Bayesian Latent Class Models in Veterinary and Human Epidemiology

Luzia Gonçalves, Ana Subtil, Nuno Brites, M. Rosário Oliveira, Ana Margarida Alho, José Meireles, Luís M. Madeira de Carvalho and Silvana Belo

Abstract Latent class models have been widely used to evaluate the performance of veterinary and human diagnostic tests, in the absence of a gold standard. In this work, we explore Bayesian latent class models, with and without restrictions, in different populations. We previously reported the Bayesian latent class analysis of the malaria dataset (n =3317) and now we apply a similar approach to a smaller dataset in the context of canine dirofilariasis (n =308). Although the last study presents a small sample size, it was important to obtain the first estimates of the prevalence (and performance measures of diagnostic techniques) of canine dirofilariasis in three districts of Portugal, taking into account the relationship with human dirofilariasis.

Key words: Latent class model, Bayesian approach, conditional dependence, diagnostic test, dirofilariasis.

Luzia Gonçalves
CEAUL and Unidade de Saúde Pública Internacional e Bioestatística, Instituto de Higiene e Medicina Tropical, Rua da Junqueira, 100, 1349-008 Lisboa, Portugal, e-mail: luziag@ihmt.unl.pt

Ana Subtil
CEMAT and Faculdade de Ciências da Universidade de Lisboa, Campo Grande Edifício C5, 1749-016 Lisboa, Portugal.

Nuno Brites
Unidade de Saúde Pública Internacional e Bioestatística, Instituto de Higiene e Medicina Tropical, Rua da Junqueira, 100, 1349-008 Lisboa, Portugal

M. Rosário de Oliveira
CEMAT and Departamento de Matemática, Instituto Superior Técnico, Universidade Técnica de Lisboa, Avenida Rovisco Pais, 1049-001 Lisboa, Portugal

Ana Margarida Alho, José Meireles and Luís M. Madeira de Carvalho
Laboratório de Parasitologia - CIISA - Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Avenida da Universidade Técnica, 1300-477 Lisboa, Portugal

Silvana Belo
UPMM and Unidade de Parasitologia Médica, Instituto de Higiene e Medicina Tropical, Rua da Junqueira, 100, 1349-008 Lisboa, Portugal

1 Introduction

An active field of biostatistical research has been Bayesian Latent Class Models (BLCM) which are used to estimate the prevalence, sensitivities and specificities in the absence of a gold standard. In this context, BLCM are widely used by the veterinary community [1, 14], where subpopulations (e.g. herds) appear naturally or are artificially created (e.g. [24]). The most popular model is the one that assumes difference in terms of prevalence and constant performance measures across populations. In practice, sometimes, this last aspect may be unrealistic and other models must be considered. Compared with a frequentist perspective, Bayesian approaches may be helpful when dealing with non-identifiable models, using informative priors, imposing constraints or admitting more than one population to avoid this problem. In case of one population, Menten el al. [10] present several BLCM in the study of the diagnosis of visceral leishmaniasis and Limmathurotsakul et al. [7] used similar models for melioidosis. In malaria context, Speybroeck et al. [20] present a contribution of a Bayesian approach to estimate the prevalence of malaria, applying ELISA, PCR and microscopy to datasets from Peru, Vietnam, and Cambodia separately. Ochola et al. [15] use a Bayesian formulation of the latent class model of Hui and Walter (two populations) to estimate the diagnostic accuracy of the malaria diagnostic techniques, based on a systematic review. Depending on the sampling scheme or the objectives of the analysis, the modelling versatility of the BLCM can be enhanced by incorporating constraints to explore differences and similarities between subpopulations in terms of prevalence, sensitivities and specificities. Recently, Gonçalves et al. [5] address this type of models and provide estimates to the malaria infection prevalence and performance measures in four subpopulations simultaneously based on a post-stratified analysis using two binary categorical variables: age groups (less than 5 years, greater than or equal to 5 years) and fever status (febrile, afebrile). In this work, we explore a dataset related with dirofilariasis diagnostic tests, in three districts of Portugal, when a stratified sampling was used and, therefore, we have natural populations (districts) under study. As the best of our knowledge, this type of models were not previously used in dirofilariasis diagnostic tests.

In this work, firstly, we give a brief description of the dirofilariasis dataset (Section 2) after we discuss the validation of the Hypothesis of Conditional Independence (HCI) in order to decide if the simplest and most parsimonious two latent class model describes the data adequately or if more complex models with dependencies between tests are needed. In the Section 4 we define the models. Finally, we present the results and final remarks.

