UNIVERSITY OF NOVA GORICA
GRADUATE SCHOOL
STUDY PROGRAMME:
ENVIRONMENTAL SCIENCES

STUDY OF ETIOLOGICAL FACTORS OF CHRONIC RESPIRATORY DISEASES IN MUNICIPALITY OF KOPER

Master’s thesis

Saša ERLIH

Mentor: doc.dr. Ivan Eržen, dr.med., spec. epid.

Nova Gorica, 2009

ABSTRACT

Background

Chronic respiratory diseases (CRD) represent a rapidly growing health problem which is gaining on importance especially in more developed parts of the world, also in Slovenia. The respiratory diseases like asthma, chronic bronchitis and other allergic respiratory diseases that are affecting airflow are becoming more and more important already in early childhood. Therefore, the objective of this study was to investigate the geographical distribution and patterns of chronic respiratory disease in population of pupils born between 1991 and 1996 in Municipality Koper of south-western Slovenia in order to identify high risk areas and possible causes.

Approaches and methods

To assess the distribution of cases, the data from previous epidemiological study Prevalence of chronic respiratory diseases among children from variously polluted areas of the municipality of Koper (2003) were processed. Children born between 1991 and 1996 attending 1st to 4th year of primary school in the territory of the Municipality of Koper were included in this research. Spatial analyses were used with an intention of identifying and assessing geographical patterns of CRD in Koper. ESDA methods, Global Moran's I and Moran Local Indicators of Spatial Associations were used to test for evidence of global and local spatial clustering, respectively. Followed by identification of potential pollution sources with geographic analysis.

Results

The Global Moran test of spatial autocorrelation for sample revealed significant clustering (P = 0.0211). The Local Moran test of spatial auto-correlations adjusted for rates variability using both boundary and k-nearest neighbor weights was significant at the spatial scale identifying few local communities as areas of higher risk (P < 0.05). Community 'Sveti Anton' had significant Moran LISA value (P = 0.0454) and defines a core of a spatial cluster extending into neighboring communities Marezige', 'Dekani', 'Pobegi-Cezarji' and 'Vanganel'.

Conclusions

The study identified spatial patterns in distribution of CRD rates. This information is important in guiding decisions on disease control strategies. The study also showed that there is benefit in performing gradual spatial analyses to better define and assess geographical patterns in disease distribution and that Local test for spatial auto-correlation adjusted for rates enhanced visualization of spatial patterns in small area studies. Based on the geographic analysis of pollution sources long range transport of pollutants from Italy and local pollution sources (traffic, industry) were highlighted as most likely risk factors.

KEYWORDS: Spatial epidemiology, Chronic Respiratory Diseases, Cluster analysis, Local Spatial Autocorrelation (LISA), Moran I

IZVLEČEK (Raziskava etioloških faktorjev obolenj dihalnih poti v občini Koper)

Ozadje

Kronična obolenja dihal so naraščajoč zdravstveni problem, s katerim se srečujemo v razvitih okoljih, prav tako v Sloveniji. Pri tem se obolenja kot so astma, kronični bronhitis ali različna alergična obolenja, ki se kažejo s prizadetostjo dihalnih poti, neredko pokažejo že zelo zgodaj v otroštvu. Zato je bil cilj te študije raziskati geografske vzorce pojavljanja obolenj dihalnih poti populacije učencev 1. do 4. razreda v občini Koper, ki leži v jugozahodni Sloveniji in namenom določiti območja povečanega tveganja z vzroki zanje.

Materiali in metode

Podatke o razporeditvi obolelih smo dobili z obdelavo vprašalnikov raziskave »Proučevanje vpliva okolja na pojav določenih bolezni in povečano stopnjo umrljivosti prebivalcev na območju dela Mestne občine Koper« iz leta 2003. V študijo so bili vključeni otroci rojeni med 1991 in 1996, ki so obiskovali 1. do 4. razred osnovne šole. Za iskanje geografskih vzorcev pojavljanja primerov smo uporabili metode prostorske statistične analize, metodo glajenja, analizo kopičenja (globalne indekse prostorske avtokorelacije) in analizo gruč (lokalne indekse prostorske avtokorelacije). Del raziskave smo namenili identifikaciji potencialnih onesnaževalcev z geografsko analizo le teh.

Rezultati

Globalni Moran I test je nakazal statistično značilno kopičenje primerov (Globalni Moran p=0.0211). Lokalni Moran I test prostorske avtokorelacije prilagojen za variabilnost stopenj zbolevnosti daje statistično značilne rezultate (p <0,05), tako ob uporabi sosedskih kot tudi mejnih uteži, za območje KS 'Sveti Anton' (lokalni Moran p = 0,0454), ki s signifikantnostjo in pozitivno prostorsko avtokorelacijo opredeljuje jedro gruče ter se razširja v okoliške krajevne skupnosti ‘Marezige’, ‘Dekani’, ‘Pobegi-Cezarji’ in ‘Vanganel’.

Zaključki

Študija je vizualno prikazala in statistično dokazala obstoj prostorskega kopičenja z višjo stopnjo zbolevnosti dihalnih poti na območju MO Koper. Ugotovili smo, da je smiseln postopen pristop k prostorski analizi, ker na tak način lažje opredelimo in določimo geografske vzorce distribucije bolezni in da z lokalnimi analizami prostorske avtokorelacije prilagojenimi za stopnje zbolevnosti izboljšamo vizualizacijo prostorskih vzorcev na manjših območjih raziskave. Študija je na podlagi geografske analize onesnaževalcev ter dosegljivih podatkov kot najverjetnejše faktorje tveganja izpostavila daljinsko širjenje onesnaževal iz Italije ter lokalne vire onesnaževanja (promet, industrija), ki najverjetneje prispevajo k povečanju tveganja.

KLJUČNE BESEDE: geografska epidemiologija, bolezni dihalnih poti, raziskovalna analiza podatkov, analiza kopičenja, lokalna prostorska avtokorelacija, Moran I

TABLE OF CONTENTS

LIST OF FIGURES:

LIST OF TABLES:

ACKNOWLEDGEMENTS

The research reported here has been supported in part by a contract from Municipality Koper as part of the research by Institute of Public Health Celje.

Additionally, I would like to thank my adviser, PhD. Ivan Eržen, for the great care he has taken to support my research and share his knowledge with me as an adviser and mentor. Special thanks go to my colleagues at work who supported me when work obligations and research didn't go hand in hand.

Finally, I would like to thank my friends and family for their encouragement and patience through the process of completing this thesis.

1 INTRODUCTION

Chronic respiratory diseases (CRD) represent a rapidly growing health problem which is gaining on importance especially in more developed parts of the world. Hundreds of millions of people suffer every day from chronic respiratory diseases. According to the latest WHO estimates (2008), currently 300 million people have asthma, 210 million people have chronic obstructive pulmonary disease (COPD) while millions have allergic rhinitis and other often under diagnosed chronic respiratory diseases. The respiratory diseases like asthma, chronic bronchitis and other allergic respiratory diseases that are affecting airflow are becoming more and more important already in early childhood.

Asthma is often regarded as most prevalent chronic obstructive respiratory disease in childhood and in period of adolescent in population of developed countries, also in Slovenia (Suskovic et. al., 2000, 2005; Kopriva and Macek, 2002). The proportion of children with asthma, aged 6-7 years was in the years 1999-2004 in Austria 7.4% and in children aged 13-14 years 15.1%. The proportion of children with allergic rhinitis and conjunctivitis in children aged 6-7 years was in the years 1999-2004 in Austria 6.1%, in children aged 13-14 years 9.7%. (ENHIS, 2007). It is to be assumed that the values in Slovenia follow a similar trend. The number of hospital admissions for asthma in Slovenia is low (Otorepec, 2009), which does not mean that numbers of asthma patients are low, it shows that children with asthma in Slovenia are well controlled, outpatient and receive appropriate therapy.

Asthma and chronic bronchitis, which are in the beginning of appearance difficult to distinguish because the symptoms are very much alike, are both chronic respiratory diseases. Asthma diagnosis is made on basis of susceptibility test - test of pulmonary hypersensitivity (Tercelj-Zorman, 2006). In case of repeated chronic purulent infections of upper respiratory tract and asthma still couldn't be diagnosed, disease is being diagnosed as chronic bronchitis. In the last decades allergic respiratory diseases are also increasing in adult population but also among children and adolescents. Most commonly those allergic diseases would appear as inflammation of nasal mucosa (Rhinitis allergica) or as polynosis which means that inflammation of conjunctiva, throat and nose appear simultaneously. Both illnesses occur mostly on seasonal basis and can be prevented or mitigated by adequate and on time medical treatment or by avoiding exposure to allergens if possible (Erzen, 1997).

Etiologically, the factors that contribute to higher prevalence of chronic airway diseases have not been entirely explained (Pearce, 1998). The fundamental causes of asthma are not completely understood. The strongest risk factors for developing asthma are a combination of genetic predisposition with environmental exposure to inhaled substances and particles that may provoke allergic reactions or irritate the airways, such as: indoor allergens, outdoor allergens and tobacco smoke, chemical irritants in the workplace and in part air pollution. Other triggers can include cold air, extreme emotional arousal such as anger or fear, and physical exercise. It has been confirmed that asthma will more frequently appear in children exposed to cigarette smoke, those with frequent viral airway infections and those with relatives suffering from allergies. It has also been established that asthma is more frequent in prematurely born children and those with allergic diseases. The connection of asthma and air pollution has not yet been thoroughly researched; the same goes for the influence of the broader environment and that of nutrition in early childhood (Weiss, 1998). Further research will be necessary for more exact determination of causative factors that could be responsible for an increased prevalence rate of chronic respiratory diseases in children.

1.1 Background

Residents of certain areas located in the west and north of the Municipality of Koper, south-western Slovenia, have been concerned because of a high rate of pulmonary diseases in these parts. They warn of episodes of heavy environmental pollution, which they have noticed especially due to a strong smell; they have also observed dust pollution, with some of them developing medical problems, which could be connected with the increased air pollution. Besides analysis of mortality in region Koper has shown higher mortality rate because of respiratory diseases then in Slovenia. Municipality Koper decided to finance a preceding study "The assessment of environmental impact on prevalence of specific diseases and mortality rate in municipality of Koper" (Erzen, 2003).

The objective of the study was to establish whether chronic pulmonary diseases appear more frequently in the part of community considered polluted than in other parts and if there are any differences in mortality rate because of respiratory diseases in different parts of the municipality Koper. The aim of former study was also suggestion of effective measures that need to be introduced in case, that the differences in health status are real.

The results of the preceding study (Erzen, 2003) showed that chronic respiratory diseases in children residing in the areas that were considered more polluted were more frequent than in the areas considered less or not polluted. Besides, an increased risk for developing specific cancer (women lung cancer) was observed among residents in more polluted parts of the municipality of Koper. The study was not designed to investigate the causes for higher morbidity rate, but on the basis of the analyses conducted is reasonable to support the hypothesis of the influence of polluted air on the prevalence of chronic respiratory diseases.

1.2 Geographical analysis

Research on prevalence of diseases and health status, that show a connection to agents in environment and their interlinked causes are main areas of environmental epidemiology (Elliot, 1996, 1997).

Distribution of environmental agents and their influence on health affect various characteristics:

- Sources of environmental agents (natural occurring, anthropogenic),
- Appearance of environmental pollution (continuous, periodic burden),
- Geographical and meteorological conditions (size and form of settlement, factors that might influence intensity of hazard factors).

Assessment of population exposure to diverse hazardous environmental factors and health risk assessment are main topics in environmental epidemiology. First step in the research is identification of hazardous factors followed by exposure assessment of the population as a whole or of specific group of the population, which may be more exposed or more susceptible to the hazardous factor in environment.

The description and analysis of geographic variations in disease with respect to environmental, behavioural risk factors is possible by using methods of spatial/geographical epidemiology (Elliot, 1997, 2001). There is always concern about the accuracy and reliability of results driven from such studies. They are usually much more accurate, if the area which is being investigated is rather small and participants in the study are dwelling predominantly in their place (Elliot, 1997).

1.2.1 Types of geographical analyses

Increased interest of the wider public and professionals, development of geographical information systems (GIS) and geographic-specific statistical tools correlated with development of powerful hardware and software solutions and expanded significantly the scope of geographic statistical tools. Given the content of research we distinguish three basic types of geographic analysis: descriptive studies, studies of clustering and ecological analysis (Lawson, 2003). Methodologically we distinguish between large and small areas geographical analyses (Elliot, 1997).

The basic purpose of descriptive studies is to show the allocation of the disease/cases in selected geographical units. Descriptive maps are often collected in the form of atlases, serving mainly to present the burden of disease and planning regionally specific public-health measures. However, the descriptive studies present the basis of all further analytical studies.

The aim of clustering studies is to define the emergence of small areas (groups) where the risk of disease is greater or lesser than average, thereby taking into account the size and age structure of the population of this area (cluster). Clustering studies may be prepared at regular intervals, when the epidemiologists are looking the public health networks for any signs of individual diseases clustering (called controlled analysis or general analysis) or performed as "post hoc" analysis. Determination of the geographical areas with an increased risk is followed by research of causes for the specifically higher risk – the clustering studies are then compounded by correlation studies.

Geographical ecological analyses or geographical correlation studies investigate the link between the geographical distribution of disease risk and risk factors for this disease on a group (cluster) level. Groups are usually defined by existing administrative units (regions, municipalities, etc.). Together with descriptive analyses the ecological studies assess to what proportion of disease in each region can be explained with the selected risk factors. Due to their rapid and inexpensive implementation in epidemiological research are very popular. On the other hand, when interpreting ecological analysis we have to be aware of the potential ecological bias: present correlation between two variables at the group level does not necessarily mean that there is a correlation at the individual level. Described bias can be reduced by adjusting the size of geographical units, or using appropriate statistical techniques. An essential feature of those techniques is the multivariate modeling by incorporating spatially dependent random variables.

The main problem at geographical analysis of the larger units is the heterogeneity of them. The values of the variables for the selected unit do reflect the average actual value of the individual components of this unit. The unit, which has in one part a large surplus of disease risk (e.g. due to a point source pollution), but otherwise the risk is below average, the analysis will appear as an element with an average risk (the perceived risk is averaged). Due to the averaging also the mentioned ecological bias may occur. Heterogeneity of the large units may be avoided by selecting studies of smaller areas.

By choosing smaller base unit the resolution is increased and the heterogeneity reduced, thus diminishing the statistical significance for rare diseases (in small areas). The probability of occurrence of rare diseases in small area over a large area with a small population is small and thus statistically non-significant, so the results of further analysis may be a result of chance (Elliot, 1997). Statistical significance may be increased by extending the observation period, or using one of the techniques of smoothing, where with the actual number of cases in each unit, data on the incidence of neighboring areas and other relevant information, assess the risk of disease in the unit (Elliot, 2001).

1.2.2 Challenges in the analysis of observational data

However we have to be aware of other occurring problems in such analyses connected to the observations of disease occurrence in a population of people "at risk." Such observational studies differ in several important ways from experimental studies common in other fields of scientific inquiry and step in where experimental studies are infeasible due to expense or ethical concerns. Observational studies require more care in designing, analysis and interpretation, since controlled environments and randomization often are not possible. Consequently, observational studies involve potential for a wide variety of misinterpretation. The nature of observational studies, particularly of epidemiological studies in the investigation of determinants of disease, provides a framework for interpretation for most spatial analyses of public health data. Central to this framework is the quantification of patterns in the frequency of disease occurrence among members of the population under observation.

However, special concerns arise in the observational setting, which often assume an experimental setting in model development and interpretation. Two common issues in epidemiological analysis are extremely important: bias and confounding (Rothman and Greenland, 1998; Waller and Gotway, 2004).

Bias

In an experimental setting, researchers control experimental conditions that could influence the observed associations between treatment and outcome, and assume that the randomized assignment of treatments to experimental subjects (randomization) allows unbiased estimation of any uncontrolled conditions affecting the association of interest. Observational studies, by their nature, do not randomize treatment; rather, the researcher observes treatment levels in study participants often for ethical reasons. As a result, observational studies contain many potential biases in their estimates of associations between treatments (exposures) and outcomes.

Selection bias

Selection bias occurs when there is a systematic difference between the characteristics of the people selected for a study and the characteristics of those who are not. In studies of children's health, where parental cooperation is required selection bias may occur. An important selection bias is introduced when the disease or factor under investigation itself makes people unavailable for study. All epidemiological study designs need to account for selection bias.

Measurement bias

Measurement bias occurs when the individual measurements or classifications of disease or exposure are inaccurate – that is, they do not measure correctly what they are supposed to measure. There are many sources of measurement bias and their effects are of varying importance. A form of measurement bias of particular importance in retrospective case-control studies is known as recall bias. This occurs when there is a differential recall of information by cases and controls; for instance, cases may be more likely to recall past exposure, especially if it is widely known to be associated with the disease under study. Recall bias can either exaggerate the degree of effect associated with the exposure or underestimate it.

Confounding

Confounding is a central notion in epidemiology, defined as "a distortion in the estimated exposure effect that results from differences in risk between the exposed and unexposed that are not due to exposure" (Rothman and Greenland, 1998). Confounding involves the biases above, but tends to focus on factors associated with the bias rather than the bias per se. That is one often refers to confounding factors or confounders as the source of differential selection, recall, or misclassification. Rothman and Greenland (1998) define two necessary conditions for a variable to be a confounder:

- The variable must be a risk factor for the disease among the unexposed (although it may not cause the disease directly).
- The variable must be associated with the exposure variable in the population providing study participants.

Suppose that we observe outcomes in subjects experiencing one level of exposure and wish to compare them to unexposed subjects. If the unexposed participants differ from the exposed participants with respect to a factor related to the disease (e.g., age), an estimate of the effect of exposure on outcome by comparisons of outcome proportions in the two exposure groups can (often, will) be different from the true effect.

1.3 Geographical analysis of studied area with emphasis on environmental factors

In the scope of the previously carried out research »The assessment of environmental impact on prevalence of specific diseases and mortality rate in Municipality of Koper« (Erzen, 2003) the following two studies were carried out »Climatic conditions on the Slovenian coast« and »Survey of the data regarding the state of the environment on the Slovenian coast«. Below is a short summary of the results of the said studies which are important for understanding the conditions in the area of the Municipality of Koper.

1.3.1 Climatic conditions

CLIMATE

For Slovensko Primorje (Slovenian Littoral) a strong influence of the sea climate is typical, and at the same time also big gradients when passing to the hinterland. Already in the area between the coast and the karst level big differences are noticeable. The coast is well-indented and especially the wind conditions as well the air temperature are locally strongly modified. In the past there was no systematic monitoring of climatic conditions in Slovensko Primorje provided by a reference climatological station that would be operating in the same location and in unchanged environment. In the past nobody thought that we would encounter climatic changes so substantially surpassing changes seen in the previous centuries. Furthermore, not enough attention was dedicated to defining gradients on the coastal hills, and to the differences between sunny and shady sides.

TEMPERATURES

On average Slovensko Primorje is the warmest part of Slovenia. The average summer temperature amounts to 23.1° C, while the average winter temperature is 2.3° C. Particularly in areas close to the coast the annual and daily movements in temperature are smaller than in areas with a strongly expressed continental climate. The sea temperature drops and rises more slowly than the temperature of land and the air above land. In autumn the temperature on the coast drops at a slower pace than in the hinterland, ground temperature inversions at sea are not so frequent (on the ground of valleys a couple of 100 m from the coast low morning temperature inversions are much more frequent than close to the sea), and in spring the highest daytime temperature does not rise as quickly as e.g. in Vipavska dolina (Vipava valley). The penetration of cooler air on the coast is not as distinctive as in the hinterland, however, the air temperature is strongly influenced by the wind direction (if it blows from the sea or land).

Also in Slovensko Primorje the air temperature strongly depends on the geographic characteristics and the distance from the sea. The average temperature is the highest at the sea level, and with increasing altitudes the temperature decreases. The difference between the coast and the areas closer to the karst level is the biggest in autumn when the sea is still relatively warm, while the land is cooling off faster. In spring the difference in the average air temperature is the smallest when the sea temperature is rising only slowly. On the coast the air temperature almost never drops to –10 °C, and there are almost no ice days, i.e. days when the temperature remains below the freezing point for the entire day.

Even more than on the average temperature, micro location has an important influence on the lowest daytime temperature, and it is in the cold half of the year when the differences are the most evident. Places situated in poorly aerated areas or at the bottom of a concavely-shaped surface have significantly lower temperatures than places on slopes or hilltops. On the coast January and February mornings have the lowest average temperatures, while in the second July decade the temperatures are the highest. Shallow ground inversions appear at night and in the morning, closer to midday they disintegrate, and these inversions must be distinguished from the widespread temperature inversions that expand over Slovensko Primorje from the Po river plain and the northern Adriatic, such inversions that result from the general synoptic situation can persist for several days and influence the entire Slovensko Primorje.

PRECIPITATION

An important indicator of climatic conditions is precipitation. For precipitation a great variability is common which is larger in annual precipitation levels than in monthly. Even though the variability is large, lower average values can be seen in the last twenty years. The amount of precipitation is expressed in height in mm, while the frequency is expressed in the number of rainy days above a certain level. The amount of precipitation varies strongly from year to year, and the month-to-month differences are also substantial. In order to define the precipitation regime a longer set of data is necessary, usually a thirty-year set of data. The amount of precipitation increases from the coast towards the karst level as this is due to the increasing altitude. Like the largest part of Slovenia, Slovensko Primorje also gets the most abundant precipitation at south-western airflows.

The long-standing average precipitation at the coast is 1000 mm, at the foot of Podgorski kras (at the altitude of 200 to 300 m) the rainfall is about 1200 mm, and in the western part of the plateau (altitude around 500 m) the rainfall exceeds 1500 mm. Slovensko Primorje has a modified Mediterranean precipitation regime with the most abundant rainfall in autumn. On average, the coast has the most rainfall in the months of September and October. The biggest year-on-year changeability is in October. In summer months there is a lot of precipitation during storms, also at the coast the maximum frequency of storms is in summer, while in winter there are almost no storms.

WIND

In addition to local winds blowing from the land to the sea and from the sea to the land, general winds are also quite common for the coast.

The most common winds are:

- Jugo blows along the coast at prevailing south-western winds, and along the Adriatic coast it turns into a south-eastern wind; it does not reach such high speeds and it is not as gusty as Bora. It usually blows prior to weather deterioration,
- Bora blows from the north-east, and it is a very gusty wind. Anticyclonic Bora wind brings cold and dry air, while cyclonic Bora brings moist air. As regards the effects on the feeling of people the two winds differ substantially, however, for the purpose of assessing the transfer and thinning of impurities in the air they can be joined.

In Slovensko Primorje several directions are evident, only in winter there is a marked prevalence of Bora, while in other seasons north-western wind is quite frequent, and in spring also south-western wind. The average daytime wind speed reaches its peak at midday. During the day the average highest wind speeds are in spring and in summer, the least expressed are the changes in speeds in winter. The minimum speeds are at sunrise and sunset.

