Multivariate spatially-structured variability of ovine helminth infections

Annibale Biggeri[1,2], Dolores Catelan[1,2], Emanuela Dreassi[1], Laura Rinaldi[3], Vincenzo Musella[3], Vincenzo Veneziano[3], Giuseppe Cringoli[3]

[1]Department of Statistics "G. Parenti", University of Florence, Florence, Italy; [2]Biostatistics Unit, CSPO, Florence, Italy; [3]Department of Pathology and Animal Health, University of Naples "Federico II", CREMOPAR Regione Campania, Naples, Italy

Abstract. A cross-sectional survey was carried out on 2004-2005 in the Campania region, southern Italy, to study the multivariate geographical distribution of four different sheep helminths, i.e. *Fasciola hepatica* (liver fluke), *Calicophoron* (*Paramphistomum*) *daubneyi* (rumen fluke), *Dicrocoelium dendriticum* (lancet fluke), and the gastro-intestinal strongyle *Haemonchus contortus*. A series of multivariate Bayesian hierarchical models based on square root transformation of faecal egg counts were performed. The results were consistent with theoretical knowledge of the biology and epidemiology of the four studied helminths. In particular, the impact of common intermediate hosts (*F. hepatica* and *C. daubneyi* share the same intermediate host species) was quantified and evidence of previously unknown ecological components was given. *D. dendriticum* was correlated to *F. hepatica* and *H. contortus* was found not to be spatially associated with the previously mentioned helminths.

Keywords: disease mapping, shared component models, hierarchical Bayesian model, veterinary epidemiology, helminths.

Introduction

Multivariate disease mapping in the context of hierarchical Bayesian modelling is a relatively new approach which is able to estimate latent common patterns of risk, and eventually relate them to the underlying prevalence of exposure. Most applications relate to cancer epidemiology (see for example Knorr-Held and Best, 2001) but there is growing appreciation for other public health and veterinary applications, particularly for parasitic infections.

Parasites are strongly dependent on environmental characteristics and ecological analysis can address relevant scientific questions (e.g. Biggeri et al., 2006; Raso et al., 2006; Gosoniu et al., 2007).

Fasciola hepatica (liver fluke), *Calicophoron* (*Paramphistomum*) *daubneyi* (rumen fluke), *Dicrocoelium dendriticum* (lancet fluke) and the gastro-intestinal (GI) strongyle *Haemonchus contortus* are widely distributed in small ruminants bred in southern Italy (Cringoli et al., 2004b, 2007b; Rinaldi et al., 2005; Maurelli et al., 2007). In addition, it is well known that:

(i) *F. hepatica* and *C. daubneyi* have the same indirect life cycle and share the same intermediate host (the amphibious snail *Lymnaea truncatula*) (Mage et al., 2002);

(ii) *D. dendriticum* has a complex life cycle involving numerous species of land snails and ants as first and second intermediate hosts, respectively (Maurelli et al., 2007); and

(iii) *H. contortus* has a direct life cycle without intermediate hosts.

The present paper was aimed at studying the degree of spatial correlation among these four

Corresponding author:
Dolores Catelan
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: catelan@ds.unifi.it

helminths in 121 ovine farms of the Campania region, southern Italy. We account for their known ecology and we utilize data from a recent (2004-2005) cross-sectional survey.

Materials and methods

Study area and sampling

The survey was carried out in the Campania region of southern Italy (latitude = 39°59'15"-41°30'25"; longitude = 13°45'25"- 15°48'23") which extends over an area of 13,590 km^2. The region is mainly hilly and extends from 0 to 1890 m above sea level. The climate is Mediterranean with dry summers and rainy winters.

A geographical information system (GIS) was established utilizing the administrative boundaries (at provincial and municipal levels) of the Campania region. In addition, as part of the regional project MAPZOO (Cringoli et al., 2007a), all the sheep farms of the region were geo-referenced. In order to uniformly sample the farms throughout the entire region, a grid, representing quadrants of 10 x 10 km, was overlaid on the region map within the GIS. As a result, the territory of the Campania region was divided into 135 equal quadrants, the centroid of each quadrant was identified and among all the farms included in the GIS database, the farm closest to the centroid in each quadrant was selected. Only pastured farms with at least 50 sheep were considered. The GIS software used for the present study was ArcGIS version 9.2 (ESRI, Redlands, CA, USA). Out of the total 135 quadrants, 121 were investigated. The remaining 14 were non-farmed areas. Faecal samples were collected from 15 adult sheep (aged >18 months) and 5 youngs (aged ≤18 months) on each farm. In total, 2420 sheep were examined.

