Geospatial Health 2(2), 2008, pp. 263-271

Sensitivity analysis of the relationship between disease occurrence and distance from a putative source of pollution

Emanuela Dreassi[1], Corrado Lagazio[2], Milena M. Maule[3], Corrado Magnani[4], Annibale Biggeri[1,5]

[1]Department of Statistics "G. Parenti", University of Florence, Florence, Italy; [2]Department of Statistical Sciences, University of Udine, Udine, Italy; [3]Cancer Epidemiology Unit, University of Turin, Turin, Italy; [4]Department of Medical Sciences, University of Eastern Piedmont, Novara, Italy; [5]Biostatistics Unit, Institute for Cancer Prevention, Florence, Italy

Abstract. The relation between disease risk and a point source of pollution is usually investigated using distance from the source as a proxy of exposure. The analysis may be based on case-control data or on aggregated data. The definition of the function relating risk of disease and distance is critical, both in a classical and in a Bayesian framework, because the likelihood is usually very flat, even with large amounts of data. In this paper we investigate how the specification of the function relating risk of disease with distance from the source and of the prior distributions on the parameters of the function affects the results when case-control data and Bayesian methods are used. We consider different popular parametric models for the risk distance function in a Bayesian approach, comparing estimates with those derived by maximum likelihood. As an example we have analyzed the relationship between a putative source of environmental pollution (an asbestos cement plant) and the occurrence of pleural malignant mesothelioma in the area of Casale Monferrato (Italy) in 1987-1993. Risk of pleural malignant mesothelioma turns out to be strongly related to distance from the asbestos cement plant. However, as the models appeared to be sensitive to modeling choices, we suggest that any analysis of disease risk around a putative source should be integrated with a careful sensitivity analysis and possibly with prior knowledge. The choice of prior distribution is extremely important and should be based on epidemiological considerations.

Keywords: case-control study, environmental pollution, absestos, focused clustering, hierarchical Bayesian models, sensitivity to prior choice.

Introduction

The relationship between disease risk and a point source of pollution is usually investigated using the distance from the source as a proxy of population exposure. The analysis may be based on case-control data (Diggle, 1990; Diggle and Rowlingson, 1994) or case-event data (Lawson, 1993).

Corresponding author:
Annibale Biggeri
Department of Statistics "G. Parenti"
University of Florence
Viale Morgagni, 59 - 50134 Florence, Italy
Tel. +39 055 423 7472; Fax +39 055 422 3560
E-mail: abiggeri@ds.unifi.it

A crucial point, that does not depend on the kind of study design, is the specification of the function describing risk decay by distance from the source. Diggle (1990) discusses some solutions and the intrinsic difficulties in estimating model parameters due to flatness of the likelihood function. Other proposals, including directional effects, can be found in Lawson (1993).

In the last years there has been an increasing interest in modeling disease risk in relation to a point source of pollution in a Bayesian framework; see, for example, Wakefield and Morris (2001), Lawson et al. (2003) and Congdon (2003). The problem of inference sensitivity to prior distributions is raised by Wakefield and Morris (2001) for models based on

aggregated data.

In this paper we investigate how the specification of the function itself relating risk of disease with distance from the source and how the choice of prior distributions on the parameters of the distance function affect the results. Without loss of generality we have restricted our attention to case-control data. A Bayesian approach to sensitivity analysis is used.

Our motivating example came from a population based case-control study on incident pleural malignant mesotheliomas in the area of Casale Monferrato, 1987-1993. Casale is a medium-size town in the North-West of Italy where a large asbestos cement (AC) plant was active from 1907 to 1985. A full description of the data and a detailed analysis can be found in papers by Magnani et al. (2001) and by Maule et al. (2007).

In the present re-analysis, information on age, sex, occupational AC exposure, occupational AC exposure of any relatives and domestic exposure to asbestos material were taken into account to adjust for known risk factors as potential confounders.

We considered four different distance functions, namely (i) exponential decay with threshold, (ii) peaked effect, (iii) plateau effect, and (iv) fixed plateau effect, to model excess risk by distance from putative source.