2 Applications

Canine dirofilariosis is caused by *D. immitis* transmitted by mosquitoes. Diagnosis of canine dirofilariasis is performed using different methods which present difficulties in the interpretation of the results [2]. In Europe, the Mediterranean countries

present the highest prevalences [11]. In Portugal, in order to improve the epidemiological knowledge of the canine and human dirofilariosis, a research project has been implemented since 2011 in three districts of the centre-south areas - Setúbal, Santarém, and Coimbra. In the next sections, the districts (and sample sizes) are denoted by 1. Setúbal ($n_1 = 40$), 2. Santarém ($n_2 = 169$) , and 3. Coimbra ($n_3 = 99$). In this work we analise a sample of 308 kennel dogs from a preliminary survey on canine population. Several diagnostic tests are under evaluation and, in this study, we explore three techniques denoted by 1. a commercial antigen kit (WITNESS *Dirofilaria*), 2. the Modified Knott's technique and 3. Blood smear.

3 Hypothesis of Conditional Independence

The HCI in some medical problems may not be a realistic assumption, for example, when the two tests are based on a similar biological phenomenon (e.g. [10, 17]). The diagnostic of local dependence has been discussed by several authors [3, 6, 16, 17] and different methods have been proposed to validate this hypothesis. Among others, Hagenaars [6] suggests the analysis of the standardized residuals for each pair of manifest variables. Garrett and Zeger [3] developed the log odds ratio check (LORC) plot, to compare the log odds ratio for the observed and predicted two-way cross classification tables for each pair of manifest variables. Qu et al. [16] propose the correlation residual plot, which is obtained by plotting residuals of pairwise correlation coefficients, defined as the difference between the observed and expected correlations. Sepúlveda et al. [17] propose the use of Biplot representations based on generalized linear models to identify conditional dependence between pairs of manifest variables within each latent class. Subtil el al. [21] developed a simulation study, considering local dependence between pairs of manifest variables, and the application of different tools revealed some problems in the detection of the violation of the HCI.

4 Bayesian Latent Class Models

The LCM approach admits a binary latent variable, Y, whose categories are called latent classes and indicate the disease status: Y takes the value 1 if the disease or infection is present and 0 otherwise. The outcomes of p diagnostic tests in the jth subpopulation/district are expressed using manifest binary variables[1], X_{ij}, assuming the value 1 if the i-th diagnostic test is positive and 0 otherwise, $i = 1,\ldots,p$ ($p = 3$ in our example), $j = 1,...,J$, where J denotes the number of subpopulations ($J = 3$). The model parameters include the prevalence in the three districts, η_j, $j = 1,2,3$,

[1] To avoid burdening the notation, the index corresponding to each individual is omitted.

the sensitivities and specificities of the three tests in the three districts, denoted by Se_{ij} and Sp_{ij}, respectively, $i = 1,2,3$; $j = 1,2,3$.

In our work, the jth subpopulation counts ($\mathbf{O}_j$) of the different patterns of test results (in a total of 2^3 possible patterns) follow a multinomial distribution:

$\mathbf{O}_j|Se_{ij}, Sp_{ij}, \eta_j \sim$ Multinomial $(\mathbf{Pr}_j, n_j)$,

where n_j is the sample size of jth subpopulation, $j = 1,2,3$, $i = 1,2,3$, and $\mathbf{Pr}_j$ is a vector of probabilities of observing the individual pattern $\mathbf{x}_j = (x_{1j}, x_{2j}, x_{3j})^t$ of test results in population j.

Under the HCI, a generic element of the vector $\mathbf{Pr}_j$ is given by

$$P(\mathbf{X}_j = \mathbf{x}_j) = \eta_j \prod_{i=3}^{p} \{Se_{ij}^{x_{ij}}(1-Se_{ij})^{1-x_{ij}}\} + (1-\eta_j) \prod_{i=3}^{p} \{Sp_{ij}^{1-x_{ij}}(1-Sp_{ij})^{x_{ij}}\}.$$

Additionally, it is possible to model conditional dependence including covariance between pairs of tests, in a similar way to the approach of [10] for a unique population. For example, if we assume that tests X_2 and X_3 are correlated in the infected class, and in a similar way in all populations, the probability of an outcome pattern $\mathbf{x_j}$ is:

$$P(\mathbf{X}_j = \mathbf{x}_j) = \eta_j \{Se_{1j}^{x_{1j}}(1-Se_{1j})^{1-x_{1j}}\}\{Se_{2j}^{x_{2j}} Se_{3j}^{x_{3j}}(1-Se_{2j})^{1-x_{2j}}(1-Se_{3j})^{1-x_{3j}} + \\ + (-1)^{(x_{2j}-x_{3j})} cov_{23|Y=1}\} + (1-\eta_j) \prod_{i=3}^{p} \{Sp_{ij}^{1-x_{ij}}(1-Sp_{ij})^{x_{ij}}\},$$

where $cov_{23|Y=1} = cov(X_2, X_3|Y = 1)$.

