ANALYSIS OF APPENDECTOMY IN BELGIUM USING DISEASE MAPPING TECHNIQUES

Setia Pramana
Karolinska Institutet

Proc. ICCS-13, Bogor, Indonesia
December 18-21, 2014, Vol. 27, pp. 417-430

ANALYSIS OF APPENDECTOMY IN BELGIUM USING DISEASE MAPPING TECHNIQUES

Mieke Nurmalasari[1] and **Setia Pramana**[2]
[1] Sekolah Tinggi Ilmu Ekonomi Indonesia, Jl Kayu Jati, 11A, Jakarta, Indonesia. Email: mieke_hafidz@yahoo.co.id
[2] Sekolah Tinggi Ilmu Statistik, Jl Otista 64C, Jakarta, Indonesia Email: setia.pramana@stis.ac.id

ABSTRACT

An appendectomy is the surgical removal of the vermiform appendix normally performed as an emergency, when the patient is suffering from acute appendicitis. The analysis of appendectomy related to the geographical distribution of disease hospital-admissions in Belgium was still under studied. This study is aimed to identify geographical differences in medical practices, and to investigate spatial and temporal distribution on appendectomy and incidental appendectomy incidence rate in Belgium for period 2001 to 2006.

Two different methods were applied to identify possible high incidence regions, used maps of the non-smoothed SIRs and Bayesian methods to smooth SIRs. Using Bayesian method, the best model (based on the smallest value of DIC) for appendectomy and incidental appendectomy cases are Poisson-Lognormal and Poisson-Gamma, respectively. The range of mean relative risks for appendectomy cases was smoothed between 0.81 and 1.32, and the range of mean relative risks for incidental appendectomy was smoothed between 0.32 and 2.67. Based on these models, we can conclude that the model of smoothed SIRs (mean relative risks) of appendectomy and incidental appendectomy cases among districts in Belgium for 2001-2006 periods are not related with the environment.

Based on smoothed SIR, three districts (Diskmuide, Tielt and Dinant) have higher relative risk than other districts in appendectomy cases. For incidental appendectomy, the highest relative risks are district Oudenaarde, Sint-Niklaas, Hoeiand and Bastenaken. The significant increased or decreased incidence of appendectomy and incidental appendectomy cases in districts are need to be investigated further. It might also be the expression of differences in medical practice.

Considering analysis of spatial temporal using Bernardinelli model, the risk of appendectomy cases significantly decreased in time where the risk was multiplied by approximately 0.9888 every year. In the other hand, the incidental appendectomy cases increased from 2001-2006 and the increase was significant for incidental appendectomy cases over time. The risk was multiplied by approximately 1.0392 every year.

KEYWORDS

Spatial Analysis, Appendectomy, disease mapping, Bernardinelli model.

1. INTRODUCTION

Since October 1990, all Belgian hospitals are subjected to compulsory registration with the health authorities of each admission through a standard form containing a define set of clinical data including ICD-coded diagnoses and procedures. These discharge abstracts are termed Minimal Clinical Data (MCD) and contain patient data (among which year of birth, gender, residence, and anonymous hospital and patient identifiers, stay data among which year and month of admission and discharge, length of stay, transfer to another hospital with specification of the type of hospital) and an unlimited number of diagnoses and procedures. This information (about 2,000,000 hospital-admissions per year) is transmitted to the authorities for compilation and processing. From MCD we can obtain the diagnose of two selected pathologies, appendectomy and incidental appendectomy. Nevertheless, the analysis of appendectomy related to the geographical distribution of disease hospital-admissions in Belgium was still under studied.

Both appendectomy and incidental appendectomy are a surgical removal of the vermiform appendix that differs on the purpose. An appendectomy is normally performed as an emergency procedure, when the patient is suffering from acute inflammation of the appendix known as appendicitis. In contrast, incidental appendectomy is performed incidental to other abdominal surgery, such as urological, gynecological, or gastrointestinal surgeries, intended to eliminate the risk of future appendicitis and to simply any future differential diagnoses of abdominal pain. For these cases, classical epidemiological measures such as Standardized Incidence Ratio (SIR) still can be calculated because the appendix of the patients only can be removed once.

The objectives of this study are (1) to carry out an analysis of the geographical distribution of disease hospital-admissions regarding appendectomy and incidental appendectomy, (2) to identify geographical differences in medical practices and (3) to investigate spatial and temporal distribution on appendectomy and incidental appendectomy incidence rate in Belgium for period 2001 to 2006.