FOG

In addition to reduced visibility (according to definition horizontal visibility is at fog less than one kilometre), fog is also important as an indicator of inversion and very weak air movements. The most important in the coastal area is the fog spreading over Slovensko Primorje from over the northern Adriatic and the Po river plain, and accompanied by a strong temperature inversion. Such fog can reach some 100 m high, sometimes the air layer mixes right above the ground and the fog lifts to low cloudiness. It can persist for several days because it can be dispersed only by the inflow of cooler air, usually at the onset of a cool front. On the coast, fog is most frequently seen in winter months, very rarely it can persist for the entire day even at the start of March.

INSOLATION AND CLOUDINESS

The duration of insolation is expressed in the number of hours of insolation per day, per month, and per year. Geographic and astronomic factors, and, of course, weather conditions have the biggest impact on the duration of insolation. The coast belongs to the sunniest parts of Slovenia. The sunniest weather is in July when the atmosphere is the most stable and less cumulus clouds are formed. In July the average insolation in the coastal areas is more than 10 hours. Usually the least sunny weather is in December when the light part of the day is the shortest.

Cloudiness is closely connected with sunny weather. Cloudiness influences the warmth of the ground obtained and released through insolation. Climatological data do not contain the data about the thickness and the type of clouds. The average cloudiness is at the minimum from July to September, and at the maximum from November to January. In November the changeability of cloudiness is increasing from year to year if compared with the remaining months.

INVERSIONS

Shallow ground inversions (inversions connected with the cold air above the northern Adriatic stretch across the entire area and there are no bigger visible local differences) in the examined area are not very distinctive; in periods when the sea is warm they cannot be observed close to the sea (in summer, in autumn, frequently also in winter). However, they appear at the bottom of valleys and in areas with a concave surface. It is almost impossible to define an entire local community with frequent inversions as the municipality has a heterogenous relief. It is estimated that shallow ground inversions are the most frequent in Vanganel and in the flat part of its surrounding area, along the Rižana valley, in the Dragonja and Rokava valley. Particularly in places where the valley is partly broadened. Settlements on tops of hills or ridges are not areas affected by ground inversions. In addition, there are shallow ground inversions in the broader lowland area of the loop between Bertoki, Dekani and Škofije. Šalara and Olmo are also appropriately situated for the occurrence of shallow ground inversions.

1.3.2 State of the environment

It was not planned for the research to include measurements of the state of the environment but rather to use the existing data collected in the scope of other studies and by regular monitoring of the conditions.

WATER

The supply of potable water in the Municipality of Koper is well organized. The Rižana water station was built in 1997 and is located in the settlement of Cepki, approximately 50 m above the main road Koper - Kozina. The technological procedure of ultrafiltration was selected for the preparation of potable water.

Outside of the peak tourist season the system supplies water to approximately 80,000 inhabitants in three coastal communities: Koper, Izola, Piran, and during the main season to over 120,000 people. During the dry summer months when the supply of water from Rižana is insufficient and the quantity of potable water does not meet the requirements, the Rižana water station provides the missing quantities of potable water through consumption of water from Kraški vodovod. In order to ensure the daily exchange of water in the network, the consumption of water is up to 130l/s in summer and about 10 l/s in winter. The potable water from Kraški vodovod is supplied to the residents of Miljski hribi and is disinfected with chlorine gas, unlike the potable water from Rižana water station.

A total of 108 out of 123 settlements are connected to the public water supply network. This means that 99 % of the coastal area is connected to the public water supply system (Rižanski vodovod, 2007). About 900 residents have no connection to the water supply system. Residents without public water supply get their water delivered in tanks or from local water sources. All public objects (kindergartens, schools, health institutions…), the food industry and other industry as well as other institutions of general interest are connected to the water supply system. The area provided with the water supply system has a sewage network. There are still some settlements without a sewage network. In these the effluent is discharged into individual or common cesspits. There are, however, individual objects without a cesspit.

The health compliance of potable water is established through water sample analyses of microbiological and physicochemical parameters in accordance with the provisions of the Rules on the Health Adequacy of Potable Water. Pursuant to the data provided by the Koper Health Institute, which regularly monitors the quality of potable water, the water supplied through Rižanski vodovod to 99% of the residents of the coastal region is health compliant.

AIR

The most significant air pollutants in Koper are industry, traffic and heating systems. Industry and traffic pollute the air over the entire year. Until the year 2004 there were only measurements of 24-hour smoke concentrations and of the acid gas index carried out in the Municipality of Koper in the scope of the state monitoring of air quality. Based on these measurements Koper was until that year one of the places with the least polluted air in Slovenia of the approximately fifty places where the measurements had been carried out. The limit value of immission concentration had actually never been exceeded. The general estimate that the air quality in the Koper area is good was the reason why the Environmental Agency of the Republic of Slovenia did not decide to expand the range of parameters so as to monitor the immission air quality.

In 1996 measurements of other pollutants in the air were carried out in Koper area (Ankaran crossing, Gabrovica na Črnem Kalu, Koper- see Figure 2) by means of a mobile measuring station. From the measurement results obtained in Ankaran crossing and in Črni Kal it is obvious that during the measuring at measuring places Ankaran crossing and Gabrovica na Črnem Kalu the hourly as well as the 24-hour limit ozone immission concentrations were exceeded. Immission concentrations of most of other pollutants were significantly lower than the limit values defined in the Regulation on Limit, Warning and Critical Immission Values of

Substances in the Air (Off.Gazette of the RS no. 73/94). In the abstract of the report on measurements in Koper (Air pollution in Slovenia in 1996, Hydrometeorological Institute of Slovenia, 1997) it is stated that the air pollution in Koper was insignificant in September 1996, only ozone pollution of air exceeded the limit 8-hour concentration twofold.

Measurements carried out in 2000 and 2003 (ARSO, 2002, 2004) have shown that in Koper and its broad surrounding area (Semedela, Dekani, Ankaran- see Figure 2), there is a serious problem of ozone air pollution, at least in the hot summer months. In the relatively short period of time when the measurements were performed the limit immission concentrations were exceeded several times, in Dekani two measurements had shown that even the critical immission values were exceeded. The pollution of air with highly volatile organic substances (benzene) in Dekani represents a special problem, which has not been well enough explained so far.

The high ozone concentrations recorded in the area of Municipality of Koper show that the problem of air pollution in this area is not marginal. Ozone in the air in urban environment is an unquestionable indicator of pollution of air with organic substances. This means that the air in Municipality of Koper is polluted but not with common pollutants as a result of the use of fossil fuels for energy generation. The high ozone concentrations and the presence of increased contents of highly volatile organic substances indicate that increased traffic and economic activity where there are emissions of highly volatile organic substances are a significant source of pollution. Considering the well-known air pollution in neighboring Italy it is obvious that the long-range transport contributes a great amount of pollutants. This presumption is confirmed by renewed measurements of air pollution from 2005 (ARSO, 2006) carried out in the cities of Koper and Ankaran, which also show that there was an increased ozone concentration. The highest concentrations are at western and north-western wind directions (Maestral), which, due to the fact that there is only sea in that direction, indicate that the pollution is transported from neighboring Italy. Based on the measurements made in 2005, the area of Koper is together with Nova Gorica the most ozone-contaminated area in Slovenia. At the time of measurements, air contamination with PM_{10} particles had exceeded for over 35 days the permissible 24-hour concentration (annual number of overstepping was exceeded) at the measuring places in Koper as well as in Ankaran.

In June 2005, the Agency of the Republic of Slovenia for the Environment set up together with the Municipality of Koper a permanent measuring station in Markovec to measure air pollution in Koper. The measurements carried out in 2006 (ARSO, 2007) show that in all measuring stations in Slovenia, the highest ozone concentrations with the most exceedings of the alert 1-hour value appear in Primorska region and on the coast.

In order to clarify the current problems of high ozone concentrations in summer time, an ecological and meteorological mobile station was set up from May 2007 to June 2008 in Lovran (officially Kolomban) situated in the local community of Hrvatini. The measuring station is located on the sparsely inhabited top of the peninsula (altitude of 156 m), separating the bays of Koper and Trieste, which is open to all sides. The measurements of pollutants in external air in the location of Lovran indicate that this area is excessively ozone polluted in summer. This is also true of the greater Primorska region, which is confirmed by the measurements in the same period in the permanent measuring stations in Koper, Nova Gorica and on Otlica. Critical conditions with regard to ozone occur in stable weather, when weather deterioration approaches slowly from the west, and weak western winds prevail over our region – along the coast the weak maestral wind prevails, while in the hinterland of Primorska the weak south-west wind blows. In these cases, it is a matter of transfer of the ozone-polluted air from the densely populated and industrial northern Italy with greater sources of emissions of ozone precursors. It is not known what portion of the measured concentrations can be attributed to this transfer because detailed analyses of emission and meteorological parameters would be required (ARSO, 2008a).

Some findings about this can be found in the study “Numerical Modelling and Scenario Analysis of Pollution Extension near Lovran Station” (ARSO, 2008b). On the basis of simulations with the numerical meteorological model ALADIN and subsequently with the transport model FLEXPART, the authors have established that in the described nine episodes with the highest measured ozone concentrations in the station Lovran it was mainly a matter of influence from near sources in Koper and Trieste. In support of this assumption are the strongly correlated values of modelled pollution concentrations from pre-set sources with the measured progress of ozone concentration. The direct influence of more distant sources is temporary. In some situations, a part of the measured concentrations can be attributed to sources in Monfalcone, more rarely to those in Umag. In these medium-distant influences, it is interesting that the 12-hour shift between concentrations of modelled and measured values cannot be noticed, which probably means that ozone is already present in the air mass that comes to Lovran from Monfalcone or Umag.

Air contamination in Koper is despite everything an environmental problem which has obviously been present in Koper for an extended period of time, however, it has become more highlighted and discussed about in recent years, when the environmental awareness of residents has grown as a result of better recognition of environmental pollution. Procedures were launched in the city council representatives of residents to recognize environmental problems and to improve conditions where this in necessary.

2 PURPOSE AND OBJECTIVES

The work proceeds former research conducted in 2003, when a study of Prevalence of chronic respiratory diseases among children from variously polluted areas of the municipality of Koper was carried out (Erzen, 2003). The purpose of present study is to assess the correlation (relationship) between etiologic factors of chronic respiratory diseases and living environment of children (1^{st} to 4^{th} grade of primary school) on the basis of the observed cases frequency in administrative units of municipality Koper. This will enable specific and more effective measures for reducing air pollution that will lead to reduction of negative health consequences due to airborne agents.

Objectives of the study are:

- To find out whether there is significant clustering of chronic respiratory diseases among children in certain small geographic areas,
- To assess the patterns (clusters) of chronic respiratory disease cases appearance in order to define areas of higher risk and,
- Establishment of potential environmental etiological factor liable for higher morbidity of chronic respiratory diseases

The specific purposes of this work are twofold: First, its results will directly outline the areas with larger than average prevalence rates and provide insight into possible agents. Second, the lessons learned from evaluating the example will provide possible solutions to solving other small scale spatial epidemiology puzzles.

3. APPROACH AND METHODS

The presented analysis is based on survey data collected in epidemiological survey on prevalence of chronic respiratory diseases among children from variously polluted areas of the municipality of Koper, carried out in the period 2002 – 2003 (Erzen, 2003). Methodologically the analysis falls into a descriptive study of small geographic units and is conducted as geographical ecological study. In order to assess the geographical distribution of chronic diseases in the Municipality of Koper data from the previous epidemiological research were reprocessed and used.

In the first part of Chapter Approach and methods we present methods behind data obtained in the second part we present the methods of data processing. Spatial analyses were used with an intention of identifying and assessing geographical patterns of CRD in Koper. ESDA methods, Global Moran's I and Moran Local Indicators of Spatial Associations were used to test for evidence of global and local spatial clustering. Followed by geographic analyses of air pollution sources with aim of analysing connections between agents and living environment of cases.

3.1 Sources of collected data

3.1.1 Data on cases of chronic respiratory diseases

Children born between 1991 and 1996 attending 1st to 4th year of primary school in the territory of the Municipality of Koper were included in this research. Teachers distributed envelopes containing questionnaires to those participating in the research, which they took to their parents or guardians to be filled in. The questionnaire included information on living conditions and leisure, on health status in relation to respiratory diseases and other factors that can be associated with the occurrence of chronic respiratory diseases. The filled-in questionnaires were returned by the parents to teachers in sealed envelopes. If there was a dilemma as to the diagnosis of the disease (entered in the questionnaire) the personal doctor was contacted with the parents' or guardians' permission in order to provide a diagnosis.

Before the beginning of data processing, certain personal details were deleted. For further processing, the following details were taken into account: questionnaire consecutive number, address and municipality of permanent residence and health condition (diagnosis).

The address of permanent residence (indirect spatial data) served as a geocoding identifier, on the basis of the Environmental Atlas of Slovenia database (ARSO). Geocoding is a process of matching or rather associating a location and the address of permanent

residence. Coordinates X and Y (Gauss-Kruegger coordination system D48/GK) from freely accessible Environmental Atlas of Slovenia database were added to the created database.
On the basis of the permanent residence (details acquired from the questionnaire) and the census of children born between 1991 and 1996 (data acquired from the Municipality of Koper), the local community of residence and postcode were additionally identified. The database, i.e. the relational database acquired in this way contains the following details: questionnaire code, geocoded location with the coordinates X and Y, local community (LC), postcode (PC) and health condition. On the basis of these details, the level of prevalence of asthma and chronic bronchitis as well as of allergic diseases and of all diseases in total has been calculated and a database with the prevalence of individual disease conditions has been prepared.

With the purpose of achieving a high level of matching and direct precision control of geolocation, it was required to re-check the geocoded coordinates of addresses on the basis of the control database, which, in Slovenia, is the EHIS database (register of house numbers, dBase, GURS – Surveying and Mapping Authority of the Republic of Slovenia).

3.1.2 Description of the examined area

The Municipality of Koper lies in the south-western part of Slovenia, and is composed of 23 smaller spatial (administrative) units - local communities. Municipality Koper encompasses 311.2 km^2 of land in the geographic region of Slovenska Istra (Slovenian Istria), and it is the largest of three Slovenian coastal municipalities. According to the data from the Central population register from June 2007 the municipality has 49,595 inhabitants, of which 23,374 live in city of Koper (STAT, 2007). The population density in the city is 158 persons per square meter (Figure 1).

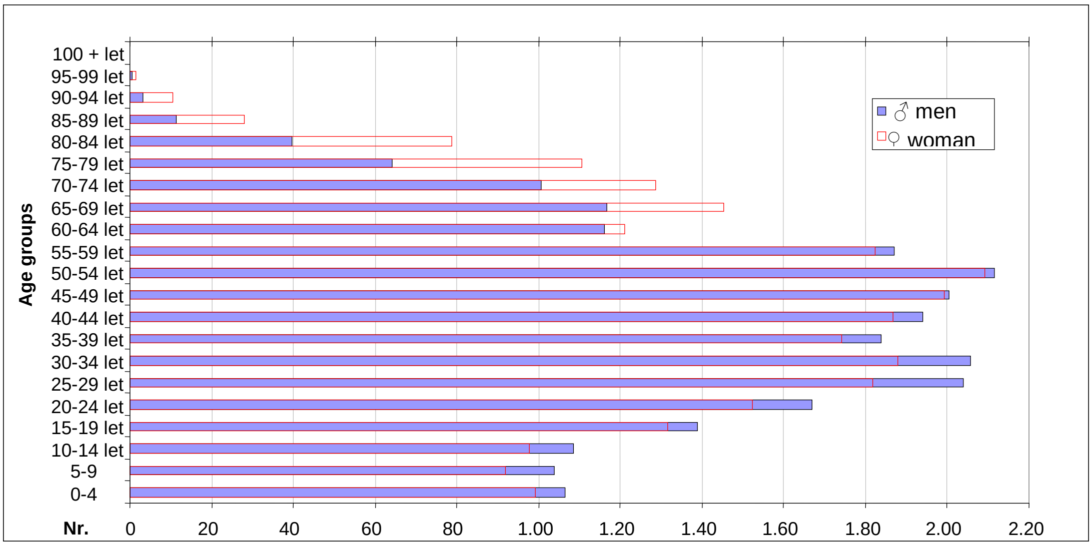


Figure 1: *Population of Municipality of Koper by five-year age groups and by sex* (Source: Municipality of Koper, 2007)

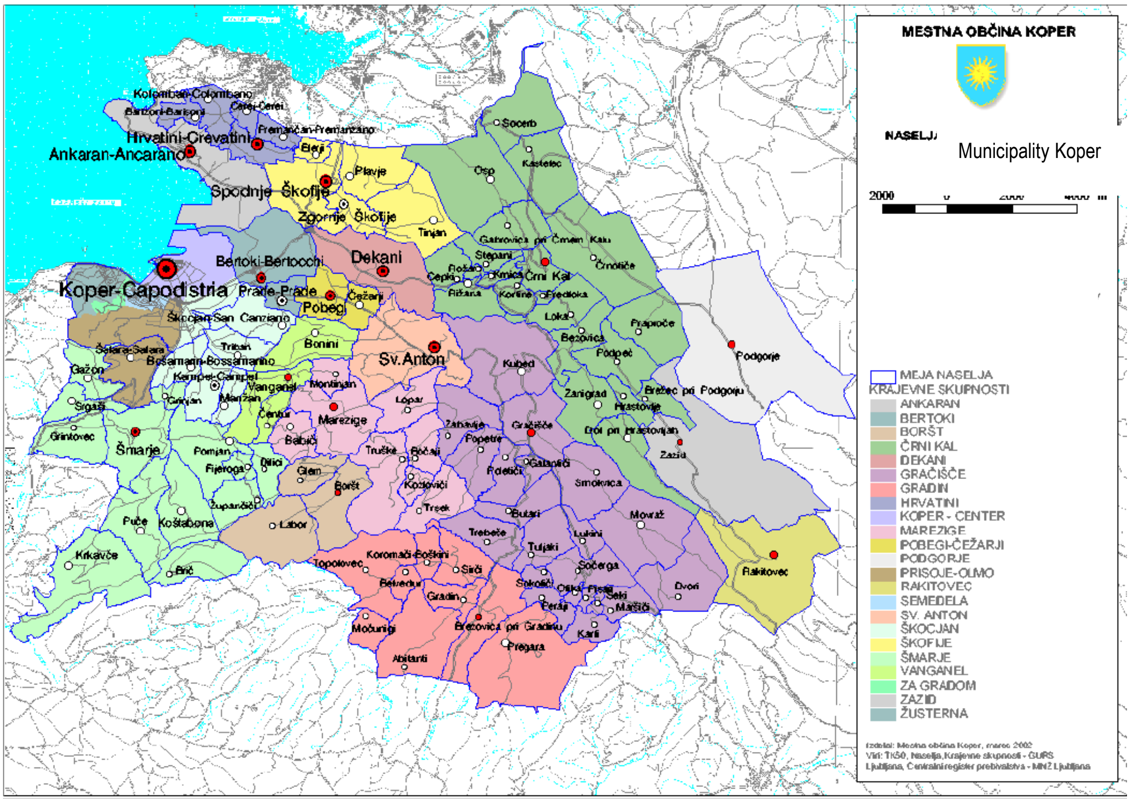


Figure 2: *Settlements and local communities in Municipality of Koper* (Source: Municipality of Koper, 2007)

The municipality is divided into 23 local communities (administrative units) with 105 settlements (Figure 2) under the Ordinance on establishing the areas of local communities in the Municipality of Koper (No. official publication 12/96). Ordinance was adopted on the basis of the Statute of the Municipality of Koper.

3.1.3 Analysis period

The first survey was conducted in years 2002 – 2003 on basis of questionnaires for Children born between 1991 and 1996 attending 1^{st} to 4^{th} year of primary school in the territory of the Municipality of Koper (Erzen, 2003). This geographic analysis followed and was conducted between 2007 and 2009.

3.1.4 Privacy

In preparation, execution and presentation of analysis, we strictly followed the Criteria for maintaining the confidentiality of personal data (Official Gazette of RS, no. 59/99, no. 57/01 and 59/01). The study was approved by the Commission for Medical Ethics at the Ministry of Health of the Republic of Slovenia.

3.2 Methods of data processing

In the subsequent sub-section first we present the methods we used in the preparation of the data, followed by the presentation of the methodology by which we analyze the data processed.

The analysis of spatial distribution of public health data involves more than just observational and visual inference. Medical science provides insight into some specific causes of disease, however much remains unknown. As a result, the analysis of public health data often builds from the statistical notion of each person having a risk or probability of contracting a disease (Waller et.al, 2004; Elliot et. al., 1997, 2001).

Therefore central role of probabilities motivates the use of statistical methods to analyze public health data and the use of spatial statistical methods to:

- Evaluate differences in rates observed from different geographic areas,
- Separate pattern from noise,
- Identify disease "clusters," and
- Assess the significance of potential exposures.

The analytic goal involves identification and quantification of any exposures, behaviours, and characteristics that may modify a person's risk. These methods also allow us to quantify uncertainty in our estimates, predictions, and maps and provide the foundations for statistical inference with spatial data. Some spatial statistical methods are adaptations of 'nonspatial' methods, other methods are used to visualize spatial data, make meaningful maps, and detect spatial patterns. Spatial statistical methods have to take the spatial arrangement, and the resulting correlations, of the observations into consideration in order to provide accurate, meaningful conclusions. Applying statistical methods in a spatial setting raises several challenges. Geographer Tobler summarized a key component affecting any analysis of spatially referenced data through his widely quoted and paraphrased first law of geography: "Everything is related to everything else, but near things are more related than far things" (Tobler, 1970, 1975). This law succinctly defines the statistical notion of (positive) spatial autocorrelation, in which pairs of observations taken nearby are more alike than those taken farther apart.

3.2.1 Prevalence of the disease

The study of disease in a population begins by addressing the occurrence of a particular outcome in a particular population over a particular time. *Prevalence* refers to the total number of *existing* cases over a specific time frame and provides a summary of the current burden of the disease under study within the population.

In general use the term *rate* defines the number of occurrences of some defined event per unit time. In contrast to the precise epidemiological use of rate, the spatial epidemiology literature tends to use disease rate to refer to the number of incident cases expected per person rather than per unit of person-time. That is, this use of disease rate refers to the total number of cases observed divided by the total number of people at risk, both within a fixed time interval.

In a closed population (no people added to or removed from the at-risk population during the study period) where all subjects contribute the same observation time, the proportion would be equal to the rate multiplied by the length of the (common) observation time for each person. Some difference between the two quantities always remains since a person stops contributing person-time to the denominator of the incidence rate the moment that person contracts the disease. However, this difference between the rate and proportion diminishes with rare diseases in the population at risk and/or short observation time per person (i.e., with less loss of observed person-time per diseased person) and provides convenient numerical approximations (Waller and Gotway, 2004).