To analyze the three districts simultaneously, a product multinomial distribution is considered simply using the product of three multinomial distributions since the subpopulations are independent. The general model (with parameters varying across districts), denoted by M1, may be simplified to obtain other simpler models, using constraints. The simplest model (denoted by M2 in next section) with constraints considers a different prevalence for each district and equal sensitivities and specificities of each test across districts, i.e., $Se_{ij} = Se_i, \forall j$ and $Sp_{ij} = Sp_i, \forall j$. This model is commonly used to evaluate diagnostic tests in two or more populations (see [22, 15, 1]). The general model (no constraints are imposed on prevalence, sensitivities and specificities across subpopulations) has 21 parameters and the simplest model has only 9 parameters to be estimated, using a Bayesian approach. Introducing different constraints into M1, several other BLCM can be fitted via MCMC techniques, using Gibbs sampling.

In terms of informative priors, in the malaria dataset, we used Beta distributions with α and β parameters – Beta(α, β) – taking advantage of its flexibility, choosing left-skewed distributions to suggest a trend to a poor performance of a test or a low prevalence and a rigth-skewed distribution when a good performance seems to be appropriate. The elicitation of an informative prior is a hard and subjective process

that needs a careful dialogue with experts and, in the canine dirofilariasis dataset analysis, the elicitation of informative priors is still in progress.

5 Results

As a first step to evaluate the validation of HCI, using the tools mentioned before (Section 2), in each district separately and considering a pool with the three districts, there is no evidence of violation of the HCI. In the malaria dataset ([5]), the decision was the same but the sample size is much larger. To evaluate if an increase of the sample size shows evidences of violation, we explore artificial situations where the sample sizes are $2n, 4n$, and $10n$ (or adding some small constant to prevent situations where the frequency is zero), and the empirical distributions are the same or similar. The biplot representation suggests a possible dependence between the Modified Knott's and blood smear tests, when the sample size is $4n$ and $10n$ (data not shown). This fact might suggest a possible model with dependence between those two tests.

We explored the general model M1 and the simplest model M2 using WinBugs (see tables 1 and 2). Inferences were based on 20,000 iterations, after discarding an initial burn-in of 5,000 iterations with convergence assessed by running multiple chains from various starting values [4]. All parameters were estimated with the highest probability density (HPD) intervals for parameters of interest, using the package BOA 1.1 7-2 [19]. Convergence was monitored by the standard diagnostic procedures based on a visual assessment of the long chains for each parameter and using the Gelman-Rubin and the Raftery-Lewis measures, also included in BOA. For both M1 and M2 there was no evidence of failure in convergence, since, in the case of the Gelman-Rubin diagnostic, $R < 1.2$ for all parameters and, for the Raftery-Lewis measures, all the dependence factors $DF < 5$ [19]. Among other, the Deviance Information Criterion (DIC) and a version of Bayesian p-value based on Pearson statistics, were also calculated as described by Nérette et al.[13]. In spite of the subjectivity, this version of Bayesian p-value suggests the lack of fit when p-values near 0 or 1 and an adequate model fit if p-value close to 0.5 [10, 12]. The values corresponding to models M1 and M2 are indicated in the caption of tables 1 and 2.

Based on this preliminary study, according to M1, the prevalence of *D. immitis* infection in dog increases from North to South districts, being higher in the district of Setúbal (18.4% – [5.8-32.1]), followed by Santarém (13.5% – [8.4-18.7]) and Coimbra (9.0% – [3.7-14.6]). As expected, some of the 95% HPD intervals are very wide, particularly, in the district of Setúbal (sample size only with 40 dogs). In terms of diagnostic tests, the commercial antigen kit (WITNESS *Dirofilaria*) presents lower values for the sensitivity and the specificity. In terms of specificities, all tests give posterior means (an medians) above 90%. The inter-district variations may be justified by ecological conditions favoring high densities of Dirofilaria vectors and that may explain the higher parasitic loads (data not shown) seen in dogs.