Parasitological analyses

Faecal samples were pooled at the laboratory in order to obtain 4 composites of 5 individual samples of the same weight, for each tested farm.

The FLOTAC technique (Cringoli, 2006) with an analytic sensitivity of 2 eggs per gram of faeces (EPG), was utilised for all coprological examinations. The flotation solutions employed were: a sucrose flotation solution (specific gravity (s. g.) = 1.250) for detecting GI strongyle eggs, and zinc sulphate plus potassium iodomercurate (s. g. = 1.450) for detecting trematoda eggs (Cringoli et al., 2004a).

In addition, coprocultures were performed and *H. contortus* third-stage larvae (L3) were identified using the morphological keys proposed by van Wyk et al. (2004).

Statistical analysis

We addressed data analysis on the square root transformation of egg counts Y^*_{ik}. Let assume that Y^*_{ik} follows a multivariate normal distribution $MVN_k (\mu, \Sigma)$ where $i = 1,..., N$ indexes farms and $k = 1,..., K$ indexes helminths. This assumption is appropriate since a certain degree of correlation among parasites is expected.

We then specified three different models for μ, as follows:

(i)   model 1: $\mu_{ik} = \alpha_k$; $\alpha_k \sim N(0,1000)$, no spatial structure was assumed;

(ii) model 2: $\mu_{ik} = \alpha_k + u_{ik}$; where u_{ik} is a helminth-specific clustering term with conditional autoregressive (CAR) specification (Besag et al., 1991); and

(iii) model 3: $\mu_{ik} = \alpha_k + u_i \delta_k + I (k \in \Gamma) \Psi_{ik}$; where u_i is a shared clustering term between the four helminths modelled as CAR; Ψ_{ik} represents a specific clustering term for *D. dendriticum* or *H. contortus* by Γ, an appropriate subset of K. This choice is motivated by the ecology of the parasites, two intermediate hosts for *D. dendriticum* and three external larval stages for *H. contortus*.

Following Knorr-Held and Best (2001), the terms $\log\delta_1, ..., \log\delta_4$, constrained to $\Sigma^4_{k=1} \log\delta_k$, were assumed to be multivariate normal distributed with zero mean and variance covariance matrix, respectively, as follows:

$$\Sigma_\delta = \sigma_\delta^2 \begin{bmatrix} 1 & -1/3 & -1/3 & -1/3 \\ -1/3 & 1 & -1/3 & -1/3 \\ -1/3 & -1/3 & 1 & -1/3 \\ -1/3 & -1/3 & -1/3 & 1 \end{bmatrix}$$

In models 2 and 3, α_k is the helminth-specific intercept whose *a priori* distribution is improper uniform.

The fit of these models was described through Pearson's correlation among predicted responses and among spatially-structured random terms (model 2 *versus* model 3). Model comparison was done by deviance information criterion (DIC; Spiegelhalter et al., 2002).

Results

Results of exploratory analysis

The parasitological results showing prevalence and infection intensities (EPG) for each studied helminth are summarised in Table 1.

With the exception of *F. hepatica* and *C. daubneyi* that exhibited a positive correlation, the other helminth infections appeared to be uncorrelated (Table 2).

In Figure 1 the geographical distribution of observed EPGs for the 4 helminths studied is shown.

Pearson's correlation coefficients among predicted responses by models 1 and 2 are shown in Table 3, i.e. the correlation among responses by models in which spatial (model 2) or non-spatial structure is present.

Both models showed a high and positive correlation between *F. hepatica* and *C. daubneyi*. With regard to the other helminths, the results differed between the two models but the one with the spatially-structured term seemed to better reproduce the observed scenario. DIC confirmed these findings (2563 in model 1 *versus* 2415 in model 2), the smaller the better fit to the data.