Materials and methods

Data

The data came from a population-based case-control study that collected subjects with pleural malignant mesothelioma newly diagnosed between 1 January 1987 and 30 June 1993 among residents in the Casale Monferrato area, comprising 52 towns and over 100,000 inhabitants (40,000 of whom from Casale). In total, 103 cases and 271 controls, are included in the present re-analysis.

Cases were retrospectively identified in the archives of the pathology units of the hospitals serving the study area and were all histologically confirmed. Controls were randomly selected either from the files of residents in the local health authority or from the mortality files of residents in the same area, and individually matched to cases by sex, birth date (± 18 months), vital status, and date of death (± 6 months). Individual matching was disregarded since in our study matching variables were spatially-neutral (Cuzick and Edwards, 1990; Diggle, 2003).

Alive subjects and the closest relative of deceased subjects were interviewed from 1993 to 1995 using a standardized questionnaire. In particular the lifelong occupational history of the subjects, their spouses, relatives or any other cohabitants, demographic characteristics, smoking habit, radiation treatments eventually received, schools attended and information on the presence and use of asbestos materials in the house or its proximities were collected.

Three main sources of asbestos exposure were identified:

(i) occupational exposure in the AC industry;
(ii) domestic exposure, with which we refer to either the indoor presence of asbestos materials such as asbestos fabrics of ironing tables, fireproof sheets for stoves and ovens, or AC materials and roofing in very close proximity to the house (e.g. garden, courtyard); and
(iii) occupation in the AC industry of relatives and cohabitants.

These variables were coded as dichotomic ("exposed"/"not exposed"). Occupational exposure in the AC industry was chosen as a proxy to asbestos occupational exposure tout court because it corresponds to very high intensity of exposure and is highly specific. Besides AC production and activities related to it (warehousing and transportation of raw asbestos and final products), no other noticeable sources of asbestos exposure of industrial origin were recorded in Casale (Magnani et al., 2001). Therefore, confounding due to residual occupational exposure is unlikely.

In the present re-analysis asbestos exposure information (occupational AC exposure, occupational AC exposure of any relatives and domestic exposure to asbestos material) were considered together with

Table 1. Data: cases and controls cross-classified by sex and occupational AC exposure (asb), occupational AC exposure of any relatives (rel) and exposure to domestic asbestos material (dom).

	Sex		asb		rel		dom	
	Males	Females	No	Yes	No	Yes	No	Yes
Controls	166	105	249	22	257	14	173	98
Cases	60	43	79	24	75	28	52	51

information on age and sex. The data are provided in Table 1.

Environmental exposure was assessed by residential distance from the source. Since this is the focus of the study, special care was dedicated to the questionnaire's section designed to reconstruct the complete residential history of all subjects, comprising all the addresses held by subjects (within and outside Casale), and a descriptions of each dwelling and its neighbourhood environment. All the residential addresses obtained from the original questionnaires were compared with, and completed by, information from the town office registries, and coded as coordinates using a global positioning system (GPS) receiver. The geographical coordinates of the AC plant location were determined in the same way. Since each subject had inhabited more than one dwelling, the address of the longest-held residence was chosen as a proxy to residential distance exposure, after exclusion of dwellings occupied in the last 20 years before the date of diagnosis for cases, or before the date of the interview or the date of death for alive and deceased controls, respectively. The spatial distribution of cases and controls and of the AC plant is shown in Figure 1.

More details of the case-control study design were described by Magnani et al. (2001) and by Maule et al. (2007).

Statistical analyses

We assumed that cases and controls are a random sample from a marked point process, the mark indicating case-control status. Conditionally on the observed locations of cases and controls, i.e. for subject i (i = 1, . . . , 374) being resident at distance d_i from the putative source, a logistic regression model was specified, modeling the log odds of being a case (Diggle and Rowlingson, 1994; Biggeri and Lagazio, 1999).