Table 1 Bayesian estimates of prevalence, sensitivities and specificities, given by posterior means, medians and 95% HPD intervals, with non-informative priors, by district, using model M1 (DIC=72.194; p-value: 0.429).

Parameters	mean	median	95% HPD
η_1	0.184	0.177	0.058 - 0.321
η_2	0.135	0.133	0.084 - 0.187
η_3	0.090	0.088	0.037 - 0.146
Se_{11}	0.579	0.583	0.252 - 0.910
Se_{12}	0.625	0.629	0.433 - 0.808
Se_{13}	0.702	0.716	0.436 - 0.947
Se_{21}	0.838	0.874	0.563 - 1.000
Se_{22}	0.957	0.970	0.873 - 1.000
Se_{23}	0.890	0.916	0.692 - 1.000
Se_{31}	0.687	0.699	0.371 - 1.000
Se_{32}	0.956	0.969	0.872 - 1.000
Se_{33}	0.891	0.918	0.692 - 1.000
Sp_{11}	0.915	0.923	0.818 - 0.999
Sp_{12}	0.973	0.976	0.947 - 0.995
Sp_{13}	0.947	0.951	0.901 - 0.987
Sp_{21}	0.923	0.932	0.824 - 1.000
Sp_{22}	0.993	0.995	0.980 - 1.000
Sp_{23}	0.989	0.992	0.968 - 1.000
Sp_{31}	0.967	0.977	0.903 - 1.000
Sp_{32}	0.993	0.995	0.980 - 1.000
Sp_{33}	0.989	0.992	0.968 - 1.000

Table 2 Bayesian estimates of prevalence, sensitivities and specificities, given by posterior means, medians and 95% HPD intervals, with non-informative priors, by district, using model M2 (DIC= 63.443; p-value: 0.3132).

Parameters	mean	median	95% HPD
η_1	0.189	0.184	0.071 - 0.314
η_2	0.135	0.133	0.085 - 0.187
η_3	0.090	0.087	0.039 - 0.147
Se_1	0.642	0.644	0.491 - 0.792
Se_2	0.972	0.981	0.919 - 1.000
Se_3	0.924	0.932	0.834 - 0.999
Sp_1	0.964	0.965	0.941 - 0.984
Sp_2	0.992	0.994	0.980 - 1.000
Sp_3	0.996	0.997	0.989 - 1.000

6 Final Remarks

Even if forthcoming field studies are needed to better understand this findings in the three districts, this preliminary study showed the existence of canine dirofilariosis. Some studies show similar prevalence in the canine and human populations [11]. Theses findings may reveal the possibility of occurrence of human pulmonary dirofilariosis in the inhabitants of those districts and physicians should be alerted to this health problem. As new studies are planned in a near future, additional statistical modelling will be explored. Although the large sample sizes are not always possible in practice, statistical analysis continues to be useful for decision-making.

Acknowledgements This work was partially sponsored by national funds through the Fundação Nacional para a Ciência e Tecnologia (FCT), Portugal, under the projects PTDC/SAU-ESA/81240/2006, PTDC/SAL-SAP/113523/2009 and PEst-OE/MAT/UI0006/2011. Ana Subtil has an FCT PhD grant SFRH/BD/69793/2010. The authors are grateful to all veterinarians who kindly collaborated in the dirofilariosis study.