3.2.2 Spatial Analyses within GIS

Many authors note the analytical and predictive capabilities of GISs. Some (e.g., Bailey and Gatrell 1995) distinguish between spatial analysis, the study of phenomena occurring in a spatial setting using the basic GIS operations outlined above, and spatial data analysis, the application of statistical description and modeling to spatially referenced data. The distinction provides a boundary between standard GIS operations and queries and the application of data analysis algorithms for estimation, prediction, and simulation familiar to many statisticians. Other authors use the term spatial analysis quite broadly as a field and include inferential statistics among the tools. For instance, Longley et al. (2001) describe six general categories for GIS-based spatial analysis: queries (enabled by the spatial relational database underlying the GIS), measurements (e.g., length, shape, distance), transformations (e.g., spatial joins, and conversion from vector to raster, or vice versa), descriptive summaries (e.g., calculating the mean response in a particular area), optimization (e.g., searching for minima/maxima across a spatial area), and hypothesis testing. The last category specifically involves inferential statistical methods, while the other categories describe a variety of tools for quantifying trends, features, and patterns within a spatial data set.

Most modern GIS's (Lee and Wong, 2001) allow little in the way of routine calculations for spatial data analysis involving statistical inference. As a result, most analyses described in subsequent chapters involve the use of separate statistical packages and are not analyzed within a GIS setting. Scripts and growing variety of applets (Web-based application tools) for both statistical and GIS computing provides toolboxes for further development.

3.2.3 GeoDa

GeoDa 0.95i (Anselin, 2004) is the latest version in a line of software tools developed by Dr. Luc Anselin's Spatial Analysis Laboratory (SAL) in the Department of Geography at the University of Illinois, Urbana-Champaign. It is designed to implement techniques for exploratory spatial data analysis (ESDA) on lattice data (points and polygons). The design of GeoDa consists of an interactive environment that combines maps with statistical graphics, using the technology of dynamically linked windows. Its origins trace back to early efforts to develop a bridge between ESRI's ArcInfo GIS and statistical software. Its immediate precursor was the DynESDA extension for ArcView 3.x, which introduced linked windows and brushing in a GIS environment. In contrast to the extension, the current software is freestanding and does not require a specific GIS system.

The free program provides a user friendly graphical interface to methods of descriptive spatial data analysis, such as spatial autocorrelation statistics, as well as basic spatial regression functionality. The latest version contains several new features such as a cartogram, a refined map movie, parallel coordinate plot, 3D visualization, conditional plots (and maps) and spatial regression. It is geared to the analysis of discrete geospatial data, i.e., objects characterized by their location in space either as points (point coordinates) or polygons (polygon boundary coordinates).

The current version adheres to ESRI's shape file as the standard for storing spatial information. It uses ESRI's MapObjects LT2 technology for spatial data access, mapping and querying. The analytical functionality is implemented in a modular fashion, as a collection of C++ classes with associated methods.

In broad terms, the functionality can be classified into six categories:

- spatial data manipulation and utilities: data input, output, and conversion
- data transformation: variable transformations and creation of new variables
- mapping: choropleth maps, cartogram and map animation
- EDA: statistical graphics
- spatial autocorrelation: global and local spatial autocorrelation statistics, with inference and visualization
- spatial regression: diagnostics and maximum likelihood estimation of linear spatial regression models

The full set of functions is documented in detail in the GeoDa User's Guides (Anselin, 2004, 2005, 2006). The software implementation consists of two important components: the user interface and graphics windows on the one hand, and the computational engine on the other hand. In the current version, all graphic windows are based on Microsoft Foundation Classes (MFC) and thus are limited to MS Windows platforms. In contrast, the computational engine (including statistical operations, randomization, and spatial regression) is pure C++ code and largely cross platform.

3.2.4 Mapping

The typical goal of mapping public health data is to provide insight into geographic variations in disease risk. Risk is simply the probability that an unfortunate event occurs. In public health, the unfortunate event is usually the contraction of or death from a specific disease. When we make a map of disease counts, proportions, or rates, we are ultimately intending to convey inferences about disease risk. However, a map of raw counts is not the best tool for inference about disease risk, since we expect regions with larger populations to have higher disease counts. We can account for population differences by using rates (disease incidence per person per time) as measures of risk. Higher disease rates reflect greater chances for contracting the disease, and thus, viewed this way, rates reflect a person's risk for disease. However, a map of rates may still obscure the spatial pattern in disease risk, particularly if the rates are based on populations of very different sizes. Since the variability in the estimated local rates depends on population size, some rates may be better estimated than others, and this may obscure spatial patterns in disease risk. Rates based on small populations or on small numbers of disease cases are likely to be elevated artificially, reflecting lack of data rather than true elevated risk; often referred to as the small number problem (Elliot, 1997; Waller and Gotway, 2004).

There are several solutions to this problem. First, we could calculate rates over larger areas (e.g., use states instead of counties), although this comes at the expense of giving up some of the geographic information we wish to convey. Second, we could make a comparative map, one that compares each rate to a common measure and, in doing so, adjusts for different population sizes. One such map is a probability map.

3.2.5 Smoothing

Yet another approach is spatial smoothing, method for reducing the noise in rates associated with geographic regions. Spatial smoothers are analogous to scatterplot smoothers in regression analysis and to moving-averaging methods in time series, adapted to two dimensions. The basic idea is to "borrow" information from neighboring regions to produce a better (i.e., more stable and less noisy) estimate of the rate associated with each region and thus separate out the "signal" (i.e., spatial pattern) from the noise.

Advantages and disadvantages of smoothing:

There are two main advantages to smoothing rates. The first is that smoothing stabilizes rates based on small numbers by combining available data at the resolution of interest. We do not have to aggregate to larger regions to achieve stable rates for mapping. The second advantage to smoothing is that it reduces noise in the rates caused by different population sizes, thus increasing our ability to discern systematic patterns in the spatial variation of the underlying risk.

Smoothing also has some disadvantages. Smoothed maps are maps of values other than the raw data values. Another disadvantage of smoothing is that it can introduce artifacts and autocorrelation into the rate map (see, e.g., Gelman and Price, 1999; Gelman et al., 2000). Thus, smoothing may replace one set of artifacts (e.g., unstable estimates) in mapped rates with another set of artifacts (e.g., correlated estimates) in altered rates. Still a lot of authors (Kafadar, 1994, 1996; Lawson, 2001) believe that the visual inferences obtained from maps of raw rates can often be so misleading that smoothing is a better choice.

On the other hand, if we are not looking for individual regions (e.g., counties, tracts) with elevated rates, but instead want to get a general assessment of broad trends and patterns, smoothing will help reduce the noise and make the trends and patterns more clear. Smoothing can reduce our attention to large rates that may be outliers by focusing it on the overall picture.

3.2.6 Smoothers

Locally Weighted Averages

We can obtain a smoothed value for each region by simply averaging the values associated with neighboring regions. With distance smoothing, a circular or disk smoothing window of specified radius is centered at the centroid of each region. The smoothed value for each region is then taken to be the average of all the values associated with centroids that lie within the disk. Thus, if r_1, r_2, ... ,r_N are the observed rates, smoothed rates can be calculated as

$$\tilde{r}_i = \frac{\sum_{j=1}^{N} w_{ij} r_j}{\sum_{j=1}^{N} w_{ij}}, \qquad \text{where the weights are given by} \quad w_{ij} = \begin{cases} 1 & d_{ij} < d \\ 0 & otherwise \end{cases} \qquad (1)$$

Here d_{ij} is the distance (Euclidean, city-block, or any other distance metric) between the centroids of regions *i* and *j* and *d* is the disk radius.

Instead of using circular disks of constant radius defined by geographic proximity, Talbot et al. (2000) recommend the use of moving windows defined in terms of constant population size. More generally, we may want to define a smoothing neighborhood that identifies, for each region, the set of regions whose values are to be used in the local average – k-nearest neighbor smoothing. A common approach to defining smoothing neighborhoods is through a spatial proximity measure that specifies the neighborhood structure and provides suitable weights for the values being averaged – rook based smoothing. Different smoothers result from different choices for the weights. Example of contiguity weights using adjacency to define neighbors. Thus, we construct weight if regions *i* and j share common boundary.

Empirical Bayes Smoothing

The smoothing approaches outlined earlier "borrow" information from nearby regions to stabilize local estimates through the use of various weighting schemes. One set of smoothed rates results from a set different weighted average of neighboring rates.

Another, somewhat more formal approach uses probability models to obtain smoothed estimates consisting of a compromise between the observed rate for each region and an estimate from a larger collection of cases and persons at risk (e.g., the rate observed over the entire study area or over a collection of neighbouring regions). The compromise combines the rate from each region, which can be statistically unstable due to the rarity of the disease and the relatively small number of people at risk, with an estimated rate from a larger collection of people which is more statistically stable but has less geographic resolution.

Clayton and Kaldor (1987) propose a Bayesian approach to this problem which defines the analytic form of the compromise estimator. Bayesian statistics in general treats all unknown model parameters as random variables, and the goal of inference is to define the distributions of these variables, thereby providing point and interval estimates, predictions, and probability calculations. Analysts derive these distributions based on a combination of prior information or beliefs regarding the variables and the observed likelihood of parameter values in light of the data. Analysts summarize prior beliefs regarding the possible values through specification of a prior distribution assigning a probability distribution without regard or reference to the data. Data inform on the variables of interest through the likelihood function (the same likelihood function used in maximum likelihood analysis). The likelihood function summarizes the conditional probability distribution of the data, given the value of the unknown parameters. The posterior distribution reflects the conditional distribution of model parameters, given the data, and represents a compromise between the prior distribution and the likelihood function, thereby summarizing the distribution of the random variable(s) of interest, taking into account both prior beliefs and the information observed in the data.

What Smoother to Use

The most important property of a smoother is its accuracy. It should correctly identify regions of high and low rates and smooth over rates that are artificially elevated due to instability. It should not indicate trends or patterns when no such trends or patterns exist. Using these criteria, Kafadar (1994) put several smoothers to the test (e.g., locally weighted average, empirical Bayes and a technique known as headbanging, applying them to carefully simulated data where the true spatial variation in the data was known. Kafadar (1994) found that the locally weighted average smoothers with weights inversely proportional to distance (distance smoothing) were the most accurate.

Edge Effects

Observations near the edges of the study area have fewer local neighbors than observations in the interior. As a result, smoothed values near the edges often average over or borrow less neighboring information than their interior counterparts. For this reason, the behavior of spatial smoothing algorithms can be suspect near the edges of the study area. Accurate assessment of the impact of edge effects and development of flexible adjustments remain open areas for further research, for which Gelman and Price (1999) and Lawson et al. (1999) provide recent developments.

3.2.7 Exploratory Spatial Data Analysis

Tukey (1977) provided the first set of tools for data exploration, now known comprehensively by the term exploratory data analysis (EDA). EDA is a set of statistical techniques, many of them graphical that allow us to look at our data in various ways. The goals of EDA include extracting important variables; identifying outliers, anomalies, and unusual observations; detecting patterns and trends; and refining scientific hypotheses. The emphasis is on descriptive rather than inferential statistical methods, with familiar techniques: five-number summaries, stem-and-leaf diagrams, box plots, scatterplots, smoothing, and the more visual and interactive ideas of brushing, rotating, and spinning (Cleveland, 1985, Bailey et al., 1995).

When analyzing spatial data EDA extends to ESDA: exploratory spatial data analysis. All EDA techniques may be used with spatial attribute data, but we need additional techniques that explicitly use the spatial arrangement of the observations. For example Haining (1990) provides ESDA techniques that can be used to detect spatial outliers, observations that may not be unusual overall but that are unusual with respect to neighbouring values. Median polish and many other ESDA methods are described and illustrated in Cressie and Read (1989) and Cressie (1993). Much recent attention has been focused on software development for ESDA, Wise et al. (1999) provide a more recent overview of scientific visualization and ESDA.

GeoDa 0.95i uses quantile, standard deviation and a box map E(S)DA methods. Quantile map provides a method of highlighting extreme values on a map, with visualization of gradual color range. Quantile map is a variety of choropleth map; probably the most common map method type for displaying areal count data. These maps use different color and pattern combinations to depict different values of the attribute variable associated with each area. Many cartographers find the choropleth map a relatively crude method of displaying data, particularly data such as disease rates or exposure values, which vary continuously in space. A similar method of displaying extreme values on map provides a standard deviation map.

A box map provides an enhanced version of a quartile map, in which the outliers in the first and fourth quartile are highlighted separately. The classification in the box map is thus identical to that used in a box plot. To confirm the classification of the outliers, we bring up a regular box plot for the same variable, using the default hinge of 1.5, and select the upper outliers. The selected points (outliers) in the box plot correspond exactly to the dark red locations in the box map. In addition to showing the high values, the map also suggests that these may be clustered in space, something which the standard box plot is unable to do.

3.3 Methods for assessment of regional spatial data

The typical data structure for assessment of spatial health patterns involves a collection of locations of incidence events over a period for a given study area. Common questions relating to the clustering of health events include:

- Do cases tend to occur near other cases (perhaps suggesting an infectious agent)?
- Does a particular area within the study region seem to contain a significant excess of observed cases (perhaps suggesting an environmental risk factor)?
- Where are the most unusual collections of cases (the most likely clusters)?

To address such questions, we need to determine whether an observed collection of cases/rates is somehow unusual (i.e., different than we expect under a hypothesis of chance allocation of cases among the population at risk). The question "Are disease cases/rates clustered?" appears to imply existence of a simple "Yes" or "No" answer and suggests a hypothesis-testing approach based on a conceptual null hypothesis of H_0 (There are no clusters of cases). The common method for operationalizing the null hypothesis for regional count data is the constant risk hypothesis, which assumes cases reflect a random sample of the at-risk population where the probability of selection is the same everywhere (perhaps varying by risk strata, e.g., age, but with spatially constant stratum-specific risks).

Detecting a cluster involves the identification of a collection of cases inconsistent with our null hypothesis of no clustering, whereas detecting clustering involves assessment of the overall propensity of cases to cluster together (i.e., detecting the tendency of cases to cluster rather than identifying a particular collection or collections of cases). A single cluster represents an anomaly in the data (i.e., a collection inconsistent with the broader pattern), when clustering represents a pattern among all or most cases. Typically, a test of clustering provides a single assessment of the statistical significance of the pattern for the entire area (e.g., a single p-value) whereas tests to detect clusters often utilize multiple tests (multiple p-values) to determine which collection of cases represents the most significant cluster. We note that the distinction between detecting clusters and detecting clustering is not always clear (e.g., some cases may occur in clusters whereas others do not, there may be more than one cluster, or only a subset of cases actually exhibit a clustering tendency) (Elliot et al., 2001; Waller and Gotway, 2004). Nonetheless, the categories focus our attention on different aspects of the observed spatial distribution of cases and different answers to the question "Is there evidence of clusters in data?"
Besag and Newell (1991) further distinguish between general tests and focused tests (of clustering or to detect clusters). General tests test for clusters/clustering anywhere in the study area; focused tests test for clusters or clustering around predefined foci of suspected increased risk of disease (e.g., a contaminated well). Statistically, the difference lies again in

the null hypotheses; contrast Ho for a general test (absence of clustering) with that for a focused test H_0 There are no clusters of cases around the foci. In a focused study, one attempts to increase statistical attention (e.g., power) on a particular set of possible clusters (those around the foci) rather than trying to find any sort of cluster in any location.

Null Hypotheses (regional counts)

In particular, many methods model the regional counts as independent Poisson random variables based on one of the basic properties of a spatial Poisson process: event counts from nonoverlapping regions follow independent Poisson distributions where the underlying intensity function defines the expected values (and variances). Some analysts prefer the binomial distribution as a probabilistic model for regional counts, since the Poisson distribution assigns a nonzero positive probability of observing more cases than persons at risk in each region. However, for rare diseases, this probability is very small, so the practical difference between the Poisson and binomial distributions is often negligible.

In some instances we may want to condition on the total number of cases in the study area, yielding a multinomial distribution for the set of counts. Since we consider the population sizes n_1, n_2, ... , n_N fixed, the total population size in the study area,

$$n_+ = \sum_{i=1}^{N} n_i \qquad (2)$$

is also fixed. If we condition on Y_+, the total number of cases observed, we are in effect setting the individual disease risk $r = Y_+/n_+$, which is assumed fixed for everyone under the constant risk null hypothesis.

$$r = \frac{Y_+}{n_+} \qquad (3)$$

For the most part, methods to assess clusters and clustering in count data assume some background information on the entire population rather than a set of controls. This is due primarily to the wide availability of regional census data. For many countries, the census provides readily available data giving detailed background information on the population at risk within a specific set of defined regions. Since regional count methods for assessing clustering or detecting clusters typically use observed incident disease counts and census-based population counts for the same set of regions, most assess the constant risk hypothesis (people are equally likely to contract the disease regardless of location) by comparing the counts observed to their corresponding counts expected based on a global incidence rate (proportion) applied to local population counts. (The term global rate refers to the rate observed across the entire data set, and local population; Count refers to the population size observed within a single enumeration district.) Since the comparison group (census data) typically includes all persons at risk (including the cases), application of the random labeling hypothesis (fixing the case and control locations and randomly assigning the case and control labels) is less straightforward for regional count data than for point data. As noted above, the constant risk hypothesis assumes a constant disease risk, r, giving $E_i = rn_i$ for i= 1,2, ... , N.

Under a heterogeneous Poisson process, our observed regional counts should appear (1) independent, (2) Poisson distributed, with (3) expectations (and variances) E_i for i = 1, ... , N. These three components represent areas for potential deviation from a heterogeneous Poisson process, and these (individually and in combination) represent alternative hypotheses for many of the statistical tests of disease clustering and clusters proposed for regional count data. Although only the third component relates directly to the constant risk hypothesis per se, the literature contains many different statistical assessments of clustering and clusters, each building on different deviations from the components noted above.

3.3.1 Categorization of methods (regional data)

When considering tests of disease clusters and/or clustering for regional count data in categories, we base them on the statistical strategy for detecting clusters and clustering:

- Methods based on scanning local rates (with an aim of identifying the areas with high or unusually low local incidence proportions -rates);
- Methods based on global indexes of spatial autocorrelation (described below in the following chapter);
- Methods based on local indexes of spatial autocorrelation (see below);
- Methods based on goodness-of-fit tests (aiming at summarizing the deviation between observed data and their expected values under specific probabilistic model).

In our study we focused primary on tests particularly addressing deviations from a heterogeneous Poisson process based primarily on either spatial correlation or lack of correspondence between observed and (constant risk) expected values.

3.3.1.1 Global indexes of spatial autocorrelation

Griffith (1992) describes several interpretations of spatial autocorrelation, including the notion of self-correlation. Under this interpretation, the term spatial autocorrelation implies correlation among the same type of measurement taken at different locations. A global index of spatial autocorrelation provides a summary over the entire study area of the level of spatial similarity observed among neighbouring observations. Statistical indexes of autocorrelation appear in a wide variety of applications (Anselin et al., 1995, 1996).

The goal of a global index of spatial autocorrelation is to summarize the degree to which similar observations tend to occur near each other. Typically, extreme values of the index in one direction suggest positive spatial autocorrelation, while extreme values in the opposite direction suggest negative spatial autocorrelation. Since global indexes are by definition summaries over the study area, most applications of global indexes of spatial correlation in the assessment of disease patterns result in tests of clustering rather than tests to detect individual clusters. Autocorrelation among disease counts or incidence proportions may reflect real association between cases due to infection, or perceived association based on a spatial aggregation of similar values.

Most indexes of autocorrelation share a common basic structure; we calculate the similarity of values at locations i and j, then weight this similarity by the proximity of locations i and j. High similarities with high weight (i.e., similar values close together) lead to high values of the index, while low similarities with high weight (i.e., dissimilar values close together) lead to low values of the index. Following Lee and Wong (2001), let sim_{ij} denote the similarity between data values Y; and Y_j, and let w_{ij} denote a weight describing the proximity between locations i and j, for i and j = 1, ... N. Most global indexes of spatial autocorrelation build on the basic form:

$$I = (\frac{1}{s^2})\frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij} sim_{ij}}{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}} \qquad sim_{ij} = (Y_i - \overline{Y})(Y_j - \overline{Y}) \qquad s^2 = \frac{1}{N}\sum_{i=1}^{N}(Y_j - \overline{Y})^2 \qquad (4)$$

the weighted similarity between observations. The statistical literature often refers to statistic of this form as general cross-product statistics. Various indexes adjust the average similarity by different multiplicative constants to rescale or otherwise normalize the index, but the basic structure remains the same.

As already noted null hypothesis of no clustering (or no clusters), operationalized through the constant risk hypothesis, and typically assumes Poisson-distributed regional counts with heterogeneous expected values (due to varying population sizes across regions) in addition to the assumed independence between regional counts. The discrete nature of the data and the relatively small counts expected violate any assumption of normality, and heterogeneous population sizes violate any assumption of constant variance among incidence proportions (even under the constant risk hypothesis).

As a result, the normality assumption is inappropriate for most public health (and many geographical) applications. In addition, variation in the expectations and variances of counts (and in the variances of incidence proportions) renders the randomization assumption inappropriate (Besag and Newell, 1991). Since the standard approximations to the null distribution of global indices of spatial autocorrelation may be inappropriate for assessing clustering in heterogeneously distributed Poisson data, we turn to Monte Carlo hypothesis tests (Besag and Newell, 1991). We simply calculate the value of the autocorrelation index for each of a number of data sets simulated under the constant risk hypothesis and distributional assumptions more appropriate to our application, and compare the index observed to the distribution defined by the simulated values.

Moran I

Moran's I (Moran, 1950) is widely used, and variations of relate to likelihood ratio tests and best invariant tests for particular models of correlation for normally distributed random variables (Haining, 1990; Tiefelsdorf, 2000). Moran's I follows the basic form (equation 4) for global indexes of spatial autocorrelation with similarity between regions i and j defined as the product of the respective difference between Y_i and Y_j with the overall mean.

$$sim_{ij} = (Y_i - \overline{Y})(Y_j - \overline{Y}) \qquad \overline{Y} = \sum_{i=1}^{N} Y_i / N \qquad (5)$$

Thus, I is a random variable having a distribution defined by the distributions of and interactions between the Y_i. We obtain the value of I observed by inserting observations into equation. When neighbouring regions tend to have similar values (i.e., the pattern is clustered), I will be positive. If neighbouring regions tend to have different values (i.e., the pattern is regular), I will be negative. When there is no correlation between neighbouring values, the expected value of I is E(I)=-1/(N-1) (approaching 0 as N increases).

When we replace regional disease counts with regional crude incidence proportions (rates), we are seeking to remove or at least lessen the impact of population heterogeneity. Consideration of incidence proportions (rates) removes heterogeneity in the value expected under the constant risk hypothesis since comparison of regional incidence proportions to the overall mean incidence proportion makes more sense than comparison of regional counts to the overall mean count. However, the variances of the local incidence proportions depend on the regional at-risk population sizes, which remain heterogeneous. Oden (1995) and Assuncao and Reis (1999) each propose adaptations of Moran's I for incidence proportions and provide derivations of the associated null distribution in the presence of heterogeneous regional population sizes.