In detail, we specified an additive-multiplicative logistic model:

$$Y_i \sim \text{Binominal}\ (\pi_i, 1)$$

where Y_i is the case-control indicator (Y_i = 1 for cases), π_i is the probability of being a case and

$$\frac{\pi_i}{1-\pi_i} = \exp(\omega_0+\omega_1 \text{sex}_i+\omega_2 \text{age}_i+\omega_3 \text{dom}_i+\omega_4 \text{rel}_i)[1+\omega_5 \text{asb}_i+f(d_i)]$$

where $\exp(\omega_0)$ is a constant term proportional to case-control ratio and ω_j (j = 1, . . .,5) is the not

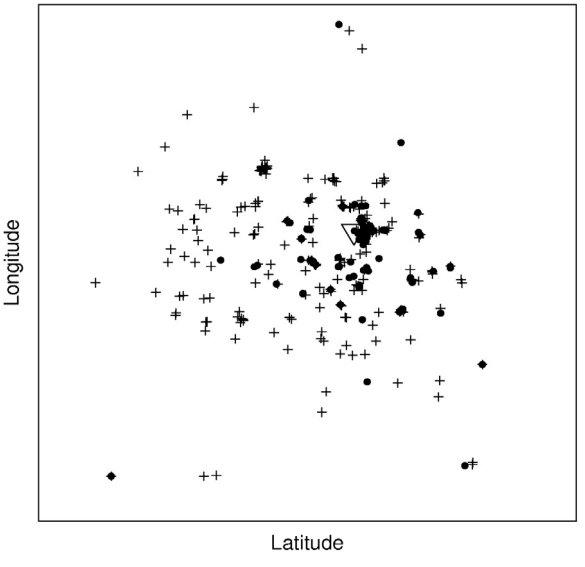

Fig. 1. Cases (•), controls (+) of residence locations and asbestos cement plant source (▽).

exponentiated effect parameter for the *j*-th risk factor. Notice that sex (female is the reference level), age (continuous), domestic exposure to asbestos material (dom, binary) and occupational AC exposure of any relatives (rel, binary) are introduced as multiplicative terms, while occupational AC exposure (asb, binary) is introduced as an additive term in the excess risk function. The effect of distance is included in the additive term, too. The function $f(\cdot)$ describes the decay of risk with increasing distance. For binary variables we have used "not exposed" as reference category.

The rationale of this modeling choice is that different exposures to the same agent sum up (i.e. combine additively, see Breslow, 1986). We considered here the two main risk factors (occupation and environmental exposure) in the additive term of the predictor. The other covariates are left in the multiplicative term, even if it can be argued that for mesotheliomas the only relevant exposure is asbestos. For environmental exposure measured by residential distance the motivation to choose an additive specification is also to ensure that the risk be unchanged at infinite distance from the source. The distance function is obviously not increasing and tends asymptotically to zero when distance goes to infinity. We specified four different functional forms for $f(\cdot)$:

(i) Model 1, an exponential decay with threshold (Diggle and Rowlingson, 1994)

$$f(d_i) = \alpha \exp(-\beta d_i^2)$$

where α is the excess relative risk at source and β the parameter of the exponential decrease function;

(ii) Model 2, a peaked effect (Lawson, 1993)

$$f(d_i) = \alpha \exp(\gamma \log(d_i) - \beta d_i)$$

where γ is the parameter which models the distance at which we have the maximum risk;

(iii) Model 3, an estimated plateau effect (Diggle et al., 1997)

$$f(d_i) = \begin{cases} \alpha & \text{if } d_i \leq \delta \\ \alpha \exp(-\beta(d_i-\delta)^2) & \text{if } d_i > \delta \end{cases}$$

where δ represents the radius of the plateau, to be estimated from data; and

(iv) Model 4, a fixed ($\delta = 5$) plateau effect

$$f(d_i) = \begin{cases} \alpha & \text{if } d_i \leq 5 \\ \alpha \exp(-\beta(d_i-5)^2) & \text{if } d_i > 5 \end{cases}$$

where the distance 5 km depends on the physical characteristic of the environment and the information based on environmental surveys (Magnani et al., 2001).