2. MATERIALS AND METHODS

Data

The data used is the number of appendectomy and incidental appendectomy cases in Belgium during 2001 to 2006, and number of Belgium population from 2001 to 2007 (calculate the middle year). The population for appendectomy is all Belgian population, whereas for incidental appendectomy the population is patients older than 65 years old. District will be the unit of analysis related to appendectomy and incidental appendectomy cases. The data set comes from 43 districts. In this report, only some variables from Minimal Clinical Data (MCD) are used i.e., sex, age group, residence, patient identifiers and year.

Standardized Incidence Ratio (SIR)

Incidence rate, a number of new cases per population at risk in a given time period, is a measure of the frequency with which a disease occurs in a population over a period of time. However, in many case a region or a district with higher population may have more cases than the one with less population. In this case the Standardized Incidence Ratio (SIR) is used instead of incidence rate.

Standardized Incidence Ratio (SIR) for each district can be defined as:

$$SIR = \frac{y_i}{E_i} = \frac{\sum_g y_{gi}}{\sum_g E_{gi}} = \frac{\sum_g y_{gi}}{\sum_g \frac{y_g^S}{n_g^S} n_{gi}} \qquad (1)$$

i = 1, 2, …, m where ygi is the number of cases in age group g for study population i, yi is the total number of cases observed on the study population, Egi expected number of cases in age group g for study population i, Ei is the overall expected number of cases for the study population, ngi is number of people at risk in age group g for study population i, and , denote the same quantities for the standard population (Waller and Gotway, 2004).

Poisson Model

The Standardized Incidence Ratio (SIR) for region i is obtained from the ratio of the observed and expected number of cases (Yi=Ei) in that region (i = 1, 2, …, m). Indeed, independently in each region i, the number of cases are supposed to follow a Poisson distribution.

The Poisson model:

$$y_i \sim Poisson\ (e_i\ ,\ \theta_i) \qquad (2)$$

where θ_i is the relative risk of disease in region *i*. The 95% confidence interval (CI) for SIR can be calculated as

$$[SIR_i * \exp(-1.96/\sqrt{Y_i}\ );SIRi * \exp(1.96/\sqrt{Y_i}\ )] \qquad (3)$$

This is equal to the CI [SIR_i/errfac; SIR_i * errfac] with *errfac* is an error factor defined as

$$errfac = \exp\left(z_{1-\alpha/2}\sqrt{\frac{1}{y}} \right) \qquad (4)$$

(Clayton and Hills, 1995).

The conventional approach of mapping standardized disease rates based on Poisson inference gives a good illustration of the geographical distribution of the underlying rates when the disease is not rare. However, for rare disease or small areas, these maps often produce a mix of colors that are difficult to interpret. Moreover, the numbers of disease cases observed in each small area are often more variable than that implied by the standard Poisson model. Bayesian models have been developed in disease mapping in order to take into account the extra Poisson variation. One way is to shrink the most unreliable standardized rates towards the overall mean rate.

Bayesian Methods

Modern approaches to relative risk θi estimation rely on smoothing methods. These methods often involve additional assumptions or model components. Here, a bayesian modeling approach was used. The bayesian method was assumed that the relative risk estimator has a distribution. In the Bayesian terms, this is called a prior distribution. In the Poisson count, the most common prior distribution is to assume that θi has Gamma distribution.

Here, three models were applied, i.e., Poisson-Gamma, Poisson-Lognormal and Conditional Autoregressive (CAR) model, to investigate spatial distribution of appendectomy and incidental appendectomy for each district in the study area for the 6 years period from 2001-2006. Model fitting was carried out using MCMC simulation methods implemented in the WinBUGS software. Two separate chains which starting from different initial values were run for each model. Convergence was checked by visual examination of ”time series” style plots of the samples for each chain, and by computing the Gelman-Rubin convergence statistic (Gelman and Rubin, 1992).