GeoDa implements the adjustment procedure of Assuncao and Reis (1999) (Anselin, 2005), which uses a variable transformation based on Empirical Bayes principle, basically comparing the Moran scatter plot for ‘raw’ rates to that for Empirical Bayes standardized rate (EB). Global spatial autocorrelation analysis of rates is handled by means of Moran’s I spatial autocorrelation statistic and its visualization in the form of a Moran Scatter Plot (Anselin 1995, 1996). The rates (on the x-axis) are standardized so that the units in the graph correspond to standard deviations. The scatter plot is also centered on the mean with the axes drawn such the four quadrants are shown. They provide a classification of four types of spatial autocorrelation: high-high (upper right), low-low (lower left), for positive spatial autocorrelation; high-low (lower right) and low-high (upper left), for negative spatial autocorrelation. The slope of the regression line is Moran's I, listed at the top of the graph (in blue). The values of Moran's I range from +1 meaning strong positive spatial autocorrelation, to 0 meaning a random pattern to -1 indicating strong negative spatial autocorrelation.

We can also interpret the Moran's I in relation to the z score. The z score for an item, indicates how far and in what direction, that item deviates from its distribution's mean, expressed in units of its distribution's standard deviation. The Moran I indicates Positive Spatial Autocorrelation if, $I > -1/(n-1)$ or $z > 0$ and proportionally Negative Spatial Autocorrelation when $I < -1/(n-1)$ or $z < 0$.
Inference for Moran's I is based on a permutation approach, in which a reference distribution is calculated for spatially random layouts with the same data (values) as observed. The randomization uses an algorithm to generate spatially random simulated data sets outlined in Anselin (1996).

Analysis of Global indexes of autocorrelation suggests evidence of clustering, but does not identify the locations of any particular clusters driving the pattern observed.

3.3.1.2 Local indexes of spatial autocorrelation

As their name implies, global indicators of spatial association assess patterns of spatial similarity summarized (often through weighted averages) over the entire study area. When such indexes indicate positive spatial autocorrelation, this autocorrelation may arise from a number of sources. If the index compares regional counts or incidence proportions to an overall mean regional count, or proportion, a spatial trend in expectation can result in index values suggesting spatial similarity. Even adjusted for trends and variance heterogeneities, global indexes remain global and only detect spatial associations averaged over the entire study area. Therefore, such indexes may have little statistical power to detect a single cluster within a study area otherwise following the null hypothesis. In addition, a global index can suggest clustering but cannot identify individual clusters. These issues led Anselin (1995), and Ord and Getis (1995) to consider local forms of the global indexes, which Anselin (1995) termed local indicators of spatial association (LISAs).

Anselin (1995) outlines the goals and structure of the class of LISAs. The main purpose of such indicators is to provide a local measure of similarity between each region's associated value (in our case, a count or an incidence proportion) and those of nearby regions. We can map each region's LISA value to provide insight into the location of regions with comparatively high or low local association with neighbouring values.

Anselin (1995) also formally links LISAs with corresponding global indicators by requiring that the LISA values from each region sum to a global indicator of spatial association (up to a multiplicative constant). This connection defines LISAs as components of a global index, and provides a means for partitioning a test of clustering (the global index) into a set of tests to detect clusters (the LISAs). As a result, most LISAs are defined as local versions of well-known global indexes. One of the most popular LISAs is a local version of Moran's I.
The typical output of a LISA analysis involves the values of the LISAs themselves, typically mapped to indicate areas with high values, suggesting stronger local correlation than others. We note that high LISA values may be due to aggregations of high counts or proportions, aggregations of low counts or proportions, or aggregations of moderate counts or proportions. As a result, high values of a LISA suggest clusters of similar (but not necessarily large) counts or proportions across several regions, and low values of a LISA suggest an outlying cluster in a single region (different from most or all of its neighbors). Other typical output includes maps of p-values associated with the probability of exceeding the observed value of each regional LISA, under a given set of assumptions that determines the distribution of the LISA under the null hypothesis. We use Monte Carlo testing to address the distribution of LISAs based on heterogeneous non-Gaussian random variables, but few issues remain, as multiple testing problem and correlation between neighbouring LISAs.

Local Moran's I

Probably the most widely used family of LISAs is the local version of Moran, defined for the *i*th region as

$$I_i = \sum_{j=1}^{N} w_{ij} sim_{ij} = (Y_i - Y) \sum_{j=1}^{N} w_{ij} (Y_j - \overline{Y}) \qquad (6)$$

Most authors (e.g., Anselin, 1995; Lee and Wong, 2001) divide each deviation from the overall mean by the overall variance of the Y_i values, yielding

$$I_{i,std} = \frac{(Y_i - \overline{Y})}{s} \sum_{j=1}^{N} w_{ij} \frac{(Y_j - \overline{Y})}{s} \qquad (7)$$

where *s* represents the square root of the sample variance of the Y_i's, and the "std" subscript represents the use of the standardized difference of each regional observation from the overall regional mean. Tiefelsdorf (2000) presents an application of $I_{i,std}$ assuming that each Y_i follows a Gaussian (normal) distribution to health data.

3.4 Methodology of research

3.4.1 Processing data on the place of residence/schooling in GIS

The established databases with prevalence's were presented in GIS (Geographic Information System) and linked to the presentation of various data groups. Afterwards the data was transferred to statistical application packages for further processing. The ESRI ArcGIS 9.2 application package requires a relational database in the dBase IV (or newer) program format (*.dbf). The database was transferred to ArcGIS 9.2 by defining the coordinates X and Y from the database.

In the geographic information system, the database (vector data) turns into a platform, which makes it possible to visually locate, i.e. to map the questionnaire participants (mapping). The data can be presented independently from the projection, on the cartographic basis (raster presentation); in this case state topographic maps (scale: DTK25 and DTK50 of the Surveying and Mapping Authority of the Republic of Slovenia) were used. In addition to this, contour lines in ESRI format (*.shp) were added: the borders of the Municipality of Koper and the borders of local communities, and the data from the Environmental survey of companies in the Municipality of Koper from 2006/07. This procedure enables visual presentation of the disease incidence variance in a two-dimensional space.

The freely accessible GeoDa 0.95i application package was used to process geospatial data and to analyze the samples and incidence distribution. The program combines mapping methods, EDA, ESDA, spatial and statistical analyses of clusters and fundamentals of spatial correlation analysis.

3.4.2 Analysis of rates (regional count data)

We started the analysis of prevalence rates with employment of Exploratory Data Analysis (EDA) included in program GeoDa. EDA methods are used for visualization of non-spatial distribution of data: histogram, scatter plot and box plot. We compared the distribution of cases per community for chronic pulmonary diseases, allergies and asthma with bronchitis, aiming at defining first order patterns of disease appearance.

The histogram is used as a discrete approximation to the density function of a random variable and is useful to detect asymmetry, multiple modes and other peculiarities in the distribution. The results are histograms with the variables (rates) classified into defined categories.

We used the EDA method called box plot. It shows the median, first and third quartile of a distribution (the 50%, 25% and 75% points in the cumulative distribution) as well as a notion of outlier. An observation is classified as an outlier when it lays more than a given multiple of the interquartile range (the difference in value between the 75% and 25% observation) above or below respectively the value for the 75th percentile and 25th percentile.

We used Exploratory Spatial Data Analysis (ESDA) dealing with the exploration of spatial aspects. Our focus was primarily on map making and geovisualization, (included in GeoDa) and comparison of distribution of events per communities for chronic pulmonary diseases, allergies and asthma with bronchitis with ESDA methods: quantile, standard deviation and box map (Anselin, 2005). We used standard deviation map and box map to highlight the extreme values on a map aiming at defining spatial outliers.
The prevalence rate of chronic pulmonary diseases was used to assess the risk for getting the disease in specific geographical region. However, as a map of rates may still obscure the spatial pattern in disease risk therefore we used the method called spatial smoothing for reducing noise in rates associated with geographic regions (by using GeoDa). We employed different smoothers: empirical Bayes smoothing, contiguity smoothing (rook based) and distance based nearest neighbor smoothing.

As an additional exploratory tool, we constructed a probability map for prevalence rates depicting the probability of observing a rate as extreme or more extreme than our original rate (under the null hypothesis that all rates are equal to a mean rate). Those methods only pose a first step to clarifying the signal to control the noise and do not addresses the fundamental issue involved – is there an evidence of clustering, clusters. The analysis was based on the results of survey of population included in the study and then on extrapolation of those data to the population of all newborns between 1991 and 1996.

3.4.3 Assessment of regional spatial data

Analysis of clustering

We used the global assessment of spatial autocorrelation as a test of clustering. With global indexes of spatial autocorrelation we summarized the degree to which similar observations or rates tend to occur near each other. Typically, extreme values of the index in one direction suggest positive spatial autocorrelation, while extreme values in the opposite direction suggest negative spatial autocorrelation. Global spatial autocorrelation analysis of rates is handled in GeoDa by means of Moran's I spatial autocorrelation statistic and its visualization in the form of a Moran Scatter Plot (Anselin 1995, 1996). We used the Moran Scatter Plot, basically comparing the Moran scatter plot for 'raw' rates to the Moran plot for weighted standardized rate. We also interpret the Moran's I in relation to the z score, indicating how far and in what direction the item deviates from its distribution's mean, expressed in units of its distribution's standard deviation.

Inference for Moran's I is based on a permutation approach, in which a reference distribution is calculated for spatially random layouts with the same data (values) as observed (Anselin, 1996). The result is shown as a statistics with a histogram for the reference distribution and with the Moran's I.

Analysis of clusters

We then moved from the global clustering to tests for detection of single cluster or clusters. Local spatial autocorrelation analysis (LISA) is based on the Local Moran LISA statistics (Anselin, 1995), which yields a measure of spatial autocorrelation for each individual location. We assessed the evidence of local clustering in GeoDa by using Moran significance map, Moran scatter plot, LISA's box plot and LISA cluster map.

The LISA significance map shows the locations with significant local Moran statistics in different shades with the corresponding p values after a number of runs for 9999 permutations to avoid too great a sensitivity on the particular randomization and thereby obtaining robust results. The LISA cluster map provides essentially the same information as the significance map, but with the significant locations color coded by type of spatial autocorrelation (9999 permutations); four categories correspond to the four quadrants in the Moran scatter plot.

The high-high and low-low locations define positive local spatial autocorrelation and are typically referred to as spatial clusters, while the high-low and low-high locations (negative local spatial autocorrelation) are termed spatial outliers. The cluster is classified as such when the value at a location (either high or low) is more similar to its neighbours (as summarized by the weighted average of the neighbouring values, the spatial lag) than would be the case under spatial randomness. While outliers are single locations by definition, this is not the case for clusters; spatial clusters shown on the LISA cluster map only refer to the core of the cluster.

<u>LISA with adjustment for rates</u>

When using the Moran's I statistic which has been computed for rates or proportions, the underlying assumption of stationarity may be violated by the intrinsic variance instability of rates. The latter follows when the population at risk (the base) varies considerably across observations.

The variance instability may lead to spurious inference for Moran's I, therefore we implemented the Empirical Bayes (EB) standardization suggested by Assuncao and Reis (1999) to correct for this. It provides a direct standardization of the variable, using a similar (but not identical) rationale. We had run all four different types of graphs and LISA maps for CRD rates with correction for rates: a significance map, a cluster map, a box plot and a Moran scatter plot.

3.4.4 Geographic analysis of air pollution sources

We wanted to analyze the impact of potential pollution sources in administrative areas (local communities) of municipality Koper. In the absence of relevant emission/immission values against which we may analyze the risk in observed communities, and lack of data from fixed (air) monitoring sites, we based our geographic analysis on expert opinions.

Maps of polluters were formed in GIS, namely at the levels of point-pollution sources (enterprises, industry), linear pollution sources (traffic, roads) and the possible effects of remote transport. In the meteorological analysis we have also identified wind-exposed areas, where it is more likely to be exposed to greater concentration of harmful substances. In cooperation with local experts (expertise), and representatives of the local population, we then on basis of those maps defined the geographical areas with different levels of risk (pollution) - geographical areas with increased pollution, lower pollution and unpolluted areas.

We followed by comparing the results of geographical analysis of cases in local communities with the geographical distribution of polluted areas in same administrative units. On this basis we draw conclusions about potential pollutants that affect the increased morbidity in localized parts of the municipality Koper.

4. RESULTS

All the children born between 1991 and 1996 attending 1st to 4th year of primary school in the territory of the Municipality of Koper, i.e. 1,727 children from 12 primary schools within the Municipality of Koper, were included in this research. There were 1,043 questionnaires returned, which corresponds to 60.4% of correctly filled-in questionnaires. Unclassified were 0.6% of all questionnaires, 0.7% were returned empty. Questionnaires were unclassified if the cases age identification was not possible or residential address has been left out.

The questionnaire included details concerning the style of living and spending spare time, the health condition with regard to respiratory organs and other factors that can be associated with the incidence of chronic respiratory diseases. Similarly, details concerning birth registration in the Municipality of Koper between 1991 and 1996 were acquired from the Municipality of Koper and on this basis, it has been established that the entire population of newly born includes 2,357 children. The selected database provided the possibility of indexing by date of birth, address, local community.

The municipality of Koper consists of 23 local communities/tracts (Figure 1 and 3). The data (Chronic Respiratory diseases - CRD cases) includes 124 cases among 1043 persons at risk (survey population): 35 allergy, 50 asthma and 39 bronchitis cases (Figure 3).

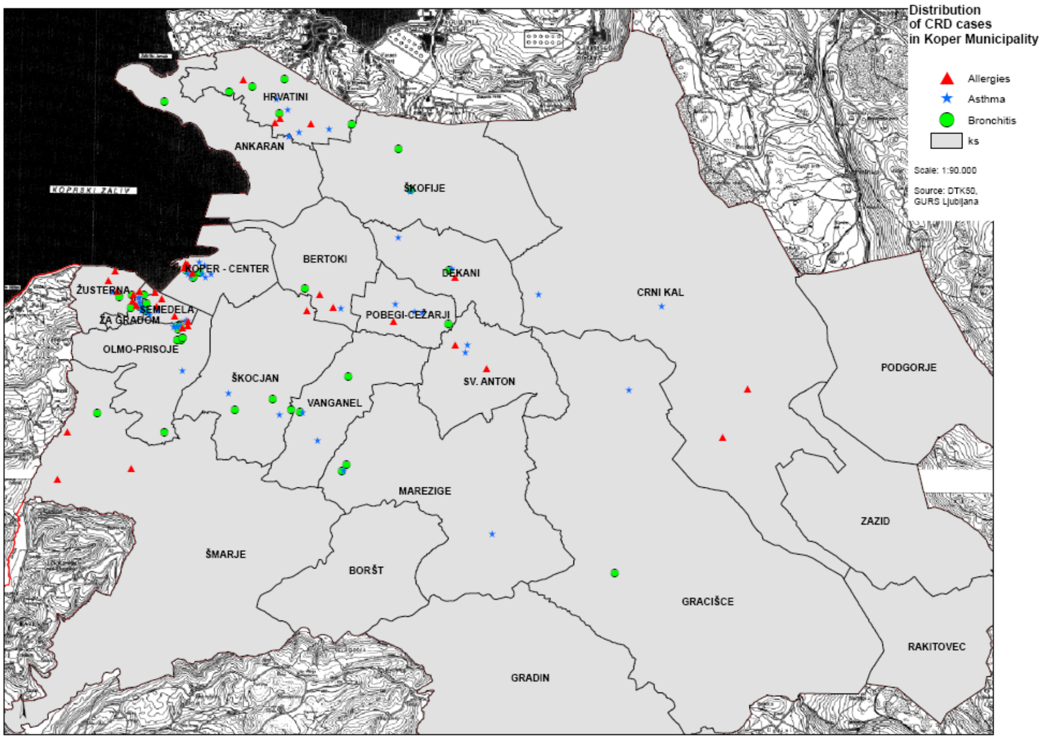


Figure 3: *Distribution map of Chronic Respiratory disease events among children included in survey in municipality Koper*

As mentioned above most of the cases were originally georeferenced to point case data and then calculated to local communities.

The examination of the geographic distribution of the permanent residence of children who reported chronic pulmonary diseases indicates that the health of children living in certain areas was more at risk than of children elsewhere. Distinctive clustering can be observed primarily in the areas of the following local communities: Pobegi-Cezarji, Hrvatini, Olmo – Prisoje, Zagrad. A relatively high rate of incidence of chronic respiratory diseases that can be observed in some local communities, e.g. Marezige, Vanganel, Dekani or Črni Kal, has to be interpreted with caution as the number of diseased children is relatively low and such distribution can be purely accidental (Figure 3).

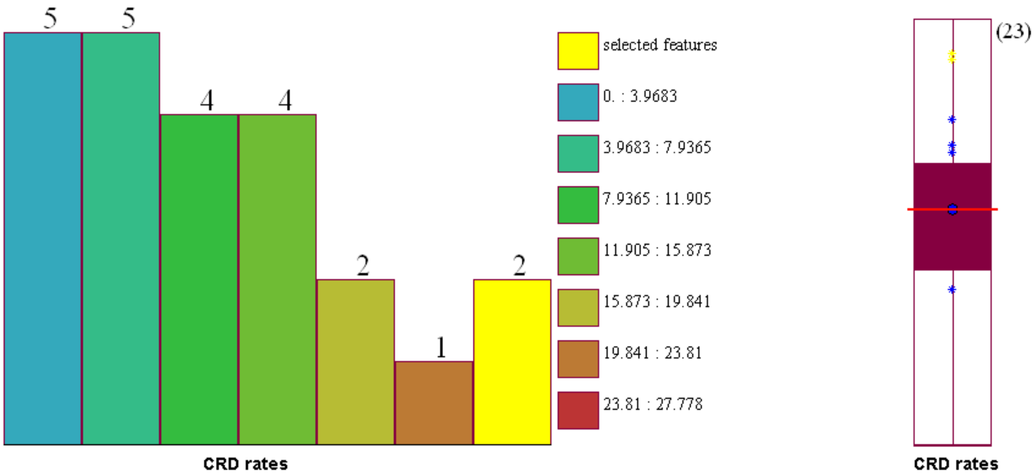
A more detailed geographic analysis of cases indicates that diseased children live in areas that are higher above sea level, i.e. in settlements on the slopes of the local hilly region. Their permanent residence is usually from 70 to 90 m above sea level.

4.1 Analysis of rates

We began the analysis of prevalence rates with employment of Exploratory Data Analysis (EDA) included in program GeoDa, used for visualization of non-spatial distribution of data. CRD rates per tract were computed by dividing the number of CRD cases (resulting from survey) by the survey population per tract (Figure 7) and then multiplied with 100 to obtain rates per 100 people per tract. The number of cases per tract ranged from 0 to 19 cases.

We started by comparing the distributions of rates for chronic pulmonary diseases, allergies and asthma with bronchitis. As results we constructed histograms with the variables classified into 7 categories (defined by user). We also used a method called box plot showing the median, first and third quartile of a distribution and expose possible outliers. Parallel to both methods we calculated the basic statistics to better define the distributions of events.

The prevalence analysis results (Figure 4) indicate that clustering of chronic respiratory disease cases is greater in the areas of some local communities in the Municipality of Koper. The data indicates that the rate of incidence of those diseases was between 0 and 27.8 in the local community of Pobegi-Cezarji, with an average rate of 10.3 cases per 100 persons. The allergic disease analysis yields similar results (Figure 5). The data indicates that the prevalence incidence range was between 0 and 8.3 in the local communities of Hrvatini and Semedela. It has been estimated that the distribution of allergic diseases is more proportionate although the trend of incidence of these diseases is, similarly, the highest in the local communities of Pobegi-Cezarji and Hrvatini.



*Legend: The red bar in the middle corresponds to the median, the dark part shows the interquartile range (going from the 25th percentile to the 75th percentile). The individual observations in the first and fourth quartile are shown as blue dots. The thin line is the hinge, corresponding to the default criterion of 1.5.

Figure 4: *Distribution of Chronic Respiratory diseases prevalence rates among children included in the study in municipality Koper (histogram with box plot map*)*

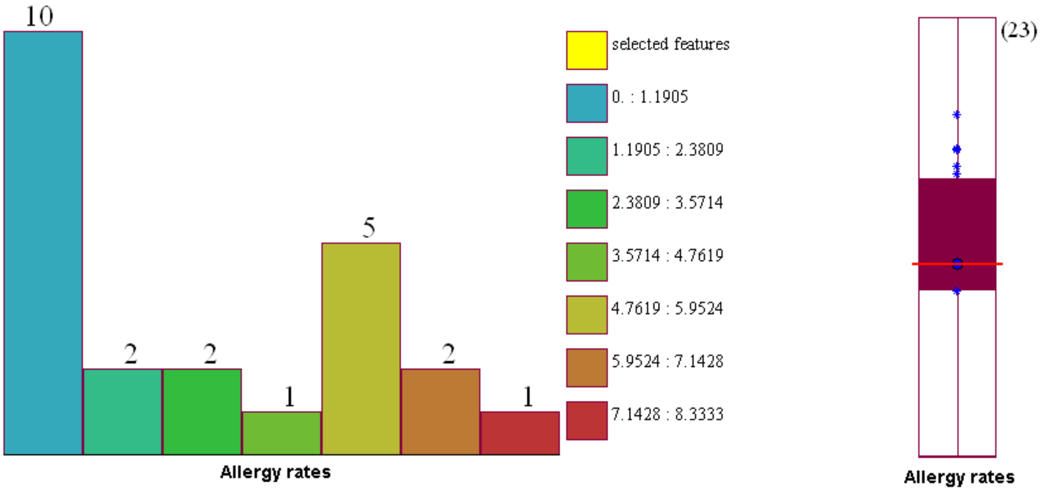


Figure 5: *Distribution of allergy prevalence rates among children included in the study in municipality Koper (histogram with box plot map)*

The asthma and bronchitis prevalence analysis results (Figure 6) indicate the rate of incidence between 0 and 22.2 cases per 100 persons in the local community of Pobegi-Cezarji. It has been estimated that the distribution of allergic diseases is more proportionate although the trend of incidence of these diseases is, similarly, the highest in the same local communities.

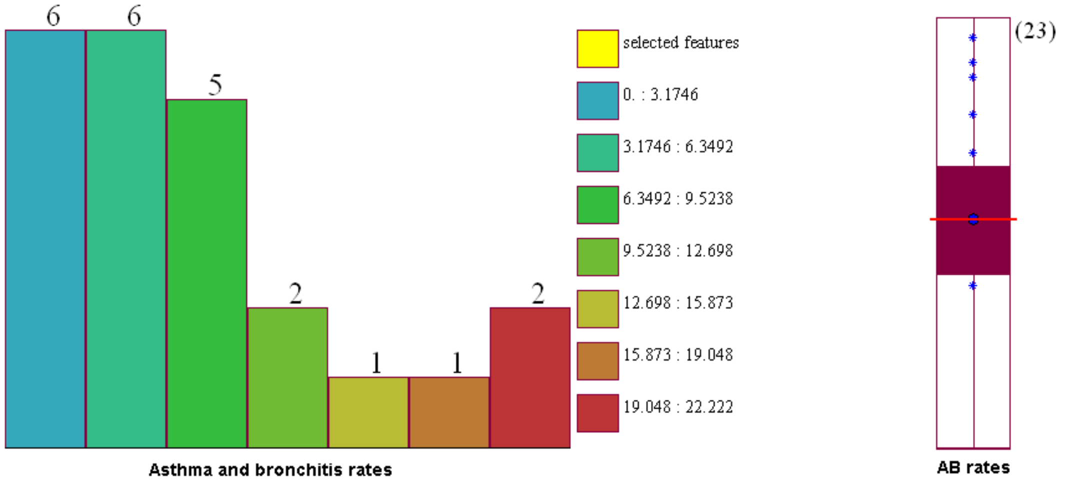


Figure 6: *Distribution of asthma and bronchitis prevalence rates among children included in the study in municipality Koper (histogram with box plot map)*

Comparing different EDA methods for: CRD events, allergy and asthma with bronchitis rates (Figures 4, 5 and 6), we may extract trends of higher prevalence rates occurring in local communities: 'Pobegi Cezarji' and 'Hrvatini'.