We first explored the characteristics of the likelihood function under the different specifications. Figure 2 reports the profile log-likelihood in proximity of the maximum. In order to map the model into the framework of Bayesian inference, we had to specify prior distributions for the parameters. Flat normal distributions centered on zero were used for ω_0, ω_1, ω_2, ω_3 and ω_4; for the parameter ω_5 a flat left truncated normal distribution was used. Informative priors for the distance functions parameters were assumed:

(i) for the parameter related to the excess risk at the source $\alpha \sim$ gamma (9, 1), with interquartile range 6.8-10.8, percentiles 1% = 3.5 and percentiles 99% = 17.4; the maximum likelihood (ML) estimates was 9.3. Maule et al. (2007) discuss the result of high risk for environmental exposure in Casale;

(ii) for the parameter related to risk decay $\beta \sim$ uniform (0, 3), the upper limit large enough for the purpose to give prior weight to rapid decay by distance;

(iii) for the parameter related to peaked risk in the decay function $\gamma \sim$ gamma (1, 1), which implies a substantial excess risk of distances >1 km from the source with a peak around 3-5 km (the function gives values around 12, using for

the other coefficient values at ML estimates). The adjusted odds ratio at distance 3-5 km was 12.1 according to Maule et al. (2007); and

(iv) for the parameter related to the plateau radius $\gamma \sim$ uniform (0, 88). Here we left the maximum possible radius as upper limit. For the model with fixed plateau effect (Model 4) we put a strong informative prior ($\delta \sim$ gamma (50, 10); mean 5, interquartile range 4.5-5.5, 10% and 90% percentiles 4.1 and 5.9).

To evaluate the sensitivity of inference to prior assumptions, we repeated the analyses using a new set of prior distributions, only slightly different from the previous one: $\alpha \sim$ gamma (5, 1), which assigns non-negligible probability to very low values of excess risk at the source (i.e. the environmental exposure is less important); $\beta \sim$ uniform (0, 10), giving even more weight to a rapid decay of risk by distance (i.e. environmental risk is concentrated); $\gamma \sim$ gamma (2, 1), which assigns less probability to excess risk at very short distance (<0.5 km) and greater probability to excess risk around 8 km from the source (i.e. there is clustering of risk at a given distance from the source).

The marginal posterior distributions of the parameters of interest were approximated by Markov chain Monte Carlo (MCMC) methods, by using WinBUGS (Spiegelhalter et al., 2004). The convergence of the algorithm was evaluated using the test proposed by Gelman and Rubin (1992) for multiple chains for all monitored parameters. We decided to discard the first 500,000 iterations (burn-in) and to store for estimation 5,000 samples (one each 100) of the following 500,000 iterations. Maximum-likelihood estimates were obtained using R software.