Poisson-Gamma model

When the disease is rare, the numbers of diseases in each area are assumed to be mutually independent and follow Poisson distributions

$$y_i = \text{Poisson}(e_i, \theta_i) \tag{5}$$

for $\forall_i, \theta_i \sim Gamma(a,b)$ with mean $m_{\theta_i} = a/b$ and variance $v_{\theta_i} = a/b^2$. As the result, the relative risk has the following distribution (posterior)

$$\theta_i \sim Gamma(a + y_i, b + e_i) \tag{6}$$

Poisson-Lognormal Model

The log-normal model for the relative risk is defined as

$$y_i = Poisson(e_i, \theta_i) \tag{7}$$

$$\log(\theta_i) = \alpha + v_i \tag{8}$$

where $v_i \sim N(0, \sigma_v^2)$ is the heterogeneity random effect, capturing extra-Poisson variability in the log-relative risks. The Lognormal model for the relative risk is more flexible (Lawson, et al., 2008). A major drawback with gamma prior is this method does not take into account the geographical location of the region. The models do not cope the spatial correlation. It is possible to account for the spatial pattern in disease by using Conditional Autoregressive (CAR) model.

Conditional Auto Regressive model (CAR)

The conditional autoregressive (CAR) model proposed by Besag et al. (1991) is used. In this model for relative risks, area specific random effects are decomposed into a component that takes into account the effects that vary in a structured manner in space (clustering or correlated heterogeneity) and a component that models the effects that vary in an unstructured way between areas (uncorrelated heterogeneity).

The model of CAR can be represented as:

$$y_i = Poisson(e_i, \theta_i), \tag{9}$$

$$\log(\theta_i) = \alpha + u_i + v_i, \tag{10}$$

where α is an overall level of the relative risk, ui is the correlated heterogeneity, and $v_i \sim N(0, \sigma_v^2)$ is the uncorrelated heterogeneity. The spatial correlation structure is used then the estimation of the risk in any area depends on neighboring areas $[u_i \mid u_j, i \neq j, \tau_u^2] \sim N(\bar{u}_i, \tau_i^2)$. The u_i is smoothed towards the mean risk in the set of neighboring areas, with variance inversely proportional to the number of neighbors (Lawson, et al., 2008).

The three models above are then compared using overall goodness of fit measures, such as Deviance Information Criteria (DIC) that has been proposed by Spiegelhalter et al. (2002)

$$DIC = 2E_{\theta|x}(D) - D(E_{\theta|x}(\theta)) \tag{11}$$

with D(.) the deviance *(-2*log(likelihood))* of the model and x the observed data. The model with a smallest DIC is the best model to predict a replicate data set of the same structure as that currently observed.

Bernardinelli Model

The most common way to consider the analysis of disease maps that have a temporal dimension is to count number of cases of diseases within small areas that are available for a sequence of T time periods. In this section we are going to analyze the space-time distribution of appendectomy cases and incidental appendectomy cases in Belgium over a period of six years (2001-2006) using Bernardinelli model.

Bernardinelli et al. (1995) suggests a model in which both area-specific intercept and temporal trend are modeled as random effects. This formulation allows for spatiotemporal interactions where temporal trend in risk may be different for different spatial locations and may even have spatial structure. All temporal trends are assumed to be linear. The Bernardinelli model is defined as

$$y_{ik} \sim Poison(e_{ik}\theta_{ik}) \tag{12}$$

$$\log(\theta_i) = \alpha + u_i + v_i + \beta t_k + \delta_i t_k \tag{13}$$

where α is an intercept (overall rate), u_i and v_i are area random effects (as defined in the CAR model), βt_k is a linear trend term in time t_k, δ_i is an interaction random effect between space and time. Prior distribution must be assumed for the parameters in this model.

3. RESULTS

During 2001 to 2006, the total cases of appendectomy and incidental appendectomy are 83981 and 1907, respectively. The incidence of appendectomy and incidental appendectomy was the highest in patients of 9-15 years and 69-75 years old, respectively. This is the same for both male and female patients. In general, there is a decreasing incidence of incidental appendectomy with age. Incidental appendectomy and appendectomy diseases affect people of certain age disproportionally.

We observed that the number of cases and the population decreases from 2001 to 2005. They then increase slightly in 2006. In the other hand, the number of Incidental Appendectomy cases increases from 2001 to 2005 before it decreases slightly in 2006.

Spatial Analysis

Next, the disease mapping technique is used as a way of presenting the results and demonstrating the geographical variation of appendectomy and incidental appendectomy incidence in each district. When comparing the incidence of both cases between two areas, or when investigating the pattern of appendectomy and incidental appendectomy cases for the same areas, it is important to adjust for differences in the age and gender of those populations. In this study, this was accomplished by gender-stratified and age standardization.