Analysis of rates based on population of survey

To define spatial distribution of data we bring ESDA methods into use. Primary we used ESDA methods,: choropleth and box mapping to visualize and highlight extreme values in spatial context. Although cartographers find the choropleth map a relatively crude method of displaying data, as disease rates vary continuously in space.

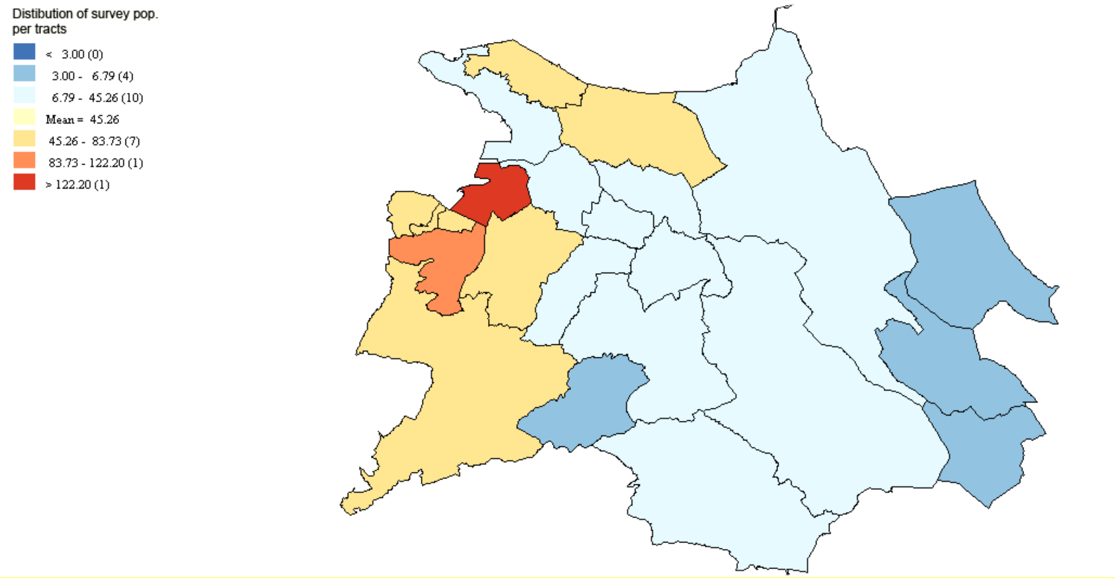


Figure 7: *Distribution of survey population (91-96) included in the study*

Regions with higher observed Chronic Respiratory Diseases (CRD) rates are of particular interest, so we will have a closer look at these (Figure 8, 9). The tracts with the highest rates per 100 people are located in 'Pobegi-Cezarji' (27.8) based on the 5 CRD events and 18 at risk and 'Hrvatini' community (27.1) based on 13 events of CRD and 113 people at risk. The average surveyed population size per tract is 45.26 people (Figure 7.

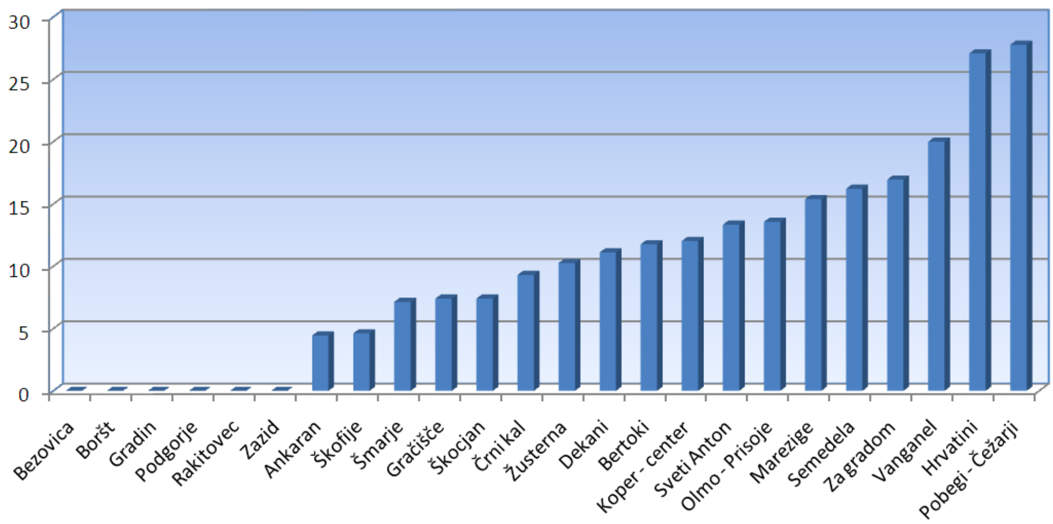


Figure 8: *Distribution of Chronic Respiratory disease prevalences per local communities in municipality of Koper*

We provide the basic statistics for the sample of CRD rates: the average value 10.3 of CRD events per 100 people at risk. Of the highest 10% of the rates, 15% of these have CRD counts 10 to 13 and one of these tracts is located within the city Koper area – 'Za gradom', the second community 'Hrvatini' is bordering Italia. There does not seem to be a trend in the rates (i.e., no apparent tendency to increase or decrease systematically in any direction), but high rates do seem to occur near the districts with higher population counts (etc. city Koper, city of Trieste in Italia).

The Figure 8 and 9 show a great deal of variability, some of which may be due to the small number of cases used to compute the rates, or to the variation in population sizes, rather than trends or patterns in the underlying CRD risk. Smoothing the rates in this situation became necessary because small variation in the number of cases may result in dramatic changes in disease rate

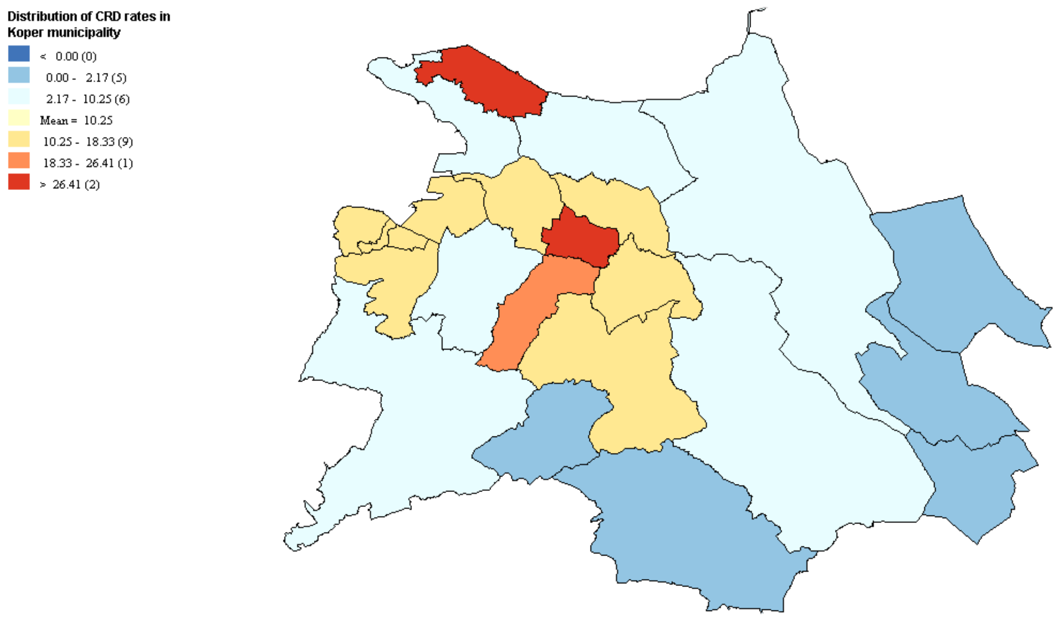


Figure 9: *Distribution map of Chronic Respiratory disease rates per communities in municipality Koper*

Therefore we first smoothed each rate using the empirical Bayes smoother. The effects of Empirical Bayes smoothing are obvious (see Figure 10): Tracts with low rates have been smoothed upward toward the mean, and tracts with high rates have been smoothed downward toward the mean. The spatial variation in the rates has been reduced considerably. The smoothed rates range in value from 10.0 to 15.0, much different than the same extrema for the original rates (ranging from 0 to 27.8 cases per 100 at risk). Even so, a few of the tracts with high original rates (e.g., those near city Koper, 'Pobegi-Cezarji' and 'Hrvatini') still appear relatively high.

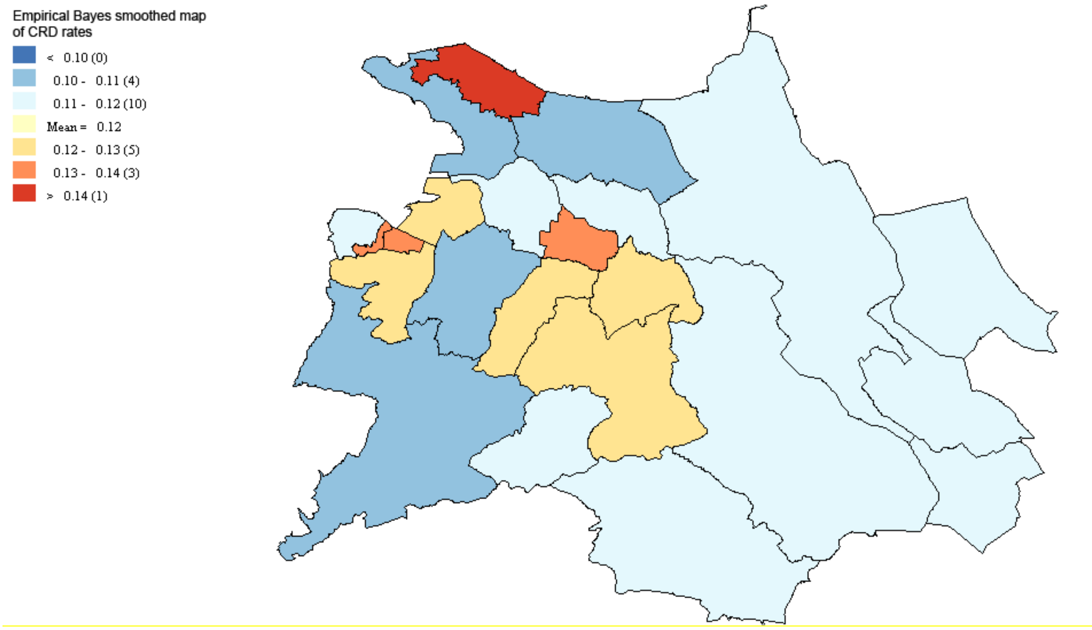


Figure 10: *Empirical Bayes smoothed map of Chronic Respiratory disease rates per communities in municipality Koper*

Next we used simple contiguity based spatial weights, where the definition of neighbor is based on sharing common boundary. The smoothed rates are computed using adjacency weights in local neighborhoods.

We used the contiguity weights and implement the rook smoothing. The resulting map appears in Figure 11. There is more variability in these smoothed rates, with values ranging from 7.1 to 15.2. Comparing this map to the map of the original rates and to the map from empirical Bayes smoothing, we can see the effects of the local smoothing. Local smoothing allows more variability in areas where adjacent tracts have moderately different rates. The same trend of tracts based in the vicinity of city Koper still appears of having higher than average rates.

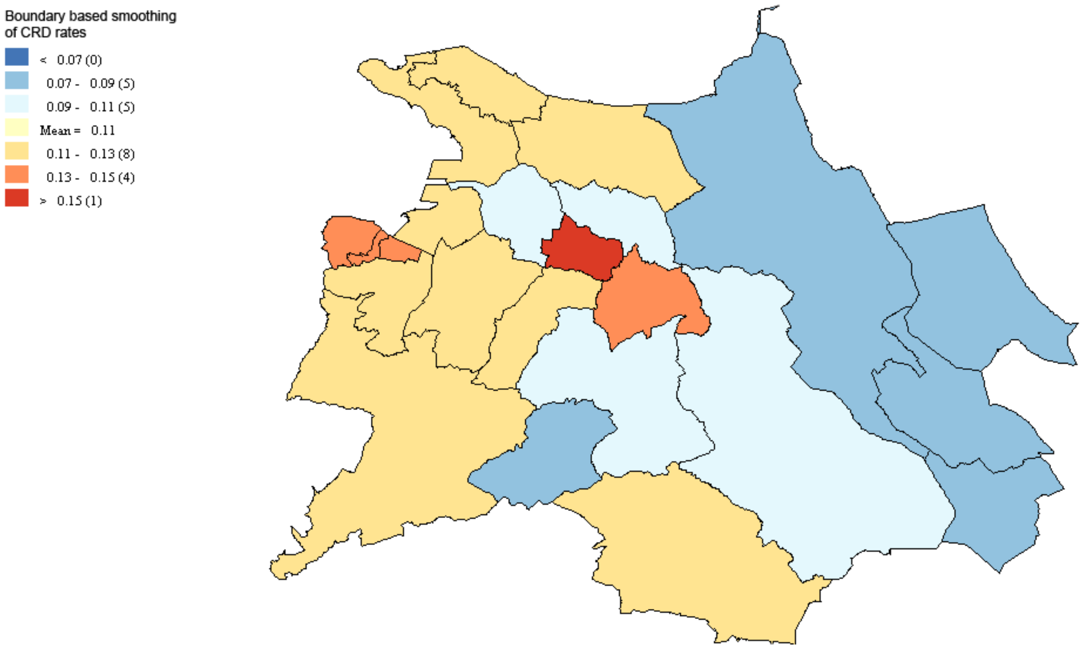


Figure 11: *Boundary based smoothed map of Chronic Respiratory disease rates in municipality Koper*

The k-nearest neighbor weights criterion (distance based smoothing) ensures that each observation has exactly the same number (k) of neighbors (Figure 12). We have used 4 nearest neighbors smoothing and still the same trend occurs.

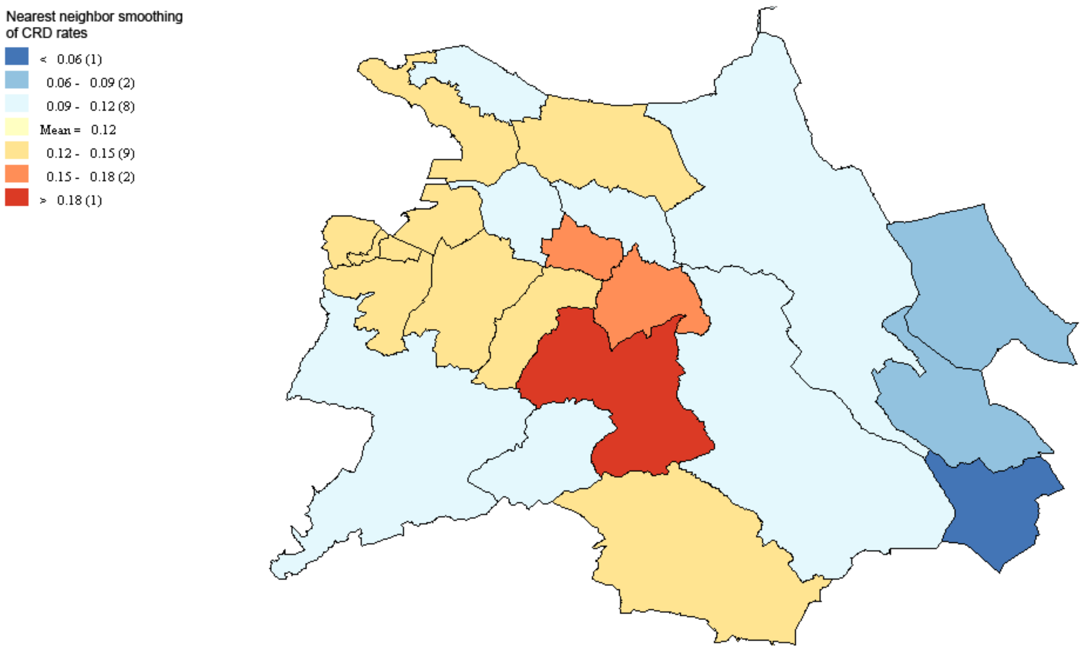


Figure 12: *Nearest neighbor smoothed map of Chronic Respiratory disease rates in municipality Koper*

Even after smoothing, a few tracts appear to have relatively high rates. This will always be the case in this type of analysis: by definition, 5% of the rates will exceed the 95th percentile of the distribution of the rate values.

Yet we have to wonder if these rates are unusually high in some sense (e.g., when compared to an overall background rate). We investigate this hypothesis more precisely later on, but as an additional exploratory tool, we construct a probability map (Figure 13). This map depicts, for each tract, the probability of observing a rate as extreme or more extreme than our original rate, under the null hypothesis that all rates are equal to a mean rate. If a probability is small, there is some evidence against this null hypothesis and we conclude that the corresponding tract has an unusually high or low rate.

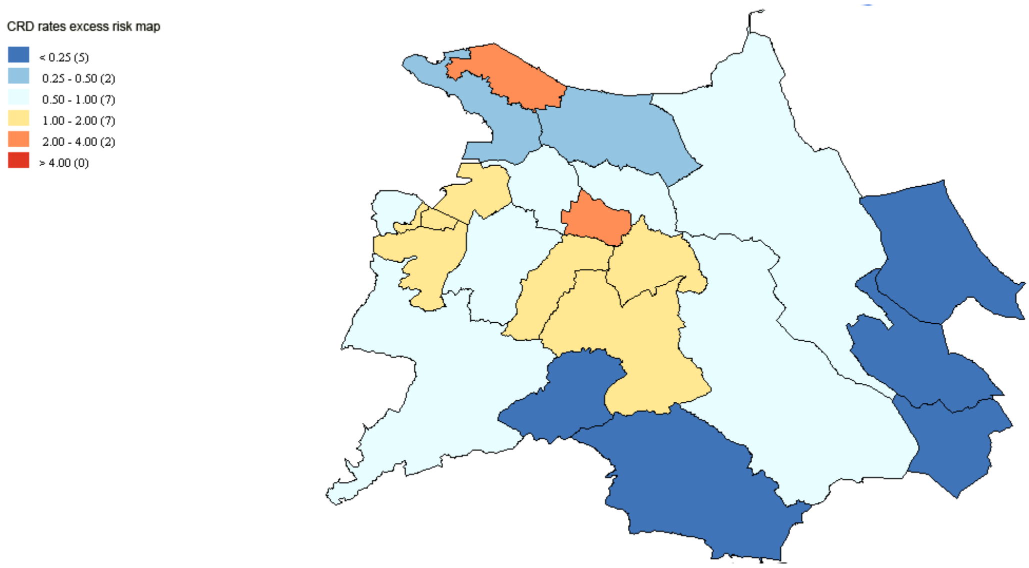


Figure 13: *Distribution map of excess risk for Chronic Respiratory disease rates per communities in municipality Koper*

Although these methods are exploratory, some areas are consistently indicated as areas of concern; communities: 'Hrvatini', ‘Semedela’ and area around ‘Pobegi Cezarji’ (‘Sveti Anton’, ‘Vanganel’). We consider additional tests and hypotheses concerning these CRD rates. In comparison to total CRD rates we introduce distribution maps for asthma with bronchitis and allergy map - figures 14 and 15.

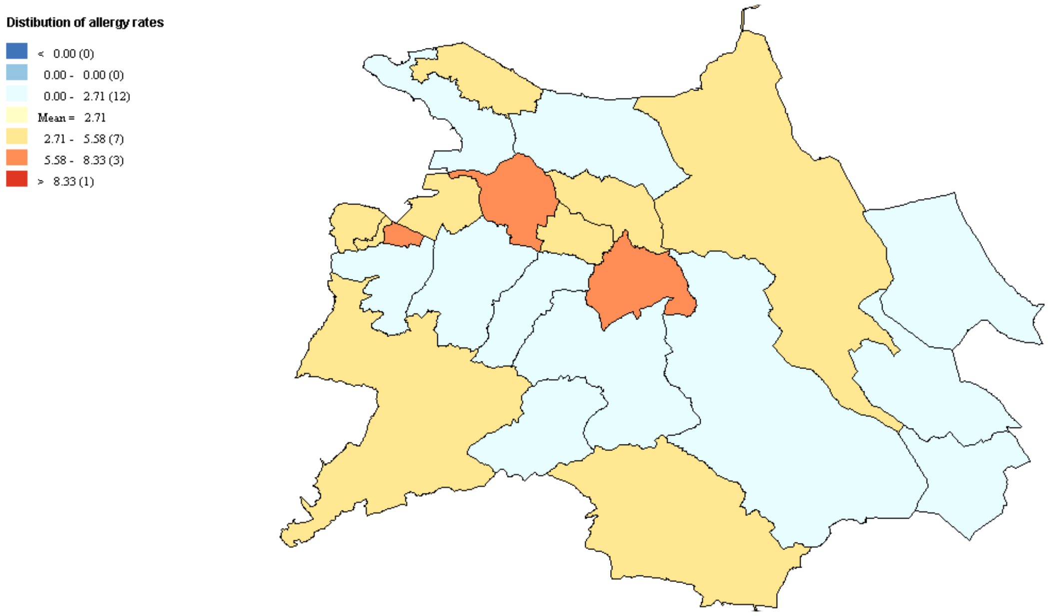


Figure 14: *Distribution map of allergies rates per communities in municipality Koper*

The results indicate that clustering or the rate of allergic respiratory diseases in children aged 7 to 11 attending 1st to 4th year of primary school is higher in some areas of the Municipality of Koper. The data reveals that the morbidity rate ranged from 0 to 8.33 cases (in the local community of Hrvatini) per 100 children, with the average rate of 2.71 (Figure 19). While for the area of Sveti Anton it can be assumed that the high morbidity rate was purely accidental as in total only two cases of disease were recorded, this can no longer be assumed for the areas of Hrvatini and Semedela.