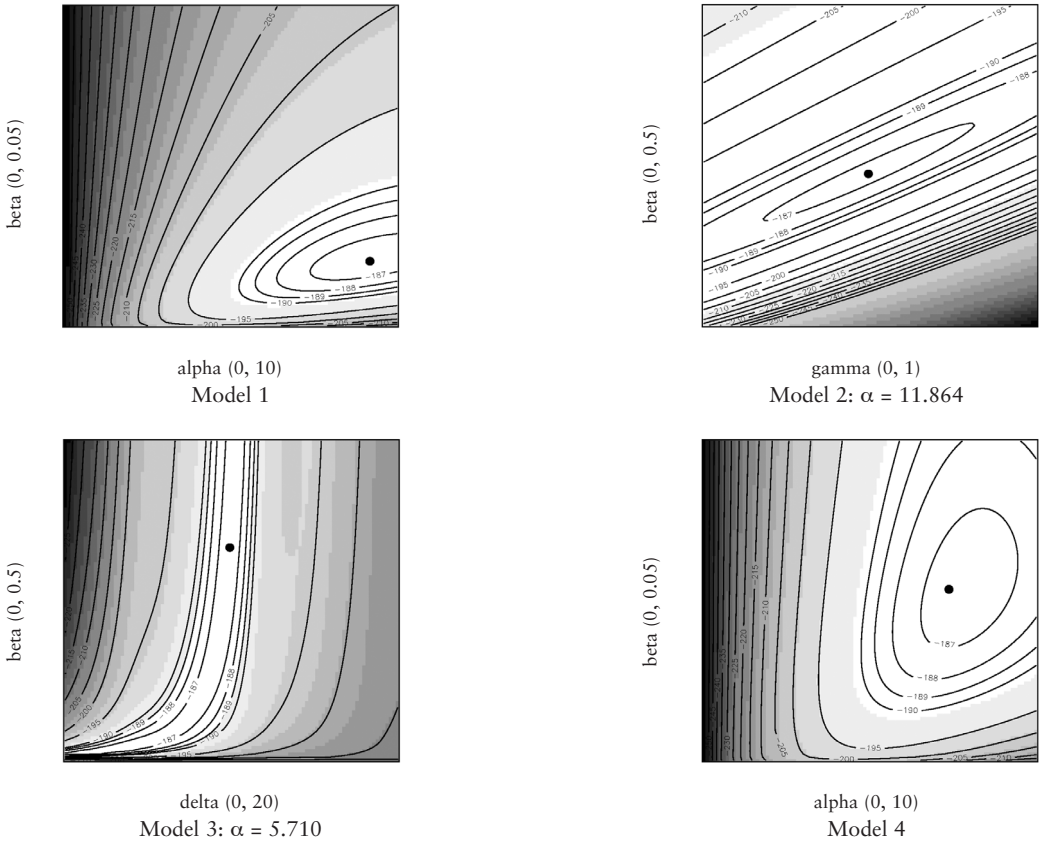

Fig. 2. Contours of the log-likelihood functions of the four models (limits for parameter values are indicated in parenthesis).

Results

The four plots of Figure 2 show a markedly flat area in the profile likelihood, indicating that the range of plausible values for the parameters (or for a subset of them) is very large. For Model 1, the bivariate profile likelihood for the excess risk and risk decay parameters has a flat area with very high plausible values for the α term. For Model 2, fixing α, the profile for the two parameter in the risk decay function highlights the strong correlation between them. For Model 3, fixing α, the profile has J-shaped form. It shows a strong uncertainty in the β term for δ around the ML value of 10 and for

Table 2. Parameter estimates (exponentiated values for the multiplicative part of the model) with maximum likelihood (ML) and Bayesian (first set of priors, B1) analysis.

Effect	Model 1		Model 2		Model 3		Model 4	
	ML	B1	ML	B1	ML	B1	ML	B1
Constant	0.040	0.050	0.036	0.052	0.058	0.055	0.046	0.051
Age	1.000	1.002	1.001	1.003	0.999	0.999	1.000	1.001
Sex	1.165	1.210	1.174	1.227	1.170	1.212	1.160	1.213
rel	1.371	1.431	1.380	1.440	1.370	1.439	1.380	1.449
dom	1.451	1.555	1.440	1.571	1.560	1.661	1.512	1.576
asb	25.81	26.90	28.50	25.64	18.85	26.38	23.03	26.450
α	9.300	9.007	11.864	9.440	5.710	7.779	7.342	8.124
β	0.010	0.014	0.230	0.296	0.335	1.291	0.026	0.173
γ			0.490	0.622	9.935			
δ						10.350		5.117
AIC	388.76		389.54		388.05		388.91	
DIC		387.42		389.15		388.72		390.06

Table 3. Parameter estimates (exponentiated values for the multiplicative part of the model) with the first (B1) and the second (B2) set of prior distributions under the Bayesian analysis.

Effect	Model 1		Model 2		Model 3		Model 4	
	B1	B2	B1	B2	B1	B2	B1	B2
Constant	0.050	0.068	0.052	0.073	0.055	0.077	0.051	0.083
Age	1.002	1.002	1.003	1.002	0.999	0.999	1.001	1.003
Sex	1.210	1.211	1.227	1.233	1.212	1.223	1.213	1.257
rel	1.431	1.586	1.440	1.598	1.439	1.688	1.449	1.543
dom	1.555	1.422	1.571	1.449	1.661	1.424	1.576	1.412
asb	26.90	21.39	25.64	20.56	26.38	20.97	26.450	17.58
α	9.007	6.177	9.440	6.139	7.779	5.180	8.124	4.356
β	0.014	0.017	0.296	0.407	1.291	4.763	0.173	2.680
γ			0.622	1.071				
δ					10.350	10.87	5.117	5.191
DIC	387.42	388.34	389.15	391.07	388.72	388.93	390.06	395.60

defined values of β (around 0.02) large uncertainty for δ <10. This means that there is information either in a risk function with stable plateau up to 10 km and a rapid decay thereafter, or in a risk plateau with a small radius and a slow decay afterward. For Model 4, fixing δ = 5, there is a large uncertainty about the value of the decay parameter β, with plausible large values for α.

