, wetlands.

### Statistical analysis

**Plotting of smoothed maps**

We fitted Poisson spatial models, which included two random effects terms that took the following into account: (a) municipal contiguity (spatial term); and (b) municipal heterogeneity. These models come under the umbrella of the so called conditional autoregressive (CAR) models for disease mapping, initially proposed by Besag and colleagues and subsequently applied in the field of epidemiology. The models were fitted using Markov Chain Monte Carlo methods with non-informative priors. A burn-in of 50,000 iterations was used to ensure convergence and 1000 iterations to obtain the posterior distribution of the estimator of relative risk for the 512 towns, using the WinBugs program.

### Explanatory variables

To ascertain the influence of land use variables on the geographical pattern, two types of models were fitted. Firstly, we fitted a classic Poisson model, with observed cases as the dependent variable, expected cases as offset, and quartiles of the proportion of municipal land devoted to each type of crop as the central variable of analysis. Then, in variables for which some type of association was found, we used Poisson spatial models of the kind described above (CAR), but including the explanatory variable as a fixed effects term. Variables were categorised in quartiles and included in the models as continuous. In the case of the “non-irrigated arable land” and “permanently irrigated land” variables, their logarithm was also used. In analysing the percentage of farmers, unemployed and illiterates, were used as centred variables by deducting their mean. These indicators were obtained from the 1991 census.

### Table 1

<table>
<thead>
<tr>
<th>Land surface use</th>
<th>Q2 (95% CI)</th>
<th>Q3 (95% CI)</th>
<th>Q4 (95% CI)</th>
<th>Trend p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>1.17 (0.91 to 1.52)</td>
<td>1.27 (1.04 to 1.55)</td>
<td>1.23 (1.02 to 1.47)</td>
<td>NS</td>
</tr>
<tr>
<td>Industrial</td>
<td>0.97 (0.81 to 1.15)</td>
<td>1.10 (0.97 to 1.25)</td>
<td>0.92 (0.79 to 1.06)</td>
<td>NS</td>
</tr>
<tr>
<td>Road and rail networks, port areas</td>
<td>1.12 (0.98 to 1.29)</td>
<td>1.24 (1.03 to 1.49)</td>
<td>0.99 (0.88 to 1.12)</td>
<td>NS</td>
</tr>
<tr>
<td>Mines and quarries</td>
<td>1.17 (0.99 to 1.37)</td>
<td>0.94 (0.81 to 1.08)</td>
<td>1.02 (0.89 to 1.05)</td>
<td>NS</td>
</tr>
<tr>
<td>Non-irrigated arable land</td>
<td>0.82 (0.58 to 1.16)</td>
<td>1.29 (1.11 to 1.49)</td>
<td>1.23 (1.05 to 1.43)</td>
<td>0.0002</td>
</tr>
<tr>
<td>Permanently irrigated land</td>
<td>1.16 (0.79 to 1.71)</td>
<td>1.31 (1.13 to 1.52)</td>
<td>0.92 (0.66 to 1.26)</td>
<td>0.013</td>
</tr>
<tr>
<td>Vineyards</td>
<td>1.05 (0.55 to 2.03)</td>
<td>1.60 (0.76 to 3.36)</td>
<td>0.57 (0.14 to 2.28)</td>
<td>NS</td>
</tr>
<tr>
<td>Non-irrigated orchards</td>
<td>0.53 (0.27 to 1.06)</td>
<td>1.12 (0.36 to 3.48)</td>
<td>0.07</td>
<td>NS</td>
</tr>
<tr>
<td>Permanently irrigated orchards</td>
<td>0.92 (0.80 to 1.06)</td>
<td>1.10 (0.93 to 1.31)</td>
<td>1.10 (0.93 to 1.31)</td>
<td>NS</td>
</tr>
<tr>
<td>Meadowland</td>
<td>0.89 (0.76 to 1.04)</td>
<td>1.04 (0.88 to 1.23)</td>
<td>0.87 (0.77 to 0.98)</td>
<td>0.046</td>
</tr>
<tr>
<td>Complex cultivation patterns</td>
<td>0.87 (0.57 to 1.30)</td>
<td>0.83 (0.71 to 0.96)</td>
<td>0.86 (0.71 to 1.03)</td>
<td>NS</td>
</tr>
<tr>
<td>Evergreen forest</td>
<td>0.90 (0.80 to 1.01)</td>
<td>1.09 (0.94 to 1.27)</td>
<td>1.07 (0.87 to 1.32)</td>
<td>NS</td>
</tr>
<tr>
<td>Deciduous forest</td>
<td>0.84 (0.73 to 0.96)</td>
<td>0.96 (0.86 to 1.09)</td>
<td>0.91 (0.78 to 1.20)</td>
<td>NS</td>
</tr>
<tr>
<td>Coniferous forest</td>
<td>1.29 (1.08 to 1.54)</td>
<td>1.11 (0.93 to 1.32)</td>
<td>1.08 (0.90 to 1.30)</td>
<td>NS</td>
</tr>
<tr>
<td>Mixed forest</td>
<td>1.06 (0.92 to 1.22)</td>
<td>0.94 (0.82 to 1.06)</td>
<td>0.80 (0.68 to 0.93)</td>
<td>0.004</td>
</tr>
<tr>
<td>Pasture</td>
<td>0.95 (0.83 to 1.09)</td>
<td>1.10 (0.82 to 1.49)</td>
<td>1.03 (0.74 to 1.45)</td>
<td>NS</td>
</tr>
<tr>
<td>Sclerophyllous vegetation</td>
<td>1.33 (1.12 to 1.58)</td>
<td>1.09 (0.93 to 1.27)</td>
<td>0.98 (0.79 to 1.20)</td>
<td>NS</td>
</tr>
<tr>
<td>Beaches, dunes, deserts</td>
<td>0.96 (0.86 to 1.08)</td>
<td>1.09 (0.88 to 1.34)</td>
<td>1.01 (0.84 to 1.22)</td>
<td>NS</td>
</tr>
<tr>
<td>Wetlands</td>
<td>0.93 (0.82 to 1.06)</td>
<td>0.84 (0.54 to 1.30)</td>
<td>0.94 (0.68 to 1.31)</td>
<td>NS</td>
</tr>
<tr>
<td>Coastal areas</td>
<td>0.88 (0.77 to 1.00)</td>
<td>0.93 (0.76 to 1.18)</td>
<td>0.85 (0.69 to 1.06)</td>
<td>NS</td>
</tr>
</tbody>
</table>