The risk surfaces for the spatially-structured terms of model 2 are displayed in Figure 2. A strong spatial gradient emerged which, as expected, was very similar for *F. hepatica* and *C. daubneyi*. The surface for *D. dendriticum* shows a spatial pattern that was surprisingly close to that of *F. hepatica*. Both flukes affect the liver but they have different seasonal patterns which could explain co-existence in a common environment of intermediate hosts (Asanji and Williams, 1984). *H. contortus* infection showed a higher prevalence along the coastal region and there was no correlation with *C. daubneyi*.

The shared clustering model 3 showed a lower DIC than model 2, and hence a better fit of the data.

Table 1. Ovine helminth infections in 121 sheep farms in the Campania region, southern Italy: descriptive analysis.

	Helminths studied in the 121 sheep farms investigated			
	F. hepatica	*C. daubneyi*	*D. dendriticum*	*H. contortus*
No. positive farms	15	17	81	109
% positive farms (95% CI*)	12.4 (7.3-19.9)	14.0 (8.6-21.8)	66.9 (57.7-75.1)	90.1 (82.9-94.5)
Mean EPG	4.2	3.5	33.1	79.0

*CI = confidence interval

Table 2. Ovine helminth infections in 121 sheep farms in the Campania region, southern Italy: Pearson's correlation coefficients between observed infection prevalence data.

Helminth species	*F. hepatica*	*C. daubneyi*	*D. dendriticum*
C. daubneyi	0.2768	-	-
D. dendriticum	-0.0802	0.0876	-
H. contortus	-0.0558	0.0681	-0.0052