This, in turn, indicates that the results by a Bayesian approach should be very sensitive to modeling strategies; both in terms of a priori specification and of distance function used (Wakefield and Morris, 2001).

ML estimates of the parameters and posterior means from Bayesian analysis with the first set of prior distributions are reported in Table 2, including AIC and DIC criteria (Spiegelhalter et al., 2002).

There is little information to discriminate between the four models. Models 1 and 3 seem slightly better than the other two models. The ML estimates of the α and β terms for Model 3 are very unstable, as compared to the Bayesian posteriors means.

Figure 3 shows the estimated $f(d_i)$ functions corresponding to Models 1-4 (using the first set of prior). Models 1 and 2 show a very similar decay with distance after the peak of Model 2, while the shape implied by Models 3 and 4 is very different.

In Table 3 results for the second set of prior distributions for the distance function parameters are summarised and compared with those given by the first set. For Model 1, giving less probability to environmental risk by point source, the α term resulted in posterior estimate around 6.2 (*versus* 9.0), a value yet very high. For Model 2 we loose goodness of fit (DIC increases), the parameter estimated values tend to compensate each other. For Model 3, the posterior mean for the α term decreases but now the risk decay term assumes very high value, compatible to a step risk function. Not surprisingly, if the excess risk at the source is higher, there is room for smoothed risk decrease. For Model 4, the behaviour is similar, but there is a lack of fit constraining the plateau at 5 km.

Figures 4, 5 and 6 show the prior and the posterior distributions of the distance functions terms α, γ and δ.

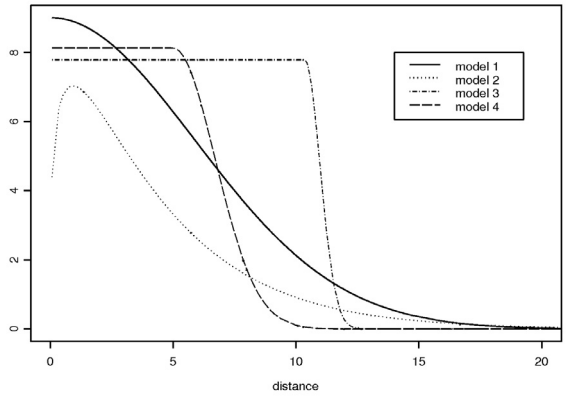

Fig. 3. Estimated $f(d_i)$ functions with the Bayesian approach, first set of priors for α, β and γ.

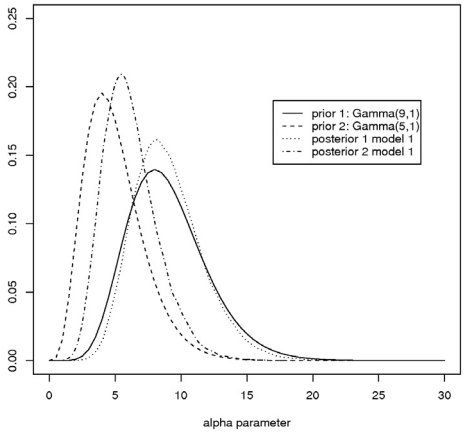

Fig. 4. Prior and posterior distributions for α parameter by Model 1.

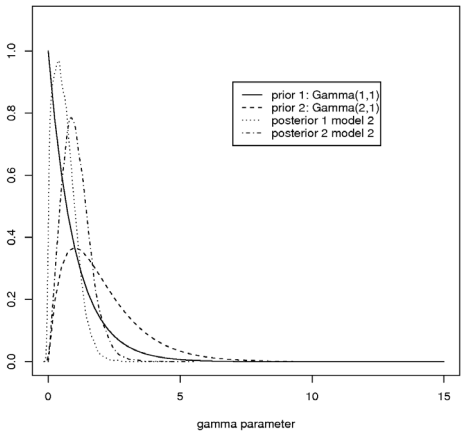

Fig. 5. Prior and posterior distributions for γ parameter by Model 2.