The reference category is the 1st quartile.

*Tests for trend were calculated assigning an increasing integer to each quartile using the Poisson regression model.

NS, not statistically significant.
attained after iteration 7000. We used a burn-in of 10 000 iterations, and 6000 iterations to obtain the posterior distribution of the parameter of the explanatory variable. In addition, we estimated the effect of the explanatory variables, adjusted for socioeconomic indicators and the proportion of illiterates and unemployed.

RESULTS
There were 1481 cases of brain cancer, distributed among the 512 towns. The number of municipal “adjacencies” was 2846. The overall age adjusted male and female incidence rates were 8.34 and 5.40 per 100 000 population, respectively.

Figure 1 shows the SIR map (left), and the smoothed map (middle). In the former it is difficult to distinguish any type of pattern, owing to the predominance of extreme values. In the smoothed map, however, it is possible to discern a pattern in which the highest risk towns are confined to the “Cantábrica-Baja Montaña” and “Media” districts of Navarre, in contrast to the lower risk observed for the “La Ribera” area. The highest risk towns form a group situated to the south of the city of Pamplona, and centred on Cizur (RR 1.28) and its neighbouring villages (Galar, Adios, Muruzabal, Elorz, Monreal, Puente La Reina). The “Cantábrica district” cluster is centred around Lesaca and Vera del Bidasoa.

The Bayesian approach allows the posterior probability of any area’s relative risks exceeding a threshold to be calculated. Taking this threshold to be unity, these posterior probabilities may be interpreted as the strength of (statistical) evidence of excess risk in each town. High probabilities can be interpreted as providing evidence of an excess risk, while low probabilities have the reverse interpretation. Figure 1 (right) shows these posterior probabilities. The pattern is very similar to that shown in fig 1 (middle), with only two towns with probabilities above 0.975 (Pamplona and Cizur).