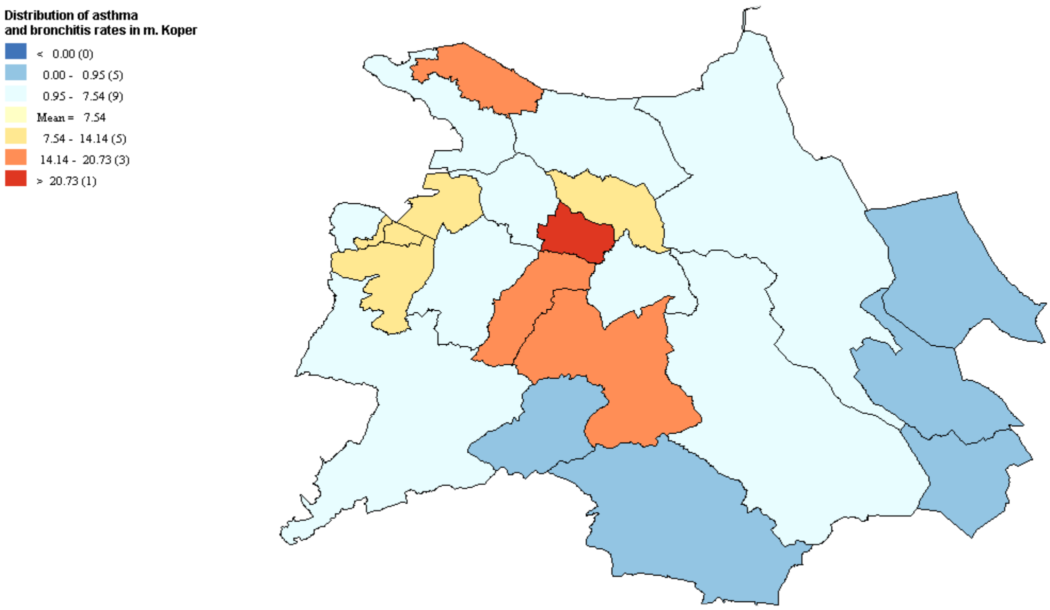


Figure 15: *Distribution map of asthma and bronchitis rates per communities in municipality Koper*

As it is often very difficult to distinguish between asthma and chronic bronchitis in children, i.e. often it is the case of the same disease – asthma, the distribution of these two chronic diseases in individual smaller areas has been checked (Figure 15). A more detailed examination of clustering, i.e. increased morbidity due to asthma and chronic bronchitis shows that there is increased morbidity in both cases in the same areas.

Analysis of rates based on population of all newborns between 1991-1996

To adjust the rates for the population of all newborns (born 1991-1996) with cases studied/observed, we did the same analysis as above for the comparison with those included in survey (the results are less stable but provide a possible output).The tract with the highest rates per 1.000 people is located in community 'Hrvatini' and is based on 13 cases of CRD on 113 persons at risk. The average population size per tract is 102.48 people at risk. The map (Figure 16) shows a great deal of variability, some of which possibly due to the smaller number of cases used to compute the rates, or to the variation in population sizes, rather than trends in the underlying CRD risk. But basically we came to the same conclusions.

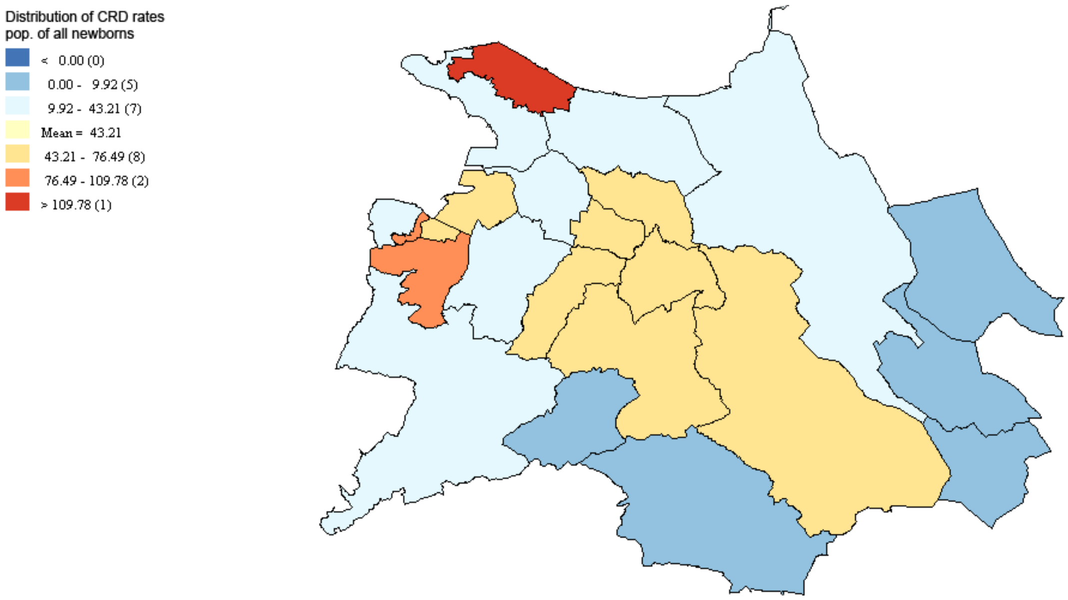


Figure 16: *Distribution map of Chronic Respiratory disease rates per communities in municipality of Koper (based on population of all newborns.)*

When comparing analysis of rates based on whole population with the ones based on survey population, we deduce the same areas consistently indicated as areas of concern, namely the communites 'Hrvatini', ‘Semedela', 'Za gradom' and the area around ‘Pobegi Cezarji’. We continue with the analyses with rates based on the survey population.

Assessment of regional spatial data

We employed indexes of spatial autocorrelation to define the degree of the adjacency between nearby regions rates. We used the global assessment of spatial autocorrelation as a test of clustering, in GeoDa handled by means of Moran's I spatial autocorrelation statistic. We used simple boundary based weights showing a good connectivity. The slope of the regression line corresponds to Moran I value of 0.2254 (Figure 17).

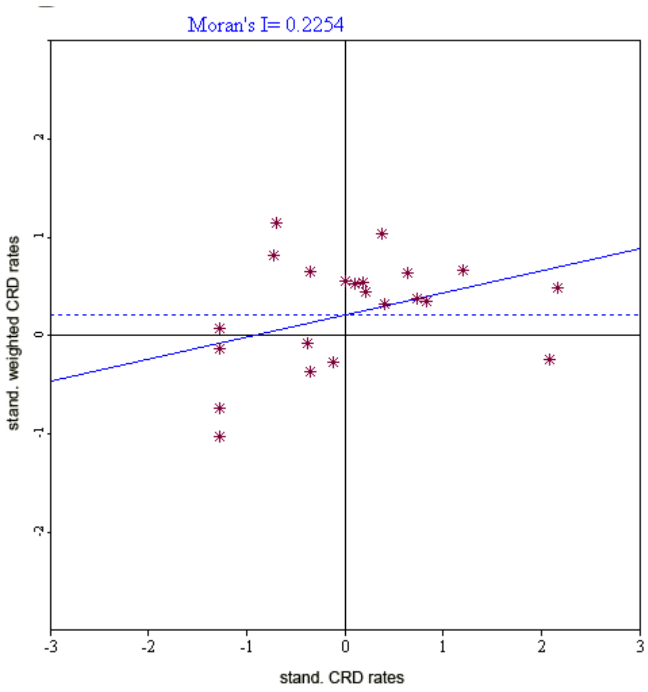


Figure 17: *Moran scatter plot for Chronic Respiratory disease rates*

The statistical significance for Moran's I is based on a permutation approach, in which a reference distribution is calculated for spatially random layouts with the same data (values) as observed. The result is a window depicting a histogram for the reference distribution, with the observed Moran's I (Figure 18).

From the summary statistics with the 9999 permutations used, the pseudosignificance level of 0.0211 is been calculated. The value I 0.2254 (Moran's I) indicates positive spatial autocorrelation; possible clustering. We also get the theoretical mean E[I]:- 0.0455 (E[I]=-1/(N-1)) and the average of the empirical distributions -0.0462 with the standard deviation (for the distribution, S) 0.1176. We calculate the z score value 2.3035 where z=(I-E[I])/√Var(I). When comparing The Moran I to z score we conclude that both values are higher then I > -1/(n-1) or z > 0 and thereby indicate the Positive Spatial Autocorrelation.

Looking at the distribution we see the value **I** lying at the tail of the reference distribution (Figure 18). As the significance is quiet high (P = 0.0211), we conclude there is a significant positive spatial autocorrelation in the data set and assume clustering for the data set.

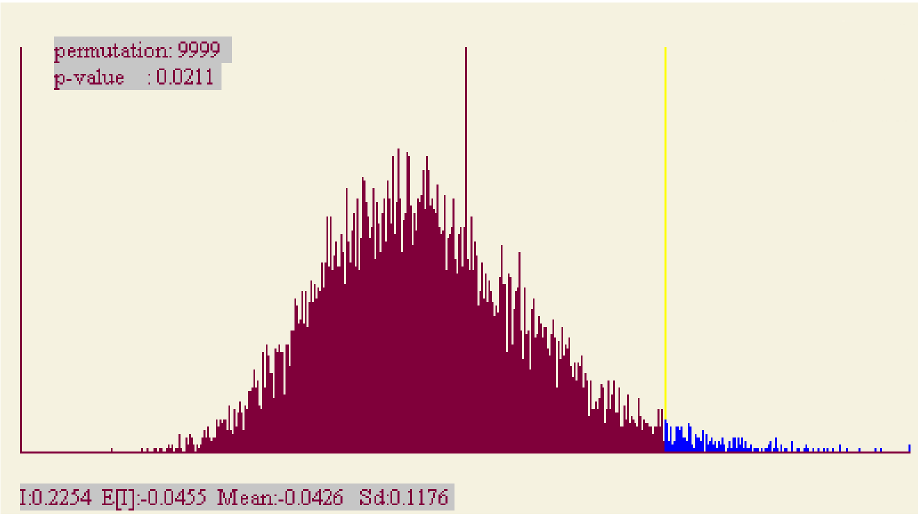


Figure 18: *Empirical distribution for Moran I after 9999 permutations (Chronic Respiratory disease rates)*

Next we used a visual method called Envelope slopes, picturing the range of autocorrelation statistics for CRD rates obtained in spatially random simulated data sets (Figure 19). Those represent, namely the slopes, the lower 5th percentile and upper 95th percentile of the reference distribution in the Moran Scatter Plot and thereby illustrating the degree of extremeness of the observed statistic.

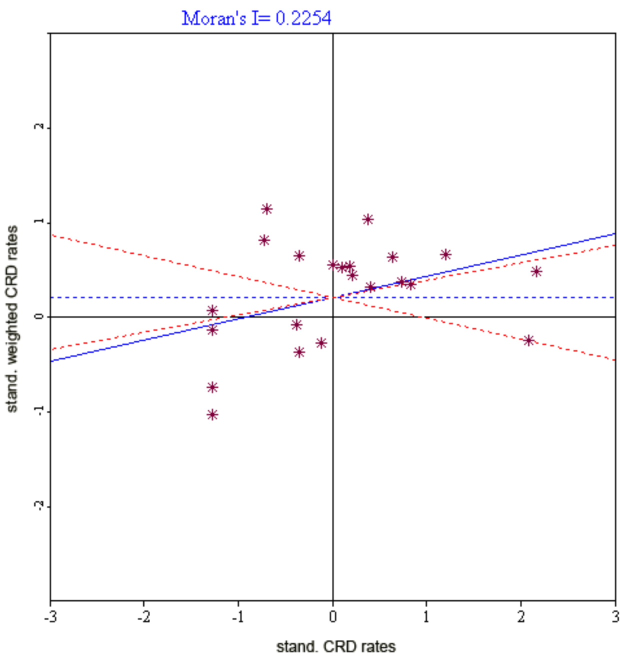


Figure 19: *Moran scatter plot of Chronic Respiratory disease rates with envelope slopes added*

Local Spatial Autocorrelation

Local spatial autocorrelation analysis (Anselin) is based in GeoDA on the Local Moran LISA statistics, which yields a measure of spatial autocorrelation for CRD rates corresponding to individual spatial location. We used four different types of graphs and LISA maps: a significance map, a cluster map, a box plot and a Moran scatter plot.

LISA Significance Map

The significance map of CRD rates, illustrated in Figure 20, shows the locations with significant local Moran statistics in different shades after a number of runs for 9999 permutations. Note the results for p > 0.05 are unreliable (ignore multiple comparisons).

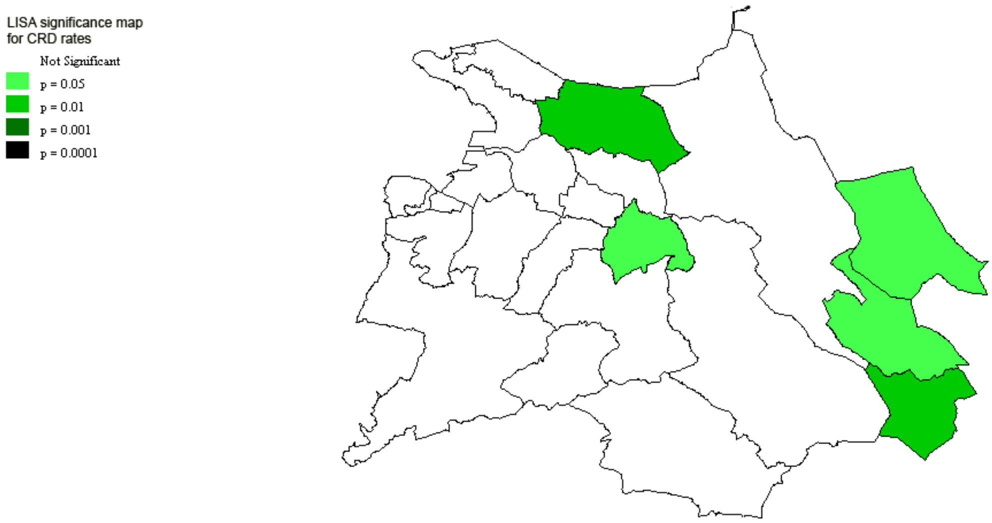


Figure 20: *Lisa significance map for municipality Koper Chronic Respiratory disease rates*

The *LISA cluster map*, shown in Figure 21, provides the same information as the significance map (two maps are linked together), but with the significant locations color coded by type of spatial autocorrelation (color categories corresponding to the four quadrants in the Moran scatter plot). When comparing LISA Significance and LISA Cluster maps, they set up the community 'Sveti Anton' as area with positive local spatial autocorrelation (high-high, P = 0.0161) and area of communities 'Rakitovec' (p=0.0039), 'Podgorje' (p=0.0276) and 'Zazid' (p=0.0275) with positive auto correlation (low-low) – as most reliable cores of the clusters. The community 'Skofije' (P = 0.0097) is set as area with negative spatial autocorrelation (Low-High) and referred as spatial outlier. The high rates cluster around the core 'Sveti Anton' most likely extends into bordering communities 'Marezige', 'Dekani', 'Pobegi-Cezarji' and 'Vanganel'.

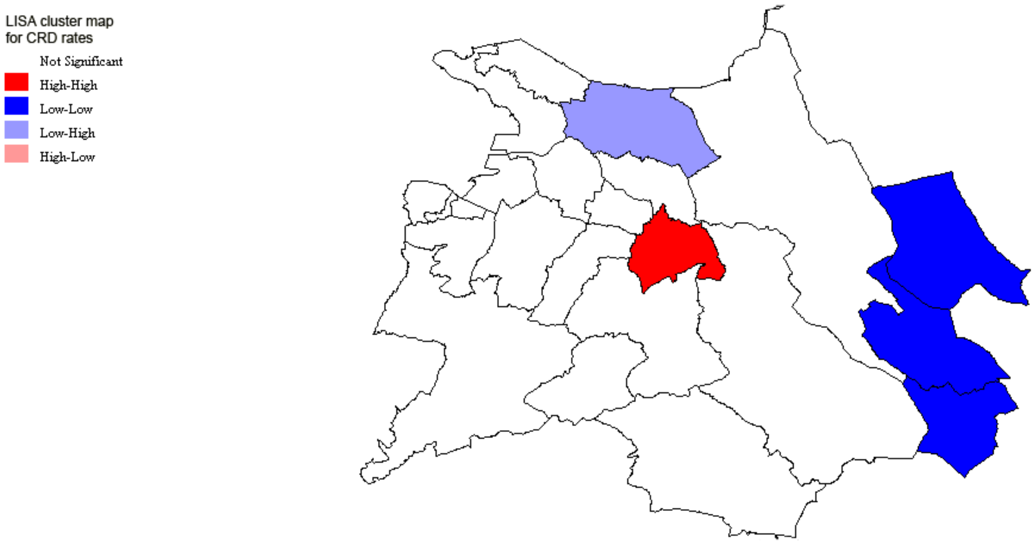


Figure 21: *Lisa cluster map for municipality Koper Chronic Respiratory disease rates*

Local Spatial autocorrelation for rates

As the variance instability may lead to spurious inference for Moran's I, GeoDa use the Empirical Bayes (EB) standardization suggested by Assuncao and Reis (1999); analysis of spatial autocorrelation for CRD rates with an adjustment for the variance instability of rates.

LISA Significance Map

The significance map, illustrated in Figure 22, shows the locations with significant local Moran statistics in different shades with the corresponding p values are given in the legend. The map shows results after a number of runs for 9999 permutations to avoid too great a sensitivity on the particular randomization. The 'Ankaran' with 'Skofije' tract are forming the most likely outlier with a pseudosignificance p= 0.0352 and p = 0.0136, with 'Sveti Anton' following as core of spatial cluster of higher rates with a pseudosignificance p = 0.0454 extending into bordering regions.

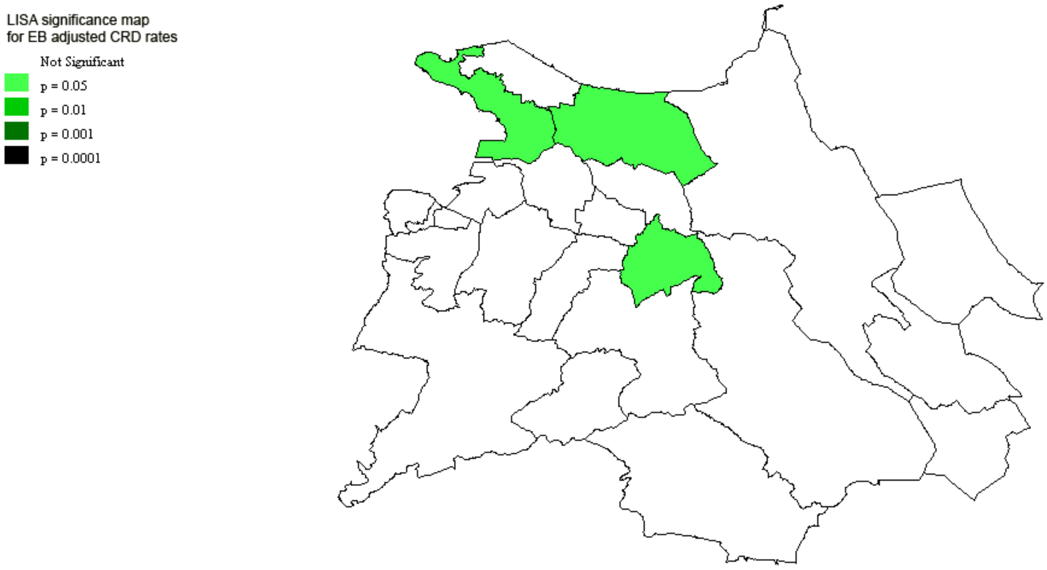


Figure 22: *Lisa significance map of Chronic Respiratory disease for municipality Koper; Empirical Bayes adjusted rates*

LISA cluster map **(**Figure 23)

When comparing LISA Significance and LISA Cluster maps, the community 'Sveti Anton' shows as area with significant positive local spatial autocorrelation (high-high) – core of the cluster. The tract 'Ankaran' and 'Škofije' has negative auto correlation (low-high) – is termed spatial outlier. The high-high locations (positive local spatial autocorrelation) are referred as spatial cluster, only define the core of the cluster.

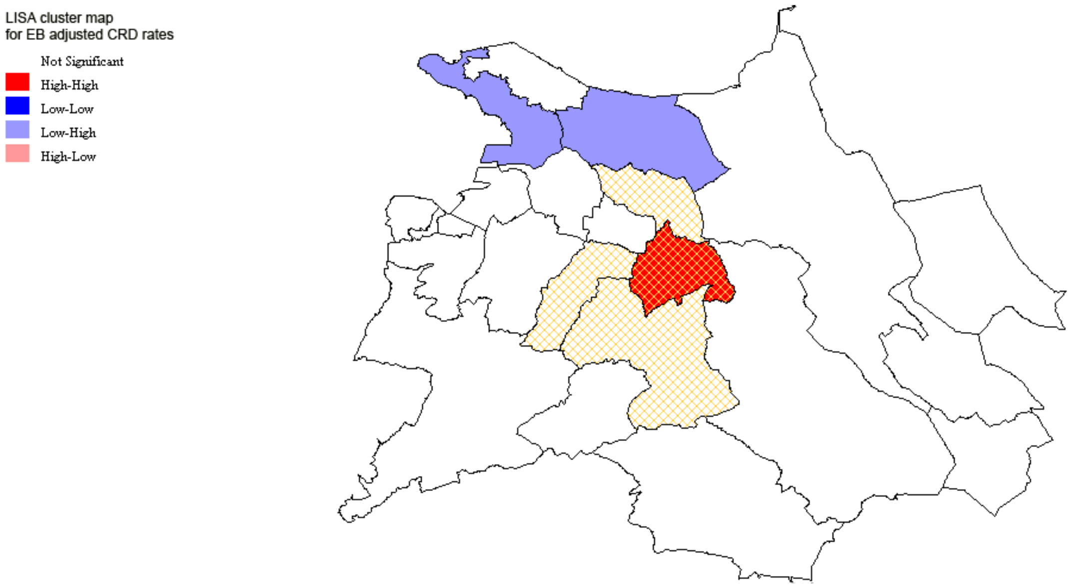


Figure 23: *Lisa cluster map of Chronic Respiratory disease for municipality Koper; Empirical Bayes adjusted rates*

Therefore, the higher rates cluster itself likely extends from the core area 'Sveti Anton' to communities 'Marezige, 'Dekani' and 'Vanganel'; outlined with yellow, Figure 28. Also Moran I value (-0.0684) differs somewhat from the statistics for the unstandardized rates (Figure 24).

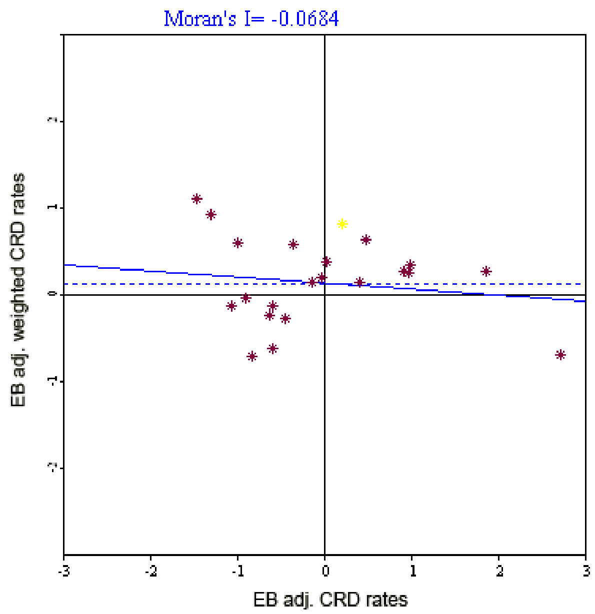


Figure 24: *Lisa box plot and Moran scatter plot for Chronic Respiratory disease Empirical Bayes adjusted rates*

4.2 Frequency of other factors that possibly contribute to CRD development

Etiologically, the factors that contribute to higher prevalence of chronic airway diseases have not been entirely explained. The examination of data on the exposure of children to cigarette smoke (Erzen, 2007) indicates that the number of people who smoke in the presence of children is still very high (Table 1). A quarter of parents of children with a chronic pulmonary disease smoke in the presence of children. The portion of parents who have healthy children and smoke in their presence is only slightly higher. When taking into consideration the statement that they smoke in residential premises when children are not present, the difference gets even smaller.