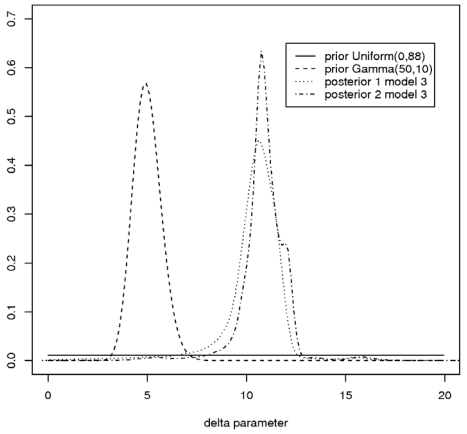

Fig. 6. Prior distributions for δ for Models 3 and 4 and posterior distribution by Model 3.

Discussion

The statistical analysis of the relationship between risk of disease and distance from a putative source of pollution is characterized by two critical points, that are strictly related.

First, it is difficult to identify a specific functional form. This is because usually data contain scarce information about (Diggle et al., 1997). The number of subject locations close to the point source is of course much less than the number of locations far from the source. The distance functions differ in the first part, close to the source, where information is scarce. Moreover, the model is non-linear and the parameters are strongly correlated. In the case of logistic regression the model is not multiplicative. Non-canonical link functions can create problems since sufficiency is not guaranteed (for examples, see Moolgavkar and Venzon, 1987). As a result, the likelihood function is flat and this is related to the second point.

Second, Bayesian inferences could be strongly dependent on the priors. When fitting non-linear models, the Bayesian approach has clear advantages in stability of estimates, and informative priors are a resource. In the example discussed here, sensible priors are easy to specify and to justify. On the contrary uninformative priors, if any, could be misleading (Biggeri et al., 2004). Inference sensitivity is an important part of the analysis. We compared two set of priors for the four models considered. Subject specific knowledge is an important issue for guiding in the choice of prior parameters and setting alternatives. See, for example, the excess risk at the source parameter. If environmental pollution is really a hazard we would expect a high value of risk at the source. Maule et al. (2007) discussed that environmental asbestos exposures gave a risk of about one-third (33%) of the occupational exposure risk, hence an important fraction. This is based on the estimated α around 9.0 and of the asb coefficient of 26.9. The Bayesian sensitivity analysis reports α values between 7.8 and 9.4 under the first prior, depending on the model, and between 4.4 and 6.2 using the second prior. Averaging over the fitted model gives 7.02. The alternative specifications are more conservative and as a result the sensitivity analysis point to an environmental risk of about 30% the occupational one (7.02 *versus* 23.2).

A possible development would be applying Bayesian model averaging (Hoeting et al., 1999). However, it could be argued that some of the models are not plausible enough and we do not aim to one overall estimate, but rather aim to evaluate and compare alternatives. How does the plateau model in the example with fixed risk distance at 5 km behave? Although the model fit is the worst among the fitted ones, this is still valuable information. Data support is weak and we can appreciate on the contrary that the risk is high up to 10 km from the source (Magnani et al., 2001).

An attempt to identify point-centered clusters and line clusters without taking into account the source location is reported by Lawson et al. (2007). The authors found evidence of one point-centered cluster located at the AC plant (the source considered in our example here). Weak evidence of two line clusters was reported, and the posterior density of line cluster centers mainly located within the city. Coherently, a simpler non-parametric cluster model highlighted a broad raised risk area. A further refinement would be to look for line clusters having

taken into account for the risk by distance from the source, using the parametric threshold Model 1.

Directional effects were excluded. Figure 2 in Maule et al. (2007), however suggest that the cluster center could be shifted from the plant location. We did not model this displacement.

In conclusion, we suggest to base the choice of the prior distributions on epidemiological knowledge. The point source models are ill-conditioned and we have to avoid flat or uninformative priors. ML estimates could not be obtained in some dataset. A Bayesian approach is important to sensitivity analysis of appropriate alternatives.