References

1. Bronsvoort, B., Toft, N., Bergmann, I., Srensen, K., Anderson, J., et al.: Evaluation of three 3ABC ELISAs for foot-and-mouth disease non-structural antibodies using latent class analysis. BMC Veterinary Research 2(2006)
2. Furtado, A.P., Do Carmo E.S., Giese, E.G., Vallinoto, A.C., Lanfredi, R.M., Santos, J.N.: Detection of dog filariasis in Marajo Island, Brazil by classical and molecular methods. Parasitol. Res.**105**(6):1509–1515 (2009)
3. Garrett, E.S., Zeger, S.L.: Latent class model diagnosis. Biometrics **56**: 1055–1067 (2000)
4. Gelman, A., Rubin, D.B.: Inference from iterative simulation using multiple sequences. Statistical Science **7**: 457–72 (1992)
5. Gonçalves, L., Subtil, A., Oliveira, M.R., Rosário, V., Lee, P., Shaio, M.-F.: Bayesian latent class models in malaria diagnosis. Under revision (2012)
6. Hagenaars, J.: Latent structure models with direct effects between indicators. Sociological Methods and Research **164**: 379–405 (1988)
7. Limmathurotsakul, D., Jamsen, K., Arayawichanont, A., Simpson, J.A., White, L.J., et al. Defining the true sensitivity of culture for the diagnosis of melioidosis using Bayesian Latent Class Models. PLoS ONE **5**: e12485 (2010)
8. Lunn, D., Thomas, A., Best, N., Spiegelhalter, D.: WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing **10**: 325–337 (2000)
9. Martinez, E., Louzada-Neto, F., Derchain, S., Achcar, J., Gontijo, R., et al.: Bayesian estimation of performance measures of cervical cancer screening tests in the presence of covariates and absence of a gold standard. Cancer Informatics **6**: 33–46 (2008)
10. Menten, J., Boelaert, M., Lesaffre, E.: Bayesian latent class models with conditionally dependent diagnostic tests: A case study. Statistics in Medicine **27**: 4469–4488 (2008)
11. Montoya-Alonso JA, Carretn E, Corbera JA, Juste MC, Mellado I, Morchn R, Simn F. Current prevalence of Dirofilaria immitis in dogs, cats and humans from the island of Gran Canaria, Spain. Vet. Parasitol. 176(4):291–4 (2011)
12. Neelon, B., O'Malley, A.J., Normand, S.L.T.: A Bayesian two-part latent class model for longitudinal medical expenditure data: Assessing the impact of mental health and substance abuse parity. Biometrics **67**: 280–289 (2011)

13. Nérette, P., Stryhn, H., Dohoo, I., Hammell L.: Using pseudogold standards and latent-class analysis in combination to evaluate the accuracy of three diagnostic tests. Preventive Veterinary Medicine **85**: 207-225 (2008)
14. Nielsen, L., Ersboll, A.: Age-stratified validation of an indirect Salmonella Dublin serum enzyme-linked immunosorbent assay for individual diagnosis in cattle. Journal of Veterinary Diagnostic Investigation **16**: 212–218 (2004)
15. Ochola, L., Vounatsou, P., Smith, T., Mabaso, M., Newton, C.: The reliability of diagnostic techniques in the diagnosis and management of malaria in the absence of a gold standard. Lancet Infectious Disease **6**: 582–588 (2006)
16. Qu, Y., Tan, M., Kutner, M.: Random effects models in latent class analysis for evaluating accuracy of diagnostic test. Biometrics **52**: 797–810 (1996)
17. Sepúlveda, R., Vicente-Villardón, J.L., Galindo, M.P.: The Biplot as a diagnostic tool of local dependence in latent class models. A medical application. Statistics in Medicine **27**: 1855–1869 (2008)
18. Lee, P., Liu, C., do Rosário, V., de Sousa, B., Rampao, H., et al.: Potential threat of malaria epidemics in a low transmission area, as exemplified by São Tomé and Príncipe. Malaria Journal **9**:264 (2010)
19. Smith, B.J.: BOA: An R package for MCMC output convergence assessment and posterior inference. Journal of Statistical Software **21**: 1–37 (2007)
20. Speybroeck, N., Praet, N., Claes, F., Hong, N.V., Torres, K., et al.: True *versus* apparent malaria infection prevalence: The contribution of a Bayesian approach. PLoS ONE **6**: e16705 (2011)
21. Subtil, A., Oliveira, M.R. and Gonçalves, L.:. Conditional dependence diagnostic in the latent class model: A simulation study. Statistics and Probability Letters **82**, 1407-1412 (2012)
22. Toft, N., Jørgensen, E., Højsgaard, S.: Diagnosing diagnostic tests: evaluating the assumptions underlying the estimation of sensitivity and specificity in the absence of a gold standard. Preventive Veterinary Medicine **68**: 19-33 (2005)
23. Toft, N., Innocent, G., Gettinby, G., Reid, S.: Assessing the convergence of Markov Chain Monte Carlo methods: An example from evaluation of diagnostic tests in absence of a gold standard. Preventive Veterinary Medicine **79**: 244–256 (2007)
24. Weber, M., Verhoeff, J., van Schaik, G., van Maanen, C.: Evaluation of Ziehl-Neelsen stained faecal smear and ELISA as tools for surveillance of clinical paratuberculosis in cattle in the Netherlands. Preventive Veterinary Medicine **92**: 256-266 (2009).