Appendectomy Cases

In the appendectomy cases (Figure 1), we observe the geographical variation SIR of appendectomy cases in district in Belgium, 2001-2006. The significance of SIR also was calculated using the 95% confidence interval of SIR, which is calculated by equation (3). In Figure 1, the map shows that the three areas with the highest SIRs are shaded in dark green. These areas are district Diksmuide, Tielt and Dinant (SIR > 1.2). Other districts that have significant increased incidence as compared to the whole study region are Mechelen, Nijvem, Roeselare, Eeklo, Moeskroen, Verviers, Borgworm, Bastenaken and Namen. Otherwise, district Leuven, Aalst, Gent, Charleroi, Zinnik, Thuin, Luik, Hasselt, Maaseik, Aarlen, Marche-en-Famenne and Virton have decrease incidence significantly or fewer cases occurred than expected.

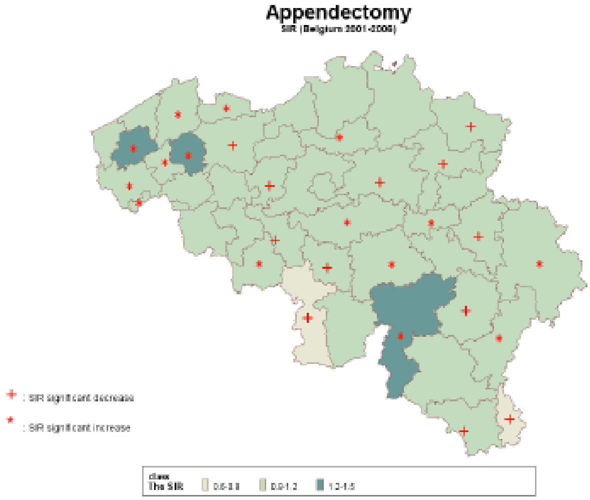

Figure 1: SIR map of the appendectomy cases 2001-2006

Incidental Appendectomy

Figure 2 shows that the value of SIR for incidental appendectomy cases have more variability of SIR as compared to the SIR in appendectomy cases. Districts with SIR > 2 are Oudenaarde, Sint-Niklaas, Hoei and Bastenaken. Some districts with significant increase on incidence are Antwerpen, Oostende, Veurnee, Bergen, Luik, Aarlen and Neufchatean. These districts indicate that more cases of incidental appendectomy occurred than expected based on the age specific incidence proportions from the standard population. On the other hand, district Turnhout, Leuven, Brugge, Roeselare, Eeklo, Charleroi, Zinnik, Hasselt, Maaseik, Tongeren, Namen and Philippeville indicate fewer cases occurred than expected.

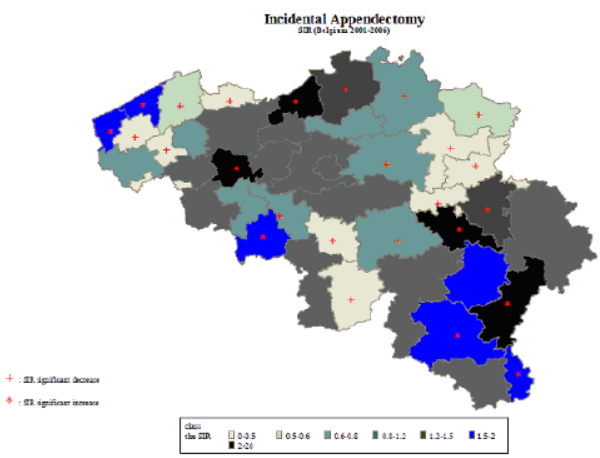

Figure 2: SIR map of Incidental Appendectomy cases 2001-2006

The result from ESDA and the non-smoothed SIRs (not shown) indicated that there is no much different for SIR of appendectomy and incidental appendectomy by gender. Hence, for the next analysis, we only consider analysis of appendectomy and incidental appendectomy that already corrected by age group and gender, not presented the SIR of these cases by gender separately.

Bayesian Methods

Appendectomy Cases

Three models were compared using Deviance Information Criterion (DIC) to determine which model gives the good estimates. They were Poisson-Gamma model (DIC=474.907), Poisson-Lognormal model (DIC=474.773) and Conditional Auto Regressive model (DIC=474.805). Poisson-LognormaL model is better than Poisson Gamma and CAR model, according to the DIC value.

The comparison map of non-smoothed SIR and smoothed SIR for the period 1996-2000 are presented in Figure 3. There is no much different between non-smoothed SIR and mean Relative Risk (smoothed SIR), but the value of SIR were smoothed by using Bayesian. The range of non-smoothed SIR is 0.78-1.36, then after smoothing method using Poisson-Lognormal model, the range of relative risk is 0.81-1.32.