Fasciola hepatica

Calicophoron daubneyi

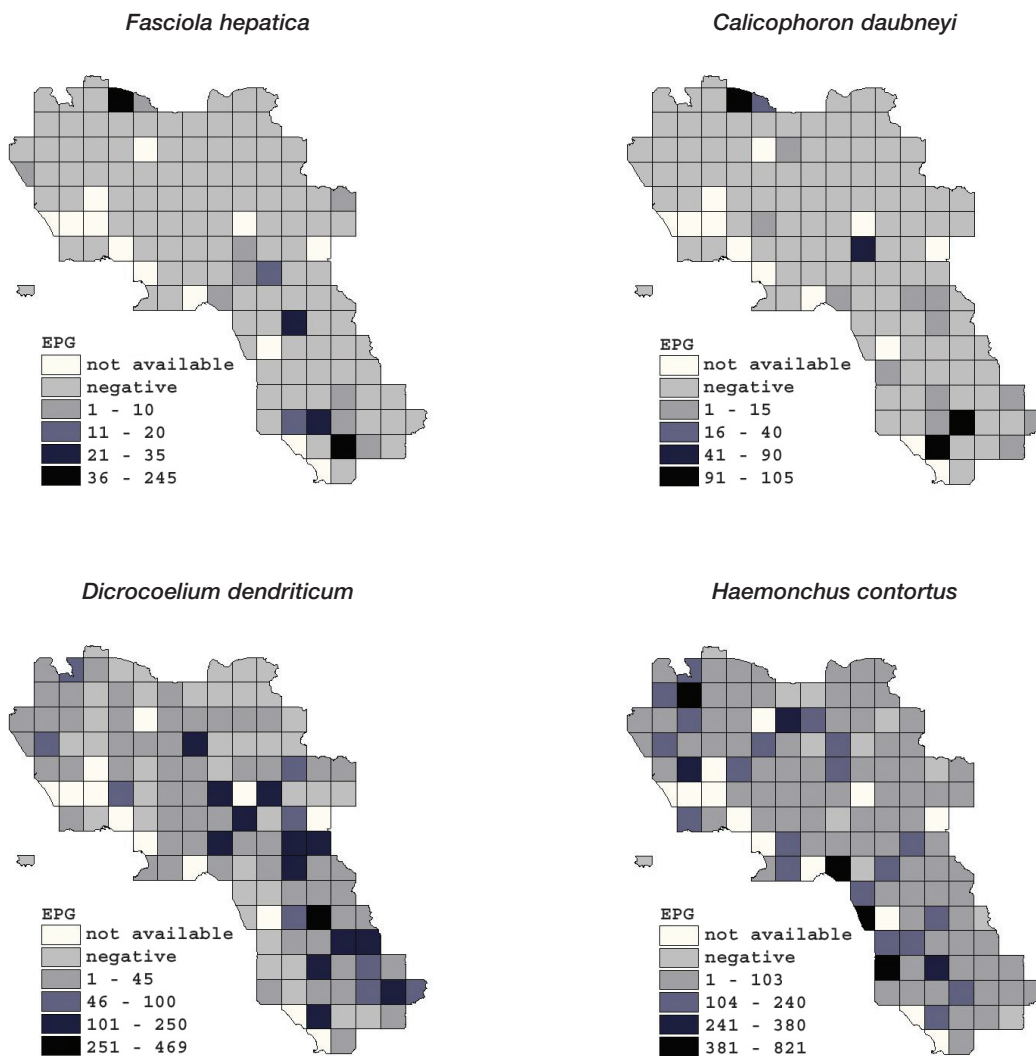**Dicrocoelium dendriticum**

Haemonchus contortus

Fig. 1. Maps of ovine helminth infections in the Campania region, southern Italy (colour scale by count of EPG).

Table 3. Ovine helminth infections in 121 sheep farms in the Campania region, southern Italy: Pearson's correlation coefficients among predicted responses and DIC values comparing model 1 with model 2.

Helminth species	Model 1			Model 2		
	F. hepatica	*C. daubneyi*	*D. dendriticum*	*F. hepatica*	*C. daubneyi*	*D. dendriticum*
C. daubneyi	0.42	-	-	0.32	-	-
D. dendriticum	-0.30	-0.28	-	-0.19	-0.03	-
H. contortus	0.08	-0.22	0.38	0.10	0.11	0.03
DIC*		2563			2415	

*DIC = deviance information criterion.

Fasciola hepatica

Calicophoron daubneyi

Dicrocoelium dendriticum
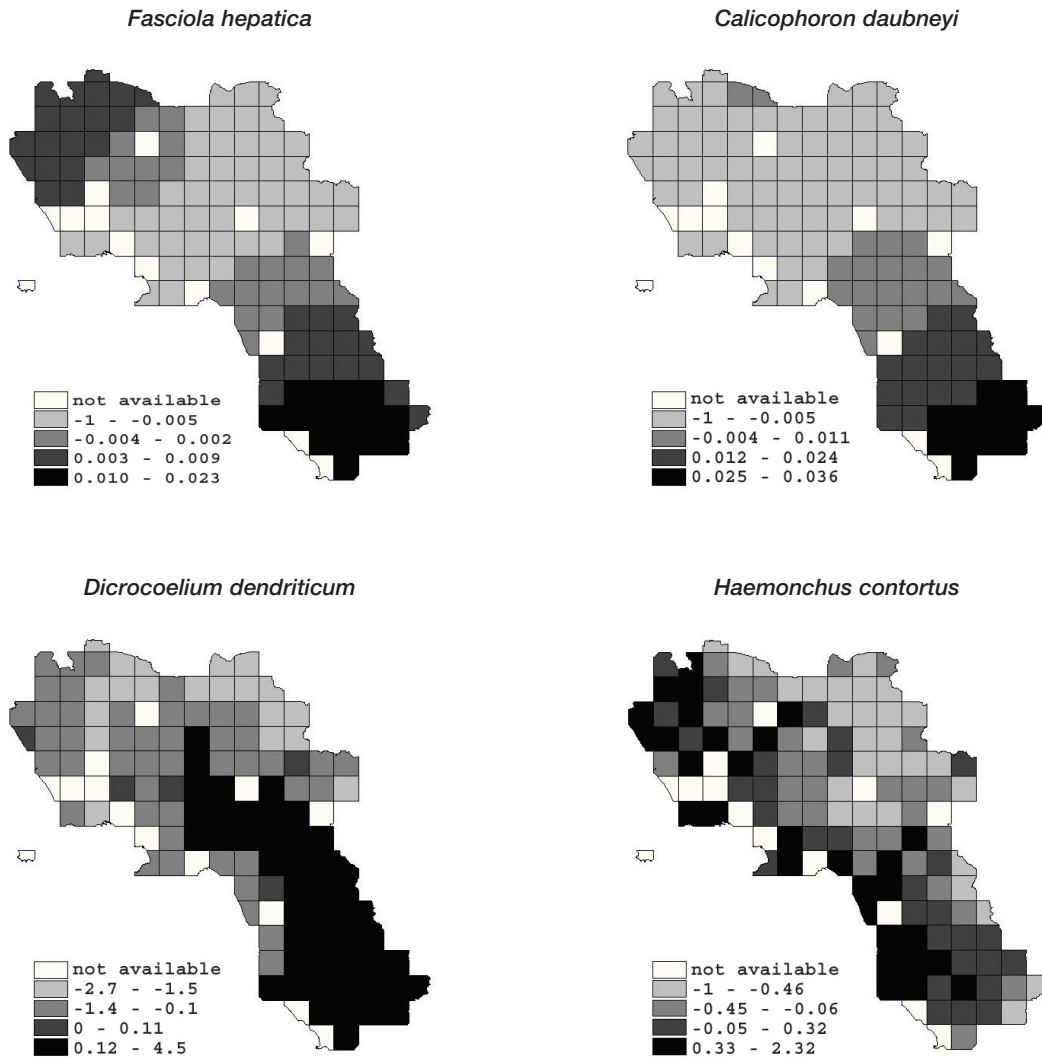
Haemonchus contortus