Table 1: *Smoking habits of adults in flats occupied by healthy and ill children*

	Habits	healthy children (%)	Children with manifested CRD (%)
SMOKING IN THE FLAT	Yes	**7**	**6**
	Yes, but ventilate the room	**13**	**8**
	Yes, periodically	**8**	**10**
	Yes, in absence of children	**5**	**7**
	No	67	69
THE NUMBER OF CIGARETTES PER DAY	0 cigarettes	71	70
	1 - 10 cigarettes	22	22
	11- 20 cigarettes	6	8
	More than 20 cigarettes	1	0

The examination of data on the distance to roads with heavy traffic, on airing methods of flats, and on how much time children spend in the open air, shows there is no essential difference between the group of children with chronic pulmonary disease and the group of healthy children. There are, however, differences with regard to physical activities of children. As expected, children with chronic pulmonary disease are less physically active and have more difficulties in carrying out physical exercise than healthy children.

The data on the incidence of allergic diseases in a family indicates that these diseases are also partly hereditarily predisposed. As much as 43% of children with chronic pulmonary disease reported that someone of their close relatives had an allergic disease and 9% reported that other relatives had allergic disease. The proportion of healthy children reporting that their close relatives had an allergic disease was substantially lower, i.e. 28% (Table 2).

Table 2: *Data on allergic diseases in a family*

	Allergies in family	Healthy children (%)	Children with manifested CRD (%)
The data on the incidence of allergic diseases	Yes, both	2	9
	Yes, one of them	18	23
	Yes, grand parents	8	11
	Yes, relatives	5	9
	No	63	46
	Unknown	4	2

4.3 Geographic analysis of air pollution sources

The most important sources of air pollution in Koper are industry, traffic and heating systems. Industry and traffic pollute the air over the entire year. Weekly and seasonal dynamics linked to the traffic density is typical for the air pollution from traffic. The largest emissions are present in summer at weekends when the traffic is the densest. Air pollution from heating systems is mainly linked to the cooler half of the year.

With regard to the air pollution from industry the Agency of the Republic of Slovenia for the Environment reported on 12 largest industrial sources of air pollution in the territory of the Municipality of Koper in 2005 (Municipality of Koper, 2005, Harmel, 2006). The quantity data show that the companies in the Municipality of Koper produce most carbon dioxide (CO2). The major carbon dioxide producer is the company Za Gradom, a little less of CO2 emissions are released into the air by Kemiplas. Kemiplas also stands out in terms of the largest annual quantity of most of the pollutants, especially carbon monoxide (CO), organic compounds (TOC) and sulphur oxides (SO2). The largest annual quantity of nitrogen oxide (NOx) is produced by the company Tomos. Port Koper (Luka Koper) causes most of the dust emissions (total dust). Companies are for the most part aware of the environmental problems and regularly ensure operational monitoring of emissions into the air which show that the emissions are within the prescribed limit values and concentrations. The report entitled Environmental Survey of Companies in 2006 (Harmel et. al. 2006) specifies only one company that exceeded the permissible limit values for the NOx parameter (Figure 25).

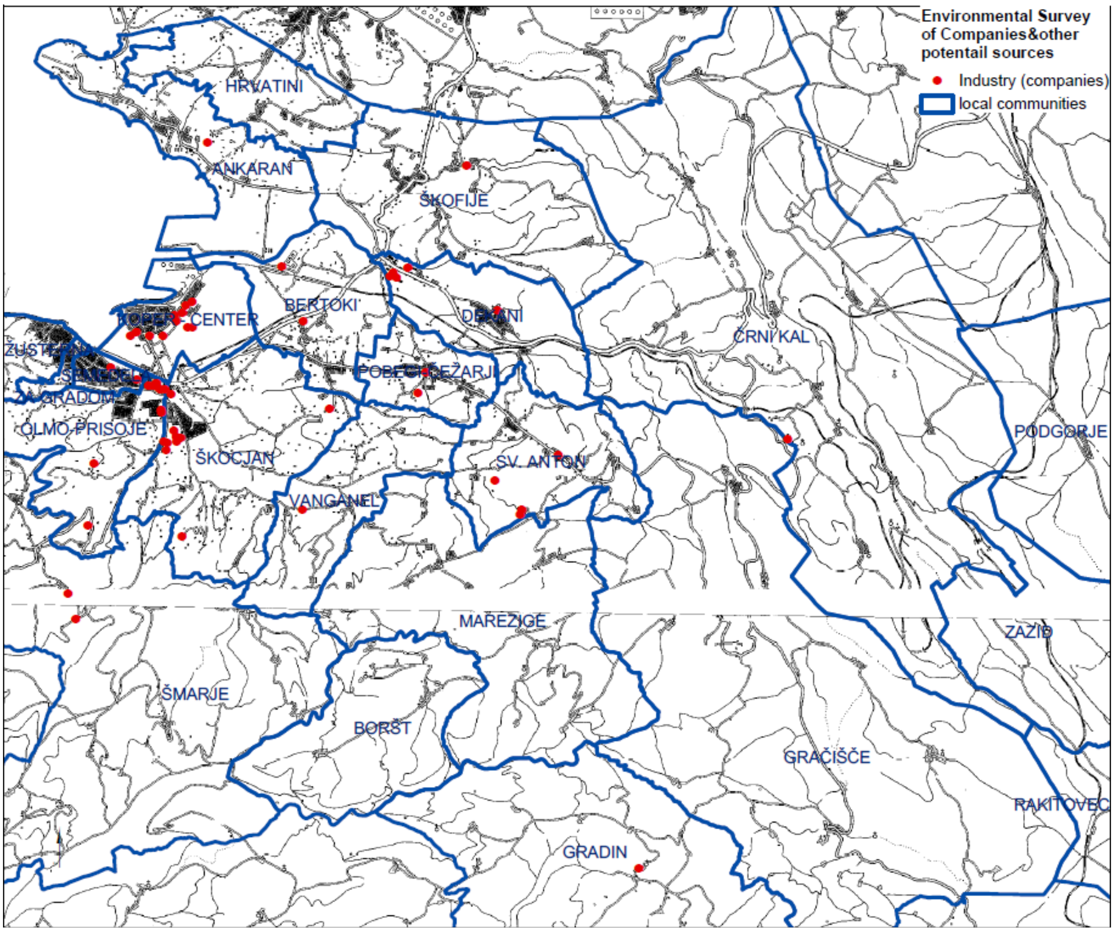


Figure 25: *Geographic distribution of companies and other potential pollution sources (traffic/roads)*

In the assessment of air pollution in the Municipality of Koper also the long-range transport of polluted air must be taken into consideration, especially the cross-border pollution from the Po River plain. The data from the studies conducted by ARSO (ARSO, 2008a, 2008b) allow an estimate that there is a certain contribution of long-range air transport of pollutants to the worse air quality in the area of the Municipality of Koper. It is not known what portion of the measured concentrations can be attributed to this transfer because detailed analyses of emission and meteorological parameters would be required (ARSO, 2008a).

Some findings support the study "Numerical Modelling and Scenario Analysis of Pollution Extension near Lovran Station" (ARSO, 2008b). On the basis of simulations with the numerical meteorological model ALADIN and subsequently with the transport model FLEXPART, the authors have established that in the described nine episodes with the highest measured ozone concentrations in the station Lovran it was mainly a matter of influence from near sources in Koper and Trieste. In support of this assumption are the strongly correlated values of modelled pollution concentrations from pre-set sources with the measured progress of ozone concentration. The direct influence of more distant sources is temporary. In some situations, a part of the measured concentrations can be attributed to sources in Monfalcone, more rarely to those in Umag. In these medium-distant influences, it is interesting that the 12-hour shift between concentrations of modelled and measured values cannot be noticed, which probably means that ozone is already present in the air mass that comes to Lovran from Monfalcone or Umag – from direction south-west and south-east.

We examined the data of previous measurements of air pollution in the area of municipality Koper and analyzed the impact of meteorological features of this area. As main sources of air pollution, we considered: Port Koper, fuel depot Srmin, industrial zone around the intersection Ankaran and main roads (traffic). We then on basis of those maps defined the geographical areas with different levels of risk (pollution) - geographical areas with increased pollution, lower pollution and unpolluted areas (Table 3).

Table 3: *Definition of polluted areas - a review based on meteorological criteria and industrial activity in municipality Koper*

Given the occurrence of shallow temperature inversions	**Given the prevailing wind direction (long range transport)***	**Industrial activity**
	Ankaran	
	Sv. Anton	Sv. Anton
Bertoki	Bertoki	Bertoki
	Črni kal	
Dekani	Dekani	Dekani
	Hrvatini	
	Koper	Koper
Pobegi Čežarji	Pobegi Čežarji	Pobegi Čežarji
	Semedela	Semedela
Škofije	Škofije	
	N. part Prisoje-Olmo (Šalara)	N. part Prisoje-Olmo (Šalara)
		Koper-center
		Škocjan

* based on the observation of meteorologic monitoring stations: Port Koper, Tinijan (for resources in the industries and the Port Zone); for traffic Koper and Žusterna

The first group represents areas with increased pollution, that include local communities Ankaran, Bertoki, Dekani, Hrvatini, Pobegi-Cezarji, Olmo-Prisoje, Semedela, Sv. Anton, Škocjan and Škofije. The second group is supposed to be unpolluted including communities: Borst, Črni kal, Gračišče, Gradin, Marezige, Podgorje, Rakitovec, Šmarje, Vanganel, Zazid and Žusterna. In addition, we created a third group namely, those that were considered between first two groups (less polluted) – Za gradom and Koper-center (Table 3, Figure). When comparing the number of cases between more polluted and communities considered less or non-polluted, then we summarize the numbers and get 98 cases in communities considered more polluted (69 polluted plus 29 less polluted) against 26 cases in areas considered non-polluted. In comparison also the prevalences of diseases show a similar trend with an average prevalence of 13.3 cases in more polluted areas to 8.6 cases per 100 at risk in non-polluted communities (considered non-polluted).

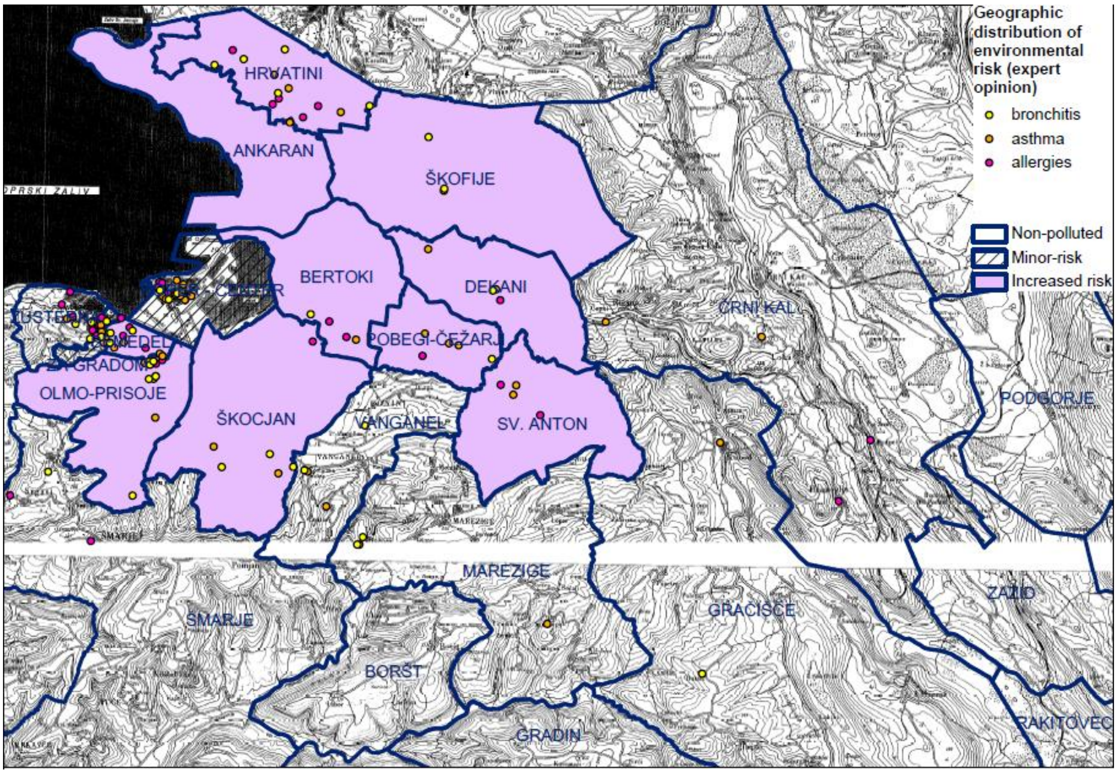


Figure 26: Geographic distribution of environmental risk based on the expert opinion

5 DISCUSSION

The present study addresses a topic that is now certainly important, as the number of chronic respiratory disease cases in population of Slovenian adolescents is perceived as rising. Asthma has become the most prevalent chronic disease in children and is the leading cause of hospitalization of children up to 15 years of age (Asher, 2006). In the years 1999-2004, the number of children suffering from asthma in some countries of Europe raised from less than 5% to more than 20% (ENHIS, 2007).

The purpose of the geographical-spatial research on increased morbidity is, in addition to proper informing of risk posed by the causes of living environment, to give out the information about spatial distribution of disease burden and thereby contribute to the accurate design of health-education programs, prevention and screening campaigns, while encouraging the formation of new hypotheses in order to clarify the causes for the observed geographic differences in localized areas.

World Health Organization (WHO) in its recommendations (WHO, 1999) advises the use of geographic analysis in assessing the possible increased risk of the disease under investigation linked to the environment. Following a recommendation the geographical analysis should be designed multi-layered: first preparing descriptive study, then setting the hypotheses, and final check with the correlation studies. In methodological approach it makes sense to use GIS and multivariate modeling by incorporating spatially dependent variables. At least as urgent as a methodological approach, however, is the ideal presentation and interpretation of results. The Slovenian public-health profession is responding to questions about the alleged prevalence of respiratory disorders in certain areas. Analyses of CRD diseases made in recent years in municipalities of Koper and Zagorje ob Savi have been primarily descriptive in nature (Erzen, 2003, 2006).

The present study provides an insight into the results of geographic analysis in municipality Koper. We analyzed the connections between available data and processed data of analysis conducted in 2003 in GIS and by using methods for assessment of clustering and individual clusters. The first part of the discussion is devoted to the methodology and their limitations. In the second part of discussion we focus on the results of our investigation, therefore, the characteristics of the spatial allocation of the selected cases and the potential impact/link with the environmental factors (geographic burden) on occurrence of these diseases in recent years in municipality Koper.

5.1 Discussion on approaches and methods

In children included in the 2002/03 research, we checked whether the occurrence of the respiratory disease is manifested in the geographical sense, and if so, where are the areas with increased morbidity. Now days the influence of environmental, behavioural, socioeconomic, genetic factors in certain area on health of people living there, is studied with help of methods, developed in spatial epidemiology. To avoid confounding factors as much as possible, the study was conducted in small area, where the living conditions are similar and ideally the only difference between different groups of participators would be the level of exposure to the factor being studied. Besides the benefit for the accuracy of the study, there are also considerable weaknesses connected with such an approach, especially because the number of people living in such area could be small and the influence of extreme values could be quite substantial (Elliot, 1996; Waller and Gotway, 2004).

The location (address) of each child that responded to the invitation and participated in the research was geocoded, geolocated, and entered in GIS. In this way we created conditions for an epidemiological analysis of the prevalence of the occurrence based on the location of residence. This approach is appropriate, especially in cases when there is a presumption that a certain environmental factor is responsible for the occurrence of the disease. The decision whether clustering occurs, was adopted on the basis of spatial statistics, mapping of prevalence rates, and spatial correlation analysis of chronic respiratory diseases in children.

Although we used residential location for geocoding, such locational assignment may not be entirely satisfactory for some applications. For instance, assigning residence location to each case ignores human mobility and may assign cases to locations far from areas where relevant (e.g., occupational or school-based) exposures occur. Also, people move from place to place during the course of the day and may receive significant environmental exposures at their workplace, in their car, or in other locations. Finally, in studies of chronic diseases (such as various other diseases) where disease onset may occur years after the suspected relevant exposure(s), appropriate locational assignment may involve collection of historical housing and occupational data for each case and any relevant noncases collected as a comparison group. These are some of the problems that are encountered when working with spatial data, it is of most importance to be aware of them, but it should not let them deter us from spatial analysis.

We began with EDA and ESDA methods aiming at defining first order patterns of disease appearance and continued with the ESDA methods for rates, respectively prevalences of chronic respiratory diseases. We proceeded with smoothing methods and probability map, both offering a first step toward clarifying the signal by attempting to control (or at least model) the noise inherent in the summary statistics and to pool values having similar expectations. Although such approaches aid us in displaying and understanding summary values from small areas, they do not address a fundamental issue involved in the analysis of geographically aggregated data, is there an evidence of clustering, clusters.

We investigated this hypothesis by using spatial autocorrelation methods as tests for defining disease clustering and clusters for regional count data. Using both boundary and k-nearest neighbor weights, we tested with global spatial auto-correlation methods at community's spatial resolution for clustering. We followed by the analysis of local indexes of spatial autocorrelation to define possible clusters and outliers.

A limitation of the LISA methods is their failure to correct for multiple comparisons while testing for clustering. Some authors have proposed that when using a significance level of 0.05 and no correction factor for multiple comparisons, up to 5% of the areas are identified by random chance variation (Weiss and Wagener, 1990). However, Rothman (1990) argues that no adjustments are needed for multiple comparisons since reducing type I error increases type II error thereby reducing the power to detect true statistically significant differences. To our knowledge, there seems to be no consensus on how to adjust for the multiple comparisons when using this kind of methodology

Due to the fact that the spatial patterns of disease distribution may change depending on the spatial scale used for mapping, a phenomenon known as the Modifiable Areal Unit Problem (MAUP). The term MAUP arises from the fact that areal units are not natural but arbitrary constructs that may not necessarily have a relationship with disease distribution. The solution to the MAUP would be in performing spatial analyses at more than one areal unit, to minimize its effects, but as the scale of research was a small area this wasn't feasible.

The findings may also be limited by the fact that the rate of prevalence of the disease was calculated on the basis of the portion of diseased children taken from a sample of returned questionnaires (only 60.4 % of the questionnaires had been returned). In some of the local communities, there was no case of asthma, which can be a result of unresponsiveness of those included in the research, but most likely it has to do with the fact that the number of children in this area is very small.

5.2 Discussion on results

The research is a continuation of the study »The assessment of environmental impact on prevalence of specific diseases and mortality rate in Municipality of Koper« carried out in 2003 whose purpose was to obtain data and to make an objective assessment whether the health of residents of individual local communities of Municipality of Koper was more at risk due to adverse environmental factors than in other areas where the environmental factors are present to a lesser extent. In the scope of the study of prevalence of chronic respiratory diseases in children, it has been established that the prevalence of asthma, chronic bronchitis and allergies represents a big burden on the health of children (Erzen 2003, 2006).

The results of previous analyses have shown that chronic respiratory diseases are more frequent in children in polluted areas than in areas classified as less polluted or unpolluted. The differences were not always statistically significant, however, they always point in the same direction. Also studies of cancer morbidity in Municipality of Koper have shown that in addition to greater skin cancer morbidity, which can be seen as a common phenomenon in the areas due to the insolation, there is also increased incidence of lung cancer in women, while leukemia morbidity was more than twice as great as expected. This last data can be purely accidental as the number of cases in the ten-year observation period is too small to allow more reliable conclusions.

The purpose of the research » Study of Characteristics of Incidence of Chronic Respiratory Diseases in Municipality of Koper from the Viewpoint of Potential Etiological Factors« was to establish whether there was clustering of patients in individual areas as regards the incidence of chronic respiratory diseases. Within the research we have studied whether there was increased risk for the occurrence of chronic respiratory diseases in certain areas of Municipality of Koper, and whether it can be linked with emissions of harmful substances from air pollution sources in the area of Municipality of Koper and in other areas that can have a negative impact on the air quality in Municipality of Koper due to specific meteorological conditions. Based on the data on the lifestyle of persons with chronic respiratory diseases, we have studied whether other harmful factors were present that could be responsible for the occurrence of the diseases. Special attention was given to checking exposure of the diseased children to passive smoking.

Two reasons were important to include children in the research to study etiological factors for the development of chronic respiratory diseases in the environment. The first is linked to the fact that small children are not exposed to the harmful influences of the work environment; Moreover, they do not yet have habits that could influence increased incidence of the disease of respiratory organs. In addition, children stay most of the time close to their homes, which suggests that the place of residence is also the place of hazard, and thus we can also draw conclusions on potential etiological factors responsible for the disease.

On the basis of the survey and spatial statistical analysis of incidence of respiratory diseases, we have come to the conclusion that there is continued clustering of these cases in certain parts of the municipality whether observing individual chronic diseases or all of them simultaneously. A marked clustering can be observed especially in the areas of local community Sveti Anton with the communities of “Marezige”, “Dekani”, “Pobegi-Cezarji” and “Vanganel”; the community of “Hrvatini” also stands out, however, the analyses do not show it as statistically significant. According to a high rate in local community ‘Hrvatini’, there is to be expected that the area through auto-correlation analyses borrowed less neighboring information than their interior counterparts. We suspect the edge effect to be responsible of over smoothing the districts rate and emphasize a possibility that the ‘Hrvatini’ may form the second cluster. Due to the small number of residents in some local communities, the number of diseased children is also relatively small, however, it should be noted that there is a permanent sample of incidence of chronic respiratory diseases in children.

There are several reasons for clustering of patients with chronic respiratory diseases, and it is impossible to define in detail the significance and contribution of a single reason as isolated observation is impossible and immission data are not available. Without any doubt external air pollution, living conditions, and, of course, hereditary factors are the main reasons. The most significant air pollutants in Koper are industry, traffic and heating systems. Weekly and seasonal dynamics linked to the traffic density is typical for the air pollution from traffic. The largest emissions are present in summer at weekends when the traffic is the densest.

The high ozone concentrations recorded in the area of Municipality of Koper show that the problem of air pollution in this area is not marginal. Ozone in the air in urban environment is an unquestionable indicator of pollution of air with organic substances. This means that the air in Municipality of Koper is polluted, even though measurements of content of certain other substances had shown that those substances did not exceed the permissible levels or were even at a very low level. The fact that the data based on the measurements of industrial emissions indicate that the emission pollution is not excessive does not mean that the air burden is really low. It is necessary to take into account the concentration of pollutants in these areas as well as the long-range transport of polluted air, especially cross-border pollution from the Po river plain, because immission concentrations indicate that the problem of air contamination in Municipality of Koper is highly relevant to the health.