District which has significant increased or decreased risk of this disease is based on the 95% Confidence Interval for SIR (CISIR) and 95% Credible Interval for mean relative risk (CI Bayes). It can be seen in Figure 3.

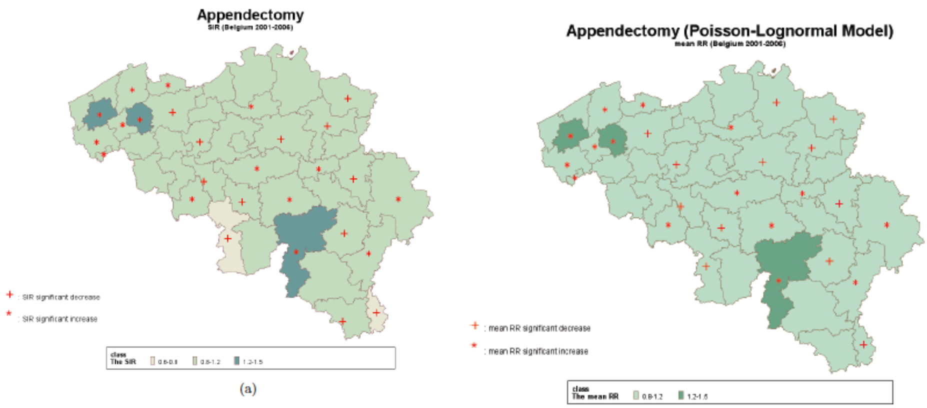

Figure 3: The comparison map of non-smoothed SIR and smoothed SIR for the period 2001-2006

Incidental Appendectomy Cases

The DIC for Poisson-Gamma model is 303.660, DIC Poisson-Lognormal model = 304.142, DIC CAR model =303.786. Based on this, Poisson-Gamma model is the best model for incidental appendectomy. Table 1 shows the posterior values for the parameters of the model after 50000 iterations.

Table 1
Posterior statistics for the parameters in the Poisson-Gamma model

	Mean	**SD**	**Credible Interval**
Alpha	3.318	0.8054	(0.01485; 5.174)
Beta	3.188	0.8305	(0.01526; 5.098)

There is few difference of map distribution (Figure 4) after using Poisson-Gamma model, the mean relative risk were smoothed. District Aarlen is not significance anymore. The range of non-smoothed SIR is between 0.18 and 3.18, and Poisson Gamma-smoothed SIR (mean relative risk) is between 0.32 and 2.67.

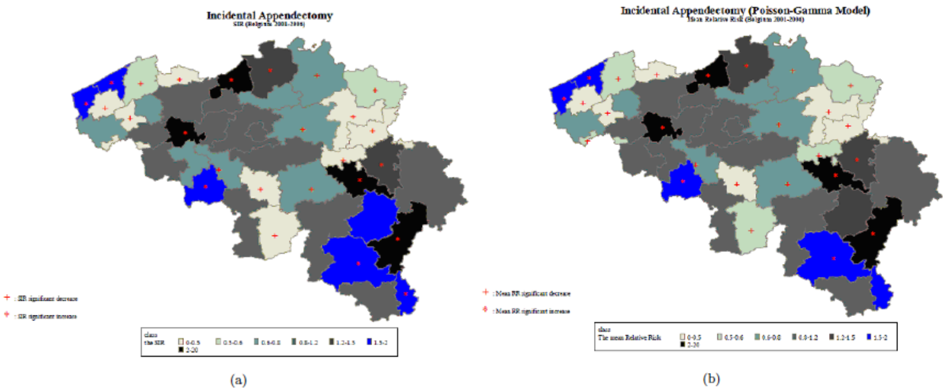

Figure 4 (a) Non-Smoothed SIRs and (b) Bayesian Mean Relative Risk (RR) of Incidental Appendectomy Cases in Belgium, 2001-2006; using Poisson-Gamma Model.

Bernardinelli Model

In this part, we considered the analysis of appendectomy and incidental appendectomy cases that have temporal dimension (from 2001-2006). Using Bernardinelli model, we obtained the estimation of the time effect for appendectomy cases as shown in Table 2.