Acknowledgements

The research was partially supported by COFIN-2004 (PRIN 2004137478) and COFIN-2006 (PRIN 2006131039).

References

Biggeri A, Dreassi E, Lagazio C, Marchi M, 2004. Bayesian focused clustering for a case-control study on lung cancer in Trieste. Statistical Modelling, Proceeding of the 19th Internalional Workshop on Statistical Modelling (IWSM). Biggeri, Dreassi, Lagazio and Marchi (eds). Firenze University Press, Italy, 91-95.

Biggeri A, Lagazio C, 1999. Case-control analysis around putative sources. In: Disease Mapping and Risk Assessment for Public Health. Lawson, Bertollini, Biggeri, Böhning, Lesaffre and Viel (eds). Wiley, London, UK, pp. 271-286.

Breslow N, 1986. Use of the power transform to discriminate between additive and multiplicative models in epidemiologic research. In: Modern Statistical Methods in Chronic Disease Epidemiology. Moolgavkor and Prentice (eds). Wiley & Sons Ltd, USA, pp. 181-196.

Congdon P, 2003. Applied Bayesian Modelling. John Wiley & Sons Ltd, USA.

Cuzick J, Edwards R, 1990. Spatial clustering for inhomogeneous populations. J R Stat Soc Ser A 157, 433-440.

Diggle PJ, 1990. A point process modelling approach to raised incidence of a rare phenomenon in the vicinity of a prespecified point. J R Stat Soc Ser A 153, 340-362.

Diggle PJ, 2003. Statistical Analysis of Spatial Point Patterns, 2nd edition. Arnold Publishers, London, UK.

Diggle PJ, Morris S, Elliot P, Shaddick G, 1997. Regression modelling of disease risk in relation to point sources. J R Stat Soc Ser A 160, 491-505.

Diggle PJ, Rowlingson BS, 1994. A conditional approach to point process modelling of elevated risk. J R Stat Soc Ser A 153, 349-362.

Gelman A, Rubin DR, 1992. Inference from iterative simulation using multiple sequences (with discussion). Stat Sci 7, 457-511.

Hoeting JA, Madigan D, Raftery AE, Volinsky CT, 1999. Bayesian model averaging. A tutorial. Stat Sci 14, 382-417.

Lawson AB, 1993. On the analysis of mortality events associated with a prespecified fixed point. J R Stat Soc Ser A 156, 363-377.

Lawson AB, Browne WJ, Vidal Rodeiro CL, 2003. Disease mapping with WinBugs and MlWin. Chichester. John Wiley & Sons Ltd, USA.

Lawson AB, Simeon S, Kulldorff M, Biggeri A, Magnani C, 2007. Line and point cluster models for spatial health data. Comput Stat Data Anal 51, 6027-6043.

Magnani C, Dalmasso P, Biggeri A, Ivaldi C, Mirabelli D, Terracini B, 2001. Increased risk of malignant mesothelioma of the pleura after residential or domestic exposure to asbestos: a case-control study in Casale Monferrato, Italy. Environ Health Perspect 109, 915-919.

Maule MM, Magnani C, Dalmasso P, Mirabelli D, Merletti F, Biggeri A, 2007. Modeling mesothelioma risk associated with environmental asbestos exposure. Environ Health Perspect 115, 1066-1071.

Moolgavkar SH, Venzon DJ, 1987. General relative risk regression models for epidemiologic studies. Am J Epidemiol 126, 949-961.

Spiegelhalter DJ, Best NG, Carlin BP, van der Linde A, 2002. Bayesian measures of model complexity and fit (with discussion). J R Stat Soc Ser B 64, 583-616.

Spiegelhalter DJ, Thomas A, Best N, Lunn D, 2004. WinBUGS User Manual, Version 1.4.1. (On-line user manual: http://www.mrcbsu.cam.ac.uk/bugs: September 13th, 2004).

Wakefield JC, Morris SE, 2001. The Bayesian modelling of disease risk in relation to a point source. J Am Stat Assoc 96, 77-91.