Fig. 2. Ovine helminth infections in 121 sheep farms in the Campania region, southern Italy: spatially-structured random terms from model 2 (see text).

Table 4. Ovine helminth infections in 121 farms in the Campania region, southern Italy: Pearson's correlation coefficients among spatially-structured terms from model 2 and δ terms from model 3 (see text).

Helminth species	F. hepatica	C. daubneyi	D. dendriticum	δ
F. hepatica	-	-	-	1.414
C. daubneyi	0.96	-	-	1.496
D. dendriticum	0.71	0.73	-	1.275
H. contortus	0.42	0.35	0.30	0.529

Shared clustering

not available
-1 - -0.005
-0.004 - 0.002
0.003 - 0.009
0.010 - 0.023

Dicrocoelium dendriticum

not available
-2.7 - -1.5
-1.4 - -0.1
0 - 0.11
0.12 - 4.5

Haemonchus contortus

not available
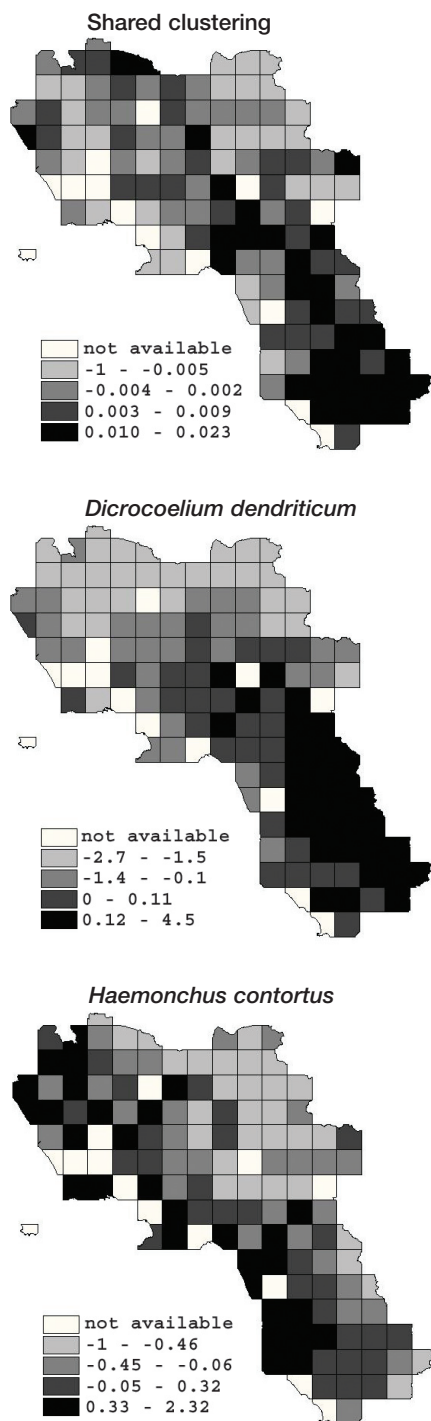-1 - -0.46
-0.45 - -0.06
-0.05 - 0.32
0.33 - 2.32

Fig. 3. Ovine helminth infections in 121 sheep farms in the Campania region, southern Italy: shared clustering terms and parasitic specific terms from model 3 (see text).