Table 2
Mean Estimates of Time effect from Bernadinelli Model for Appendectomy

	Mean	Lower	Upper
Alpha	-0.5523	-4.378	0.0886
Beta	-0.0112	-0.0156	-0.00688
σ_u^2	0.0193	2.607E-4	0.0648
σ_v^2	1.882	9.397E-4	19.87
σ_{delta}^2	0.0016	6.766E-4	0.0033

Table 2 shows that the risk of appendectomy cases small significant decrease over time. This can be observed from the obtained trend term in time t_k which equal -0.0112. From this value, we can also obtain the ratio between two consecutive years which equal

0.9888 (*exp* [-0.0112]). Hence, the risk was multiplied by approximately 0.9888 every year. Maps of spatially smoothed relative risk of appendectomy cases at different time points are shown in Figures 5. A map of spatially smoothed time trends as shown in Figure 6 provides a visual impression of the small decreased in incidence occurring.

The map of temporal trend as derived from Bernardinelli model showed slightly decrease in risk of appendectomy cases (see Figure 6).

In contrast with the appendectomy cases, the risk for incidental appendectomy shows a small significant increase from 2001-2006. Based on Table 3, the ratio between two consecutive years was 1.039 (*exp* [0.0385]), thus the risk was multiplied by approximately 1.039 every year. A map of spatially smoothed relative risk of incidental appendectomy over time is presented in Figures 7.

Table 3
Mean Estimates of Time Effect from Bernadinelli Model for Incidental Appendectomy

	Mean	Lower	Upper
Alpha	-563.7	-1097.0	-26.75
Beta	0.0385	0.0052	0.0713
σ_u^2	33330	18.22	1.07E+5
σ_v^2	443200	685.6	1.405E+6
σ_{delta}^2	0.2504	0.1399	0.4244

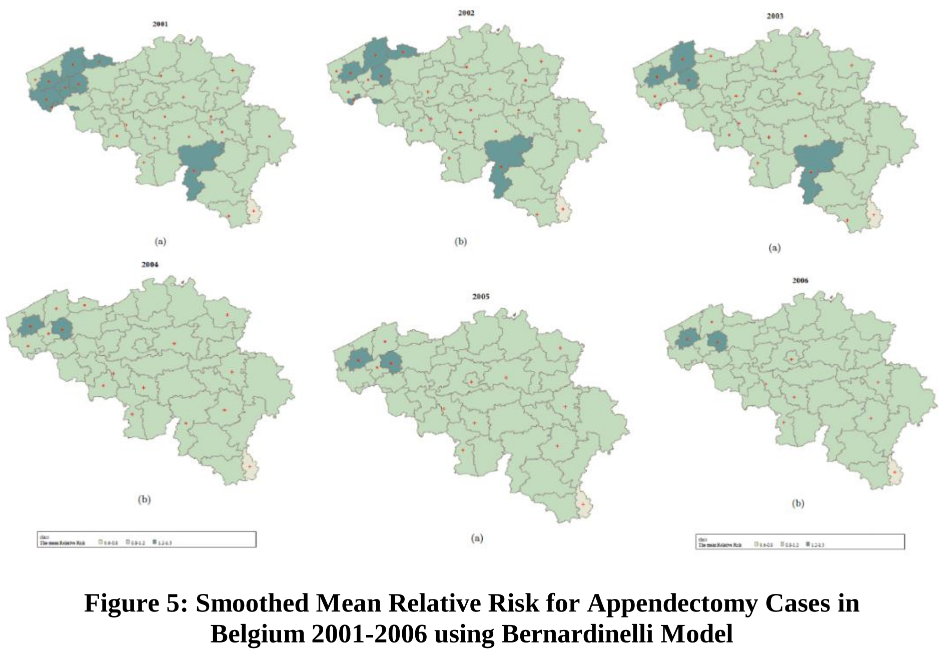

Figure 5: Smoothed Mean Relative Risk for Appendectomy Cases in Belgium 2001-2006 using Bernardinelli Model

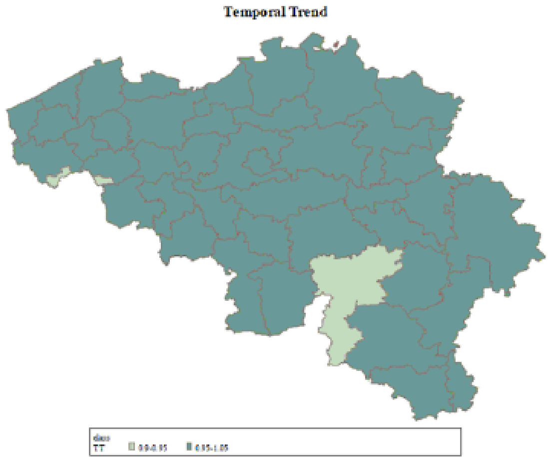

Figure 6: Temporal Trend for Appendectomy cases in Belgium 2001-2006 using Bernardinelli model.