Explanatory variables
Table 1 shows the estimated RR by quartile, obtained by fitting the classic Poisson models. The variables that registered a statistically significant rise in relative risk with increase in surface area were “non-irrigated arable land” and “permanently irrigated land”. The more urban towns also registered

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Estimation of relative risk (RR) and 95% credibility interval (CI) for study variables from Bayesian spatial models</th>
</tr>
</thead>
<tbody>
<tr>
<td>RR (95% CI)</td>
<td>RR* (95% CI)</td>
</tr>
<tr>
<td>Non-irrigated arable land† (quartiles)</td>
<td>1.041 (0.949 to 1.161)</td>
</tr>
<tr>
<td>Log(non-irrigated arable land)‡</td>
<td>1.046 (0.953 to 1.127)</td>
</tr>
<tr>
<td>Dichotomous non-irrigated arable land¶</td>
<td>1.121 (0.807 to 1.523)</td>
</tr>
<tr>
<td>Permanently irrigated land† (quartiles)</td>
<td>1.039 (0.846 to 1.232)</td>
</tr>
<tr>
<td>Log(permanently irrigated land)‡</td>
<td>1.010 (0.878 to 1.168)</td>
</tr>
<tr>
<td>Unemployed (centred)§</td>
<td>1.006 (0.962 to 1.050)</td>
</tr>
<tr>
<td>Illiterates (centred)§</td>
<td>0.977 (0.848 to 1.122)</td>
</tr>
<tr>
<td>Dichotomous urban¶</td>
<td>1.170 (0.807 to 1.523)</td>
</tr>
<tr>
<td>Farmers (centred)</td>
<td>1.000 (0.989 to 1.007)</td>
</tr>
</tbody>
</table>

*Adjusted for unemployment and illiteracy. †Percentage of municipal land surface area devoted to this type of crop categorised in quartiles, and included in the model as continuous. ‡Logarithm of the percentage of municipal land surface area devoted to this type of crop. §Percentage in the population of each town minus overall mean. ¶Dichotomous variable, quartiles 3 and 4 versus 1 and 2 (reference).

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Evaluation of the goodness of fit of the different models, using the deviance information criterion (DIC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>Bayesian deviance D(E)</td>
</tr>
<tr>
<td>CAR model</td>
<td>327.0638</td>
</tr>
<tr>
<td>CAR + non-irrigated arable land</td>
<td>326.9857</td>
</tr>
<tr>
<td>CAR + log(non-irrigated arable land)</td>
<td>327.5457</td>
</tr>
<tr>
<td>CAR + permanently irrigated land</td>
<td>326.4911</td>
</tr>
<tr>
<td>CAR + log(permanently irrigated land)</td>
<td>327.3546</td>
</tr>
</tbody>
</table>
excess risk. Risk proved lower in towns having a greater proportion of meadow and woodland. The grey images in fig 2 depict the geographical distribution of non-irrigated and permanently irrigated arable land on the basis of information drawn from the CORINE pages.

Bayesian spatial models were used to test whether the association held between urban land, on the one hand, and non-irrigated and permanently irrigated arable land, on the other. Table 2 shows the results. The variables were categorised in quartiles and included in the models as continuous. We also analysed their logarithmic transform, and non-irrigated arable land and urban land as dichotomous variables. None of these attained statistical significance. The variables of unemployment, illiteracy, and percentage of farmers displayed no association with the distribution of brain cancer. Table 3 shows the criterion of goodness of fit (DIC). According to this criterion the model including the “non-irrigated arable land” covariate proved to be the best. The explanatory variables failed to make any substantial improvement to the fit of the model and the credibility interval included unity, which would lead one to conclude that these seem not to explain the municipal pattern of brain cancer incidence.

**DISCUSSION**

The smoothed maps shown here reveal that certain towns situated in the “Media” and “Cantábrica-Baja Montaña” districts of Navarre are areas of highest risk. Yet despite these results, the pattern of brain cancer incidence in Navarre and the Basque Country cannot be conclusively said to be determined by any specific type of land cover and/or crop. It was very surprising to find an association, on the first analysis, with the “non-irrigated arable land” covariate and not with “irrigated land”. However, the assumption of a greater application of pesticides in irrigated areas in general could be misleading, since the quantity and type of pesticides used depends on the type of crop.

The data presented are incidence data, drawn from internationally standardised population based cancer registries subject to quality controls. It therefore seems most unlikely that the pattern exhibited would be affected by misclassification problems stemming from inclusion of metastases pertaining to other sites. The study period is 1986–93, and is unlikely that differences in diagnostic procedures or introduction of new technologies could explain a pattern.

The models employed belong to the family of generalised linear mixed models (GLMM). They have become popular in the study of geographical patterns, their success being partly a result of the fact that they afford a solution to various problems besetting classic analyses: spatial autocorrelation and extra-Poisson variation, ecological regression analyses, and the wide random variability of indicators which necessitates the use of smoothed estimators for plotting maps. Furthermore, their success is also a result of the availability and widespread use of the BUGS and WinBUGS computer software programs.