In the model, the shared term captured a major part of the association between *F. hepatica* and *C. daubneyi*. The correlation among spatially structured terms from model 2 and δ terms of model 3 are reported in Table 4. The correlation between specific clustering terms of *F. hepatica* and *C. daubneyi* was 96% but only 30% among *D. dendriticum* and *H. contortus*. Consistently, the importance of shared clustering terms was greater for *F. hepatica* and *C. daubneyi* (1.414 and 1.496, respectively) than for *H. contortus* (0.529).

The shared clustering terms and the specific clustering terms for *H. contortus* and *D. dendriticum* are reported in Figure 3. These last terms exhibited the same spatial pattern as that produced by model 2.

Discussion and conclusions

We aimed to assess the extent of spatial co-variation among the risk of different helminths in ovine farms. Liver and rumen flukes (*F. hepatica* and *C. daubneyi*) share the same intermediate hosts, i.e. amphibious snails; thus, we expected a strong spatial correlation among the two helminths. Lancet fluke (*D. dendriticum*) has two intermediate hosts, i.e. land snails and ants; thus we expected absence of correlation with the other two trematoda. *H. contortus* has no intermediate hosts; three different larval stages live in the external environment and we were interested in looking at possible correlations with the other helminths. To test these hypotheses we specified a model with shared clustering components for the four helminths and specific clustering for *D. dendriticum* and *H. contortus* (model 3).

In particular, recall that for *F. hepatica* and *C. daubneyi* assumed $\mu_{ik} = \alpha_k + u_i \, \delta_k$ where u_i is CAR modelled, and for *D. dendriticum* and *H. contortus* $\mu_{ik} = \alpha_k + u_i \, \delta_k + \Psi_{ik}$, where Ψ_{ik} is a CAR process specific for each of the two parasites.

Our analysis highlighted that:
(i) there is evidence of shared spatially-structured variability among the four helminths studied;

(ii) *F. hepatica* and *C. daubneyi* were highly correlated; and

(iii) *D. dendriticum* and *H. contortus* were not correlated and showed a specific spatial structure.

Shared terms can be interpreted as a latent covariate (Knorr-Held and Best, 2001). Therefore *F. hepatica* and *C. daubneyi* prevalence can be interpreted as surrogate or observable covariate of a latent variable (assuming a classical error model, with $z_i = u_i / \delta_k$ and $\beta_k = \delta_k^2$ the model becomes $\mu_{ik} = \alpha_k + u_i \delta_k + \Psi_{ik} = \alpha_k + \beta_k z_i + \Psi_{ik}$).

Our findings modify previous expectations since we found a common spatial structure driven by *F. hepatica* and *C. daubneyi* and not only a correlation between these two parasites. Coverage or effectiveness of veterinary surveillance is traced by the prevalence of those two preventable diseases. The shared clustering model 3 results, can be interpreted as evidence in favour of the prevalence of *F. hepatica* and *C. daubneyi* infection being a proxy for coverage of veterinary preventive or prophylactic treatments.

With regard to statistical issues, the shared clustering model adopted is appropriate when spatially-structured variability is the subject-specific goal. Alternatively, the generalized multivariate CAR model could be considered (Jin et al., 2005). The model has one more parameter than generalized proper MCAR (Gelfand and Vounatsou, 2003) and the spatial structure for given response category depends on other response categories.

Table 1 showed that EPG values presented strong zero excess for *F. hepatica* and *C. daubneyi*. A subsequent paper will address statistical modelling of zero-inflated counts (Lambert, 1999).

In conclusion, multivariate disease-mapping can be helpful for testing hypotheses related to ecological patterns of disease risk. For *F. hepatica* and *C. daubneyi* the results were consistent with the theoretical expectation; for *H. contortus*, the findings suggested coastal environmental factors in the transmission of the infection. *D. dendriticum* showed a structure similar to *F. hepatica*, while the baseline prevalence was very different.