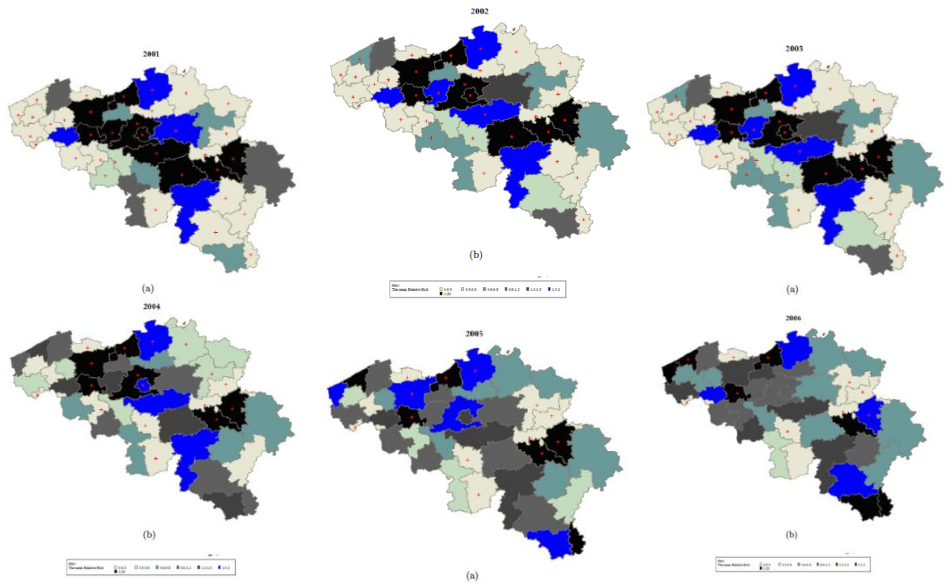

Figure 7: Smoothed Mean Relative Risk for Incidental Appendectomy Cases in Belgium, 2001-2006 using Bernardinelli Model

The map of temporal trend as derived from Bernardinelli model is shown in Figure 8.

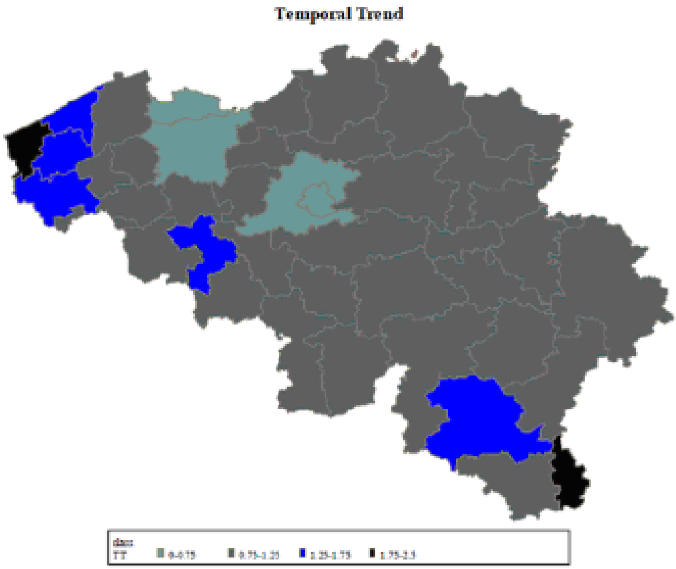

Figure 8: Temporal Trend for Incidental Appendectomy Cases in Belgium 2001-2006 using Bernardinelli Model

4. DISCUSSIONS AND CONCLUSION

The result shows that for the appendectomy cases, using non-smoothed SIRs, district Diksmuide, Tielt and Dinant have the higher SIR (SIR>1.2) than the other districts. This is due to the fact that these three districts have a rather small population and thus more prone to have more extreme values. These three districts have significant increased incidence of appendectomy disease as compare to the whole study region. For example, district Dinant has SIR=1.27, it is interpreted as 27% more cases observed than expected number.