In line with Morgenstern, the two analyses presented here come under the category of ecological studies with an exploratory and analytical multiple group design. Despite their practical advantages, such studies pose many methodological problems, which can limit their ability to infer causality. With reference to the problem known as the “ecological fallacy”, there are studies which show that consideration of spatial components in models reduces the disparity between the “ecological” risk and the individual. Moreover the reduction of the spatial unit of analysis also attenuate this type of bias. There remains the problem of confounding variables, other than age and sociodemographic indicators, which are difficult to control for in this type of study.

From an epidemiological point of view, however, the inclusion of spatial terms entails controlling for potential confounding factors that are linked to geographical distribution and cannot be taken into account on the basis of the information available. When the pattern of variation of the covariate is similar to that of disease risk, location may act as a confounder. The mixed model contains a clustering term that stands for the effect of location, thus the regression coefficient for the main covariate is controlled for the effect of location. When the location acts as a confounder, it is not surprising to see a change on introduction of the clustering term. This is one of the causes that explain the change in the statistical significance of the second analysis with the mixed models for the covariate “non-irrigated arable land”. Estimates of regression coefficients in this model may be overly conservative.

To analyse the explanatory variables of the geographical pattern, we relied on information on land use drawn from the CORINE project. This project furnishes material originally obtained by means of remote sensing, and while it is information generated on the basis of images that date from the late 1980s, it may nonetheless be adequate for the purposes of this study, in view of the long latency periods observed for this particular type of tumour. However, such information does pose a number of problems, namely: (1) we believe that these types of crop have traditionally been fairly stable in these regions, yet changes may have been made in crop type and treatment over time; (2) we do not know whether the CORINE classification has been validated by a ground truth survey; and (3) the cost of the files is high in Spain, which reduces the efficiency of this type of study.

The only two clearly established risk factors for brain cancer are certain hereditary syndromes and ionising radiations, but these can only explain a small percentage of cases. In the literature, associations have been described between brain cancer and occupational exposures in the following cases: the petrochemical and rubber industries; electromagnetic fields; and among farmers. Insofar as farmers are concerned, there is a strong suspicion that the excess risk may be caused by their exposure to phytosanitary products. However, this has been the subject of discussion, owing to the fact that farmers routinely perform tasks linked to other occupations, in the course of which they are exposed to petrochemical substances, engine exhaust fumes, organic and mineral dust, and different biological agents. In an occupational mortality study conducted in Spain, farmers in the Castile-León region registered an OR of 1.71 (95% CI 1.07 to 2.71), a finding that was repeated for proportional mortality nationwide for this same occupational group. In the study of child and adolescent mortality in Spain, the provincial distribution pattern not only coincided with that registered for all age groups, but also proved to be correlated with the three socioeconomic factors studied—that is, non-cancer-related infant mortality, percentage of provincial land surface area devoted to agriculture, and industrial/construction activity.

Further risk factors studied were exposure to N-nitroso compounds in drinking water and diet, with inconsistent results. Recently, there has been a report of increased risk of this tumour in farmers traceable to the use of pesticides in vineyards. However, our results do not support this finding, since the La Ribera district in Navarre, an area rich in vineyards, has one of the lowest incidences of brain cancer. In summary, smoothed maps reveal certain towns situated in the “Media” and “Cantábrica-Baja Montaña” districts of Navarre as being areas of highest risk. Among the towns in question, those in the “Media” district lie very close to the city of Pamplona. Notwithstanding these results, the pattern of brain cancer incidence in Navarre and the Basque Country cannot be conclusively said to be determined by any specific type of land cover and/or crop, even though our results suggest a probable increase of risk linked to areas devoted to a high percentage of non-irrigated arable land. Furthermore, this study explores new ways of making use of information drawn from the CORINE pages.
from population based cancer registries, which we feel ought to play an important role in discovering the environmental risk factors responsible for the distribution of cancer in the population.

ACKNOWLEDGEMENTS

This study was financed in part by Grant No. 98/1091 from the Fondo de Investigación Sanitaria (Health Research Fund). We are grateful to Stefanie Weykam for her help with the crop type data processing.

References

Geographical pattern of brain cancer incidence in the Navarre and Basque Country regions of Spain
G López-Abente, M Pollán, E Ardanaz and M Errezola

*Occup Environ Med* 2003 60: 504-508
doi: 10.1136/oem.60.7.504

Updated information and services can be found at:
http://oem.bmj.com/content/60/7/504

These include:

**References**
This article cites 23 articles, 8 of which you can access for free at:
http://oem.bmj.com/content/60/7/504#BIBL

**Email alerting service**
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

**Notes**

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/