References

Asanji MF, Williams MO, 1984. The effect of sex on seasonal variation in single and double infection of cattle in Sierra Leone by *Dicrocoelium hospes* and *Fasciola gigantica*. Vet Parasitol 15, 247-255.

Besag J, York JC, Mollié A, 1991. Bayesian image restoration, with two applications in spatial statistics (with discussion). Ann I Stat Math 43, 1-59.

Biggeri A, Catelan D, Rinaldi L, Dreassi E, Lagazio C, Cringoli G, 2006. Statistical methods for geographical surveillance in veterinary epidemiology. Parassitologia 48, 73-76.

Cringoli G, 2006. FLOTAC, a novel apparatus for a multivalent faecal egg count technique. Parassitologia 48, 381-384.

Cringoli G, Rinaldi L, Musella V, Veneziano V, Maurelli MP, Di Pietro F, Frisiello M, Di Pietro S, 2007a. Geo-referencing livestock farms as tool for studying cystic echinococcosis epidemiology in cattle and water buffaloes from southern Italy. Geospatial Health 2, 105-111.

Cringoli G, Rinaldi L, Veneziano V, Capelli G, Scala A, 2004a. The influence of flotation solution, sample dilution and the choice of McMaster slide area (volume) on the reliability of the McMaster technique in estimating the faecal egg counts of gastrointestinal strongyles and *Dicrocoelium dendriticum* in sheep. Vet Parasitol 13, 121-131.

Cringoli G, Taddei R, Rinaldi L, Veneziano V, Musella V, Cascone C, Sibilio G, Malone JB, 2004b. Use of remote sensing and geographical information systems to identify environmental features that influence the distribution of paramphistomosis in sheep from the southern Italian Apennines. Vet Parasitol 10, 15-26.

Cringoli G, Veneziano V, Rinaldi L, Sauve C, Rubino R, Fedele V, Cabaret J, 2007a. Resistance of trichostrongyles to benzimidazoles in Italy: a first report in a goat farm with multiple and repeated introductions. Parasitol Res 101, 577-581.

Gelfand AE, Vounatsou P, 2003. Proper multivariate conditional autoregressive models for spatial data analysis. Biostatistics 4, 11-25.

Gosoniu L, Vonatsou P, Sogoba N, Smith T, 2007. Bayesian modelling of geostatistical malaria risk data. Geospatial Health 1, 127-139.

Knorr-Held L, Best NG, 2001. A shared component model for detecting joint and selecting clustering of two disease. J

R Stat Soc Ser B 164, 73-85.

Jin X, Carlin B, Banerjee S, 2005. Generalized hierarchical multivariate CAR models for areal data. Biometrics 61, 950-961.

Lambert D, 1999. Zero-inflated Poisson regression, with an application to defects in manufacturings. Technometrics 34, 1-14.

Mage C, Bourgne H, Toullieu JM, Rondelaud D, Dreyfuss G, 2002. *Fasciola hepatica* and *Paramphistomum daubneyi*: changes in prevalences of natural infections in cattle and in *Lymnaea truncatula* from central France over the past 12 years. Vet Res 33, 439-447.

Maurelli MP, Rinaldi L, Capuano F, Perugini AG, Veneziano V, Cringoli G, 2007. Characterization of the 28S and the second internal transcribed spacer of ribosomal DNA of *Dicrocoelium dendriticum* and *Dicrocoelium hospes*.

Parasitol Res 101, 1251-1255.

Raso G, Vounatsou P, Singer BH, N'Goran EK, Tanner M, Utzinger J, 2006. An integrated approach for risk profiling and spatial prediction of *Schistosoma mansoni*-hookworm coinfection. Proc Natl Acad Sci USA 103, 6934-6939.

Rinaldi L, Perugini AG, Capuano F, Fenizia D, Musella V, Veneziano V, Cringoli G, 2005. Characterization of the second internal transcribed spacer of ribosomal DNA of *Calicophoron daubneyi* from various hosts and locations in southern Italy. Vet Parasitol 131, 247-253.

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, 583-639.

van Wyk JA, Cabaret J, Michael LM, 2004. Morphological identification of nematode larvae of small ruminants and cattle simplified. Vet Parasitol 6, 277-306.