The SIR for incidental appendectomy cases has more variability of SIR as compared to the SIR in appendectomy cases. This is because of the incidence of incidental appendectomy cases is rare. Districts with SIR> 2 are Oudenaarde, Sint-Niklaas, Hoei and Bastenaken. Other districts with significant increase incidence are Antwerpen, Oostende, Veurnee, Bergen, Luik, Aarlen and Neufchatean. A different case for district Turnhout, Leuven, Brugge, Roeselare, Eeklo, Charleroi, Zinnik, Hasselt, Maaseik, Tongeren, Namen and Philippeville. The increased or decreased incidences of incidental appendectomy cases in these districts need to be investigated further. It might also be the expression of differences in medical practice. Observing the appendectomy and incidental appendectomy cases by gender, the maps showed quite similar result. Furthermore, the difference in incidence between males and females does not seem too essential for both cases.

Using the non-smoothing SIRs, having risks when dealing with relatively small districts or for districts with relatively low numbers of cases or disease, the disease incidence rates tends to differ largely due to random error and may have misleadingly high or low values. To prevent the misleading result from non-smoothed SIRs the Bayesian smoothing with three different models was carried out. Result shows that the DIC differences between the 3 models for appendectomy and incidental appendectomy seem very small (< 0:2%), for example, in case of appendectomy: Poisson-Gamma model (DIC=474.907), Poisson-Lognormal model (DIC=474.773) and Conditional Auto Regressive model (DIC=474.805). We agreed to choose the smallest DIC, because it is estimated to be the model that would best predict a replicate data set of the same structure as that currently observed, hence the Poisson log-normal is the best model for appendectomy cases. Using this method, the value of mean relative risks were smoothed, the range of non-smoothed SIRs is 0.78 and 1.36, then after smoothing method is 0.81 and 1.32.

We repeated the analysis in incidental appendectomy. Based on the value of DIC, Poisson-Gamma model is the best model for incidental appendectomy. The value of mean relative risks were smoothed, districts Aarlen were not significant anymore compare with value of non-smoothed SIR before. The range of SIR is 0.18 - 3.18 and the range of mean RR is 0.32 - 2.67.

Finally, we can conclude that the model of smoothed SIRs (mean relative risks) of appendectomy and incidental appendectomy cases among districts in Belgium for 2001-2006 periods are not related with the environment. The best model for both cases are Poisson Lognormal and Poisson Gamma, respectively.

When we considered the analysis of appendectomy and incidental appendectomy cases which have temporal dimension using the Bernardinelli model, the risk of appendectomy cases significantly decreased in time where the risk was multiplied by approximately 0.9888 every year. In the other hand, the incidental appendectomy cases increased from 2001-2006 and the increase was significant for incidental appendectomy cases over time. The risk was multiplied by approximately 1.0392 every year.

REFERENCES

1. Bernardinelli, L., Clayton, D., Pascutto, C. Montomoli, C., Ghislandi, M. and Songini, M. (1995). Bayesian Analysis of Space-Time Variation in Disease Risk. *Statistics in Medicine*, 14, 2433-2443.
2. Besag, J., J. York, and A. Mollie (1991). Bayesian Image Restoration with Two Applications in Spatial Statistics. *Annals of the Institute of Statistical Mathematics*, 43, 1-59.
3. Clayton, D. and Hills, M. (1995). *Statistical Model in Epidemiology*. Oxford University Press, Oxford.
4. Clayton, D. and Bernardinelli L. (1992). *Bayesian Methods for Mapping Disease Risk. In Geographical and Environmental Epidemiology: methods for small-area studies*. (P. Elliot, J. Cuzick, D. English, and R. Stern ed.), 205-220. Oxford University Press.

5. Faes, C. and Abrahantes, J.C. (2009). Lecture Notes for Disease Mapping Course. Censtat - University of Hasselt. Belgium.
6. Gelman, A. and Rubin, D.B. (1992). Inference from iterative simulation using Multiple Sequences. *Statistical Science,* 7, 457-72.
7. Lawson, A., Browne, W.J, and Vidal Rodeiro, C. (2003). *Disease Mapping with WinBUGS and MLwiN* John Willey & Sons Inc. USA.
8. Spiegelhalter, D.J., Best N.G., Carlin B.P. and Van der Linde, A. (2002). Bayesian measures of model complexity and fit. J. *Roy. Stat. Soc. Ser. B-Stat. Met*. 64: 583-616, Part 4.
9. Waller L.A. and Gotway C.A. (2004). *Applied Spatial Statistics for Public Health Data*. John Willey & Sons, Inc. New Jersey.