The influence of polluted air is additionally increased by specific weather conditions. For Slovensko Primorje a strong influence of the sea climate is typical, and at the same time also big gradients when passing to the hinterland. Already in the area between the coast and the karst level big differences are noticeable. The coast is well-indented and especially the wind conditions as well the air temperature are locally strongly modified. Very important in the coastal area is the fog spreading over Slovensko Primorje from over the northern Adriatic and the Po river plain, and accompanied by a strong temperature inversion. Such fog can reach some 100 m high, sometimes the air layer mixes right above the ground and the fog lifts to low cloudiness. It can persist for several days because it can be dispersed only by the inflow of cooler air, usually at the onset of a cool front. Important are also shallow ground inversions. They appear at the bottom of valleys and in areas with a concave surface. Specific clustering of chronic respiratory diseases could at least to some extent be attributed to weather conditions.

The comprehensive geographic/spatial analysis of pollution sources and meteorological conditions suggests that the influence of cross-border pollution is quiet significant, especially in the direction from Trieste and Monfalcone also from Umag. These findings are supported by the incidence of the disease in the local community of Hrvatini and in the surrounding area of Koper in the local community of “Marezige”, “Dekani”, “Pobegi Cezarji” and “Vanganel”. The analysis of locations also confirms the assumption that most live in the part that overlooks the Bay of Trieste and in the hinterland of the city of Koper. Based on the meteorological data from the measuring station in Koper and the mobile station in Lovran (ARSO 2007, 2008a) we can conclude that the main reason for pollution in Municipality of Koper is the transfer of the ozone-polluted air from the densely populated and industrial northern Italy with greater sources of emissions of ozone precursors. This hypothesis is confirmed by the results of the study Numerical Modelling and Scenario Analysis of Pollution Extension near Lovran Station (ARSO, 2008b), and the fact that cases have occurred in higher lying areas of Municipality of Koper. These data suggest that the contribution of emission/immission concentrations is definitely a significant risk factor and most probably the key factor for the occurrence of chronic respiratory diseases in children in the Municipality of Koper, which is indicated by the cross-border data on the health consequences that are attributable to polluted air in the cities of northern Italy (Martuzzi et al. 2006).

The harmful effects of the strong industrial activity in the part of Trieste situated below Hrvatini have already been studied by the Italians who found that the incidence of cancer is increased in the population who live in the proximity of the incineration plant (Barbone et al.1995). This fact and the data on health consequences attributable to the polluted air in the cities of northern Italy show that the pollution is heavy, therefore adverse effects can be expected in the Koper area as well.

Factors contributing to increased incidence of chronic respiratory diseases do not exist only in the external environment. There are confirmed assumptions that asthma is more frequent in those children that are exposed to cigarette smoke, frequent virus infections of respiratory organs, and in children whose relatives have allergies. It has been established that asthma is more frequent in children with any type of allergic disease as well as in prematurely born. Neither the connection between asthma and air pollution, nor the influence of the broader living environment, and the significance of the diet in early childhood have been explained yet.

The examination of data on the exposure of children to cigarette smoke indicates that the number of people who smoke in the presence of children is still very high. A quarter of parents of children with a chronic pulmonary disease smoke in the presence of children. The portion of parents who have healthy children and smoke in their presence is only slightly higher. When taking into consideration the statement that they smoke in residential premises when children are not present, the difference gets even smaller. Also the hereditary factor is very important. The data on the incidence of allergic diseases in a family show that in the group of children with chronic pulmonary disease as much as 43% reported that someone of their close relatives had an allergic disease, and 9% reported that other relatives had allergic disease. The proportion of healthy children reporting that their close relatives had an allergic disease was substantially lower, i.e. 28%. The influence and significance of other factors for the development of chronic respiratory diseases, especially genetic ones, can in some cases explain the clustering of patients in a small area, as it still happens that in certain places larger groups of relatives live together. Other possible factors that we could not take into account for the prevalence of CRD diseases is exposure to indoor pollutants, duration of residence and concentration values of other possible pollutants.

In cases, where agent is airborne the exposure assessment is exceptionally demanding and unpredictable since the adequate data are usually missing. Often only data on emission concentration are available and even those data are incomplete. The conclusions regarding health effects that appear due to exposure to airborne hazardous substances with help of those data are rather weak, only indirect and reliability of conclusions is falling with distance to measurement site. One of the most important prerequisite for the accuracy of the study is a good determination of the exposure (as an example: place of residence). The accuracy of the study and its results depends on how much the measured environmental factor and actual exposure of participants corresponds.

The next very important limitation of this approach, which is also characteristic for other methods that study impact of environmental factors on health, is due to the fact that harmful effect under study has influenced the health in the past and is not necessary present at the time when study is being conducted. It is of utmost importance to take those limitations into account while interpreting the results. This is necessary although there is a reasonable assumption that environmental conditions haven't changed that much, that this would have an important impact on rate of disease or health impairment.

The major limitations of the research are nonexistent emission and immission values making the geographic modeling of pollution sources not possible and as a consequence prevent us to locate the most significant pollution sources. Even so, when comparing results of studies basing their results on pollution modeling conducted by Sashsurvaroglu (2009) and Oyana (2005), they acknowledged the limitation of not being able to control the effect of possible confounders when interpreting the results of correlation between prevalence and modeled risk.

5.3 Moving forward

At this moment it still cannot be exactly specified which sources of air pollution are the most significant factors for the increased incidence of morbidity of children in chronic respiratory diseases. Without any doubt the problem is a complex one. Despite this it can be concluded from the geographic analysis of increased incidence of morbidity and meteorological data that the long-range transport and the cross-border influence from the near Trieste area are very significant factors endangering the health of children and probably of other residents as well. Due to the fact that the polluted air endangers the health of residents of Trieste we believe it would be reasonable to establish contacts and cooperation with the health institutions and civilian associations in Trieste, and to manage activities together – in obtaining additional data and information as well as in taking measures to reduce pollution.

In Municipality of Koper it would be reasonable to continue identifying any possible sources of air pollution, and to strive to reduce pollution, regardless of whether a direct connection with the deteriorated health condition has been established or not.

There are probably several reasons for clustering of patients with chronic respiratory diseases, and it is impossible to define in the detail the significance and contribution of each single reason because isolated observation is impossible. Without any doubt external air pollution, living conditions, and of course hereditary factors are the main reasons. Air pollution in Municipality of Koper as indicated by the high ozone concentrations is not marginal. Ozone in the air in urban environment is an unquestionable indicator of pollution of air with organic substances, even though measurements of content of certain other emissions (substances) had shown that those substances did not exceed the permissible levels and were at a low level. The studies show the effect of long range transport of pollutants is not to be ignored. Also there are possible connections between the traffic pollution and cases, but because of missing immission data was not possible to make any define assumptions.

6 CONCLUSIONS

Among children that were included in the study »The assessment of environmental impact on prevalence of specific diseases and mortality rate in Municipality of Koper« carried out in 2002/03, we analysed the geographical allocation of Chronic Respiratory diseases and examined the areas with higher risk (morbidity).

The study visually demonstrated and statistically evaluated the existence of spatial clustering and a spatial cluster of CRD higher rates in the municipality. A marked clustering can be observed especially in the areas of local community Sveti Anton with the communities of "Marezige", "Dekani", "Pobegi-Cezarji" and "Vanganel"; the community of "Hrvatini" also stands out, however, the analyses do not show it as statistically significant.

In scope of geographic analysis of pollutant sources (pollutants), we compared the results of cases distribution with potentially more vulnerable areas/parts of the municipality of Koper. Despite the complexity of the problem the results of geographic analysis of incidence suggest the long-range transport, traffic and transboundary impact from the Trieste area as major factor threatening the health of children and probably the majority of the population of the Northern Adria region.

Information obtained from this study may be useful in planning public health measures and further research. However further epidemiological studies, based on the monitored pollution data, should identify risk factors responsible for the observed patterns of disease occurrence and provide an argument by which further action will be guided. This suggest the urging need of regular air quality monitoring of mayor traffic routes with mobile monitoring units and a collection of data from both permanent air monitoring stations Koper and Lovran.

Given that air pollution also threatens the health of the population of Trieste, we consider it appropriate to establish contacts and cooperation with health authorities and with civil society associations of Trieste.

SUMMARY

In children included in the 2002/03 research we checked whether the incidence of respiratory diseases is clustered in the geographical sense and if so, where are the areas of increased morbidity. Based on the geographic analysis, we studied whether there is increased risk for the incidence of chronic respiratory diseases in certain areas of Municipality of Koper, and whether it can be linked with the emissions of harmful substances from the sources of air pollution in Municipality of Koper and other areas that can due to specific meteorological conditions negatively influence the air quality in Municipality of Koper. Based on the data of the way of life, we studied in persons with chronic respiratory diseases whether other harmful factors are present that could also be responsible for the disease. Special attention was given to checking exposure of sick children to passive smoking.

Each of the children that responded to our invitation to participate was entered in GIS. Thus we established conditions for an epidemiological analysis of prevalence of the occurrence on the basis of location of residence. The definition whether there is clustering of the disease was adopted on the basis of the geographic and spatial analysis of chronic respiratory diseases in children. EDA and ESDA analysis of events and rates defined first order patterns of disease appearance. We extracted the ongoing trend of few local communities 'Hrvatini', ‘Semedela’ and area around ‘Sveti Anton’ appearing with relatively high rates. The Local Moran test of spatial auto-correlations adjusted for rates variability using both boundary and k-nearest neighbor weights were significant ($P < 0.05$). Few districts had significant Moran LISA values, ‘Sveti Anton' community (Local Moran $P = 0.0454$) had significant positive local spatial auto-correlation and defines a core of a spatial cluster and is extending into surrounding local communities ‘Marezige’, ‘Dekani’, ‘Pobegi-Cezarji’ and ‘Vanganel’.

We analysed the conditions in the area and have established that the influence of cross-border pollution from Monfalcone and Trieste is very significant. These findings are supported by the incidence of the disease in local communities near Koper and in the local community of Hrvatini. Also the results of geographic analysis show possible connection with higher prevalence levels of CRD diseases. The harmful effects of the strong industrial activity in the part of Trieste situated below Hrvatini have already been studied by the Italians who found that the incidence of cancer is increased in the population who live in the proximity of the incineration plant. The results show that cross-border influence of the activities in the core and greater area of the Bay of Trieste is not reflected only in Hrvatini but also in other areas in Koper, especially because it was established that cases of chronic respiratory diseases in children appear in increased numbers particularly in higher lying areas. In addition to this source of air pollution, local pollution from traffic and industrial activity must be particularly highlighted. Even though the Agency of the Republic of Slovenia for the Environment established through measurements that from the viewpoint of emissions

the pollution is not excessive, the concentration of pollutants and certainly the long-range transport of harmful substances through the air in this area are the factors due to which immissions of harmful substances are excessive. And it is the immissions that are relevant to the health.

POVZETEK

Med otroki, ki so bili vključeni v raziskavo leta 2002/03 smo preverjali, ali se pojav bolezni dihal grupira v geografskem smislu in če se, kje so območja povečane zbolevnosti. Z geografsko analizo smo proučili smo, ali obstaja na določenih območjih MO Koper povečano tveganje za pojav kroničnih bolezni dihal in ali ga je lahko povežemo z emisijami škodljivih snovi iz virov onesnaženja zraka na območju MO Koper in na drugih območjih, ki lahko zaradi specifičnih meteoroloških razmer negativno vplivajo na kvaliteto zraka na območju MO Koper. Na osnovi podatkov o načinu življenja smo pri osebah, ki so obolele zaradi kronične bolezni dihal, ugotavljali ali so prisotni še drugih škodljivi dejavniki, ki jim je mogoče pripisati odgovornost za obolevanje. Posebno pozornost smo namenili preverjanju izpostavljenosti obolelih otrok pasivnemu kajenju.

Vsakega od otrok, ki so se odzvali na vabilo in sodelovali v raziskavi smo vnesli v GIS. Na ta način smo vzpostavili pogoje za epidemiološko analizo razširjenosti pojava na osnovi lokacije bivanja. Opredelitev, ali prihaja do grupiranja obolenj smo sprejeli na osnovi geografske oziroma prostorske analize kroničnih bolezni dihal med otroki. EDA in ESDA analiza primerov in stopenj zbolevnosti je bila uporabljena z namenom definiranja primarnih vzorcev pojavljanja bolezni. Na podlagi rezultatov je razviden trend pojavljanja višjih stopenj zbolevnosti v krajevnih skupnostih 'Hrvatini', ‘Semedela’ in v okolici ‘Svetega Antona’. Lokalni Moran I test prostorske avtokorelacije prilagojen za variabilnost stopenj zbolevnosti daje statistično značilne rezultate ($p < 0{,}05$), tako ob uporabi sosedskih kot tudi mejnih uteži, za območje KS 'Sveti Anton' (lokalni Moran $p = 0{,}0454$), ki s signifikantnostjo in pozitivno prostorsko avtokorelacijo opredeljuje jedro gruče ter se razširja v okoliške krajevne skupnosti ‘Marezige’, ‘Dekani’, ‘Pobegi-Cezarji’ in ‘Vanganel’.

Analizirali smo razmere na tem območju in ugotavljamo, da je zelo pomemben vpliv čezmejnega onesnaževanja Tržiča in Trsta. Podpora tej domnevi oziroma tem ugotovitvam je zlasti pojavljanje obolevanja na območju krajevnih skupnosti ob mestu Koper in KS Hrvatini. Tudi rezultati geografske analize virov onesnaževanja kažejo na možno povezavo s pojavnostjo bolezni dihalnih poti na območju opazovanja. Škodljive vpliva močne industrijske dejavnosti v tem delu Trsta, ki leži pod Hrvatini, so proučevali že Italijani in ugotovili, da je tveganje za razvoj raka povečano med prebivalci, ki živijo v bližini sežigalnice odpadkov.
Rezultati kažejo, da se čezmejni vpliv dejavnosti na območju ožjega in širšega Tržaškega zaliva ne kaže samo na območju Hrvatinov temveč tudi na drugih območjih v Kopru, predvsem, ker je bilo ugotovljeno, da se primeri kroničnih bolezni dihal med otroki pojavljajo v povečanem številu zlasti na višje ležečih območjih. Poleg tega vira onesnaževanja zraka pa je potrebno posebej izpostaviti tudi lokalno onesnaževanje, ki je v veliki meri pogojeno s prometom in industrijsko dejavnostjo. Čeprav ARSO na osnovi opravljenih meritev ugotavlja, da z vidika emisij ne gre za prekomerno onesnaževanje, pa sta koncentracija

onesnaževalcev ter seveda daljinski transport škodljivih snovi po zraku na tem območju tista, zaradi katerih so imisije škodljivih snovi prekomerne. Prav imisije pa so tiste, ki so relevantne za zdravje.

7 REFERENCES

Alexander FE, Boyle P. Methods for Investigating Localized Clustering of Disease. IARC Scientific Publication 13. Lyon, France: International Agency for Research on Cancer, 1996.

Anselin L. Exploring Spatial Data with GeoDa: A Workbook. Urbana: University of Illinois, Center of Spatially Integrated Social Science, 2005.

Anselin L, Sybri I, Kho Y. GeoDa: An Introduction to Spatial Data Analysis. Geogr Anal 2006; 38: 5-22.

Anselin L. Local indicators of spatial association: LISA. Geogr Anal 1995; 27: 93-116.

Anselin L, Bera AK, Florax R, Yoon MI. Simple diagnostic tests for spatial dependence. Reg Sci Urban Econ 1996; 26: 77-104.

Anselin L, Kim YW, Syabri I. Web-based analytical tools for the exploration of spatial data. J Geogr Syst 2004; 6: 197-218.

ARSO. Meritve onesnaženosti zraka v Kopru od 18. januarja do 26. junija 2005. Ljubljana: Agencija RS Slovenije za okolje, 2006.

ARSO. Kakovost zraka v Sloveniji v letu 2006. Ljubljana: Agencija RS Slovenije za okolje, 2007.

ARSO. Numerično modeliranje in analiza scenarijev razširjanja onesnaženja v okolici postaje Lovran. Ljubljana: Agencija RS Slovenije za okolje, 2008.

ARSO. Meritve onesnaženosti zraka v Lovranu nad Ankaranom od maja 2010 do junija 2008. Ljubljana: Agencija RS Slovenije za okolje, 2008.

Assuncao RM, Reis EA. A new proposal to adjust Moran's I for population density. Stat Med 1999; 18: 2147-2162.

Bailey TC, Gatrell AC. Interactive Spatial Data Analysis. New York: John Wiley and Sons; 1995.

Barbone F, Bovenzi M, Cavallieri F, Stanta G. Air pollution and lung cancer in Trieste, Italy. Am J Epidemiol 1995; 141: 1161-69.

Besag J, Newell J. The detection of clusters in rare diseases. J R Stat Soc Ser A Stat Soc 1991; A 154: 327-333.

Bukowski et al. Risk factors for childhood asthma. Hum Ecol Risk Assess 2002; 8: 737-761.

Clayton DG, Kaldor J. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. Biometrics 1987; 43: 671-682.

Cleveland WS. Elements of Graphing Data. Monterey. CA: Wadsworth, 1985.

Cleveland WS, Devlin SJ. Locally weighted regression: An approach to regression analysis by local fitting. J Am Stat Assoc 1988; 83: 596-610.

Cressie NAC. Statistics for Spatial Data, rev. ed. New York: John Wiley & Sons, 1993.

Cressie N. Change of support and the modifiable areal unit problem. J Geogr Syst 1996; 3: 159-180.

Cressie N, Read TRC. Spatial data analysis and regional counts. Biometrical 1989; 31: 699-719.

Cromley EK, McLafferty SL. GIS and Public Health. New York: Guilford Press, 2002.

Elliot P, Wartenberg D. Spatial Epidemiology: Current Approaches and Future Challenges. Environ Health Perspect 2004; 112: 998-1006.

Elliot P, Cuzik J, English D, Stern R. Geographical and Environmental Epidemiology: Methods for Small-Area Studies. New York: Oxford University Press, 1997.

Elliot P, Wakefield J, Best N, Briggs D. Spatial Epidemiology: Methods and Application. New York: Oxford University Press, 2001.

Elliot P, Martuzzi M, Shaddick G. Spatial statistical methods in environmental epidemiology: A critique. Stat Methods Med Res 1996; 4: 137-159.

ENHIS, 2007. World Health Organization, Europe. Prevalence of asthma and allergies in children. 18.9.2009 Web page: http://www.enhis.org/object_document/o4735n27382.html.

English P, Neutra R, Scalf R, Sullivan M, Waller L, Zhu L. Examining associations between childhood asthma and traffic flow using a geographic information system. Environ Health Perspect 1999; 107: 761-767.

Eržen D. Epidemiologija kronične obstruktivne pljučne bolezni (KOPB), astme in emfizema v Sloveniji in v svetu. Obravnava bolnika s KOPB. – Ljubljana: Golnik, Pnevmološka sekcija Slovenskega zdravniškega društva, 1997: 22-36.

Eržen I, et al. Proučevanje vpliva okolja na pojav določenih bolezni in povečano stopnjo umrljivosti prebivalcev na območju dela Mestne občine Koper. End report. Koper: Mestna občina, 2003.

Eržen I, Podkrajšek D, Juričič M, Uršič A, Zadnik V, Zaletel-Kragelj L. Proučevanje vpliva okolja na pojav določenih bolezni in povečano stopnjo umrljivosti prebivalcev na območju občine Zagorje ob Savi. End report. ZZV Celje, 2006.

Eržen I, Zaletel-Kragelj L, Erlih S. Proučevanje značilnosti pojavljanja bolezni dihal na območju Mestne občine Koper iz vidika potencialnih etioloških dejavnikov. End report. Koper: Mestna občina, 2007.

ESRI - Environmental Systems Research Institute. Getting to Know Arc View GIS. CA: Redlands, ESRI Press, 1999.

Gelman A, Price PN, Lin C. A method for quantifying artefacts in mapping methods illustrated by application to headbanging. Stat Med 2000; 9: 17-18.

Gelman A, Price PN. All maps of parameter estimates are misleading. Stat Med 1999; 18: 3221-3234.

Elliot P. et al. (Ed.). Geographical and environmental epidemiology: methods for small-area studies. New York: WHO Regional Office for Europe by Oxford University Press, New York, 1996.

Griffith D. What is spatial autocorrelation?. Espace Geogr 1992; 21: 265-280.

Haining R. Spatial Data Analysis in the Social and Environmental Sciences. Cambridge: Cambridge University Press, 1990.

Harmel M et al. Okoljski pregled podjetij v občini Koper- končno poročilo. Koper: Mestna občina, 2006.

Kafadar K. Choosing among two-dimensional smoothers in practice. Comput Stat Data Anal 1994; 8: 419-439.

Kafadar K. Smoothing geographical data, particularly rates of disease. Stat Med 1994; 15: 2539-2560.

Kopriva S. Vpliv kajenja staršev na razvoj in bolezni dihal otrok. Za srce 2002; 3: 12-13.

Kopriva S, Maček V, Śupevc M, Kos M, Kopriva-Pirtovšek K. Epidemiologija astme pri otrocih v Sloveniji. Astma pri otroku. Ljubljana: Klinični center, Pediatrična klinika, Služba za pulmologijo, 2003.

Krieger N, Waterman P, Lemieux K, Zierler S, Hogan LW. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. Am J Public Health 2001; 91 (7): 14-16.

Kvamme K. et al. GIS - Geografski informacijski sistemi. Ljubljana: Znanstveno-raziskovalni center SAZU, 1997.

Lawson AB, Biggeri A, Dreassi E. Edge effects in disease mapping. In: Lawson A,. Biggeri A, Bohning D, Lesaffre E, Viel JF, Bertollini R (Eds.), Disease Mapping and Risk Assessment for Public Health, 85-98. Chichester: West, 1999.

Lawson AB. Statistical Methods in Spatial Epidemiology. Chichester, West Sussex: John Wiley & Sons, 2001.

Lawson AB, Browne WJ, Vidal Rodeiro CL. Disease Mapping with WinBUGS and MLwiN. Chichester: John Wiley & Sons, 2003.

Lee J, Wong DWS. Statistical Analysis with ArcView GIS. New York: John Wiley & Sons, 2001.

Lonhley PA, Goodchild MF, Maguire D, Rhind DW. Geographic information Systems and Science. Chichester, West Sussex: Wiley & Sons, 2001.

Martuzzi M, Mitis F, Lavarone I, Serinelli M. Health impact of PM10 and ozone in 13 Italian cities. Geneve: WHO Europe, 2006.

Moran PAP. Notes on continuous stochastic phenomena. Biometrika 1950; 37: 17-23.

Oden N. Adjusting Moran's I for population density. Stat Med 1995; 14.

Ord JK, Getis A. Local spatial autocorrelation statistics: Distributional issues and an application. Geogr Anal 1995; 27: 286-306.

Otorepec P. KOS – Astma in alergijske bolezni pri otrocih. 18.9.2009 Web page: http://kazalci.arso.gov.si/?&ind_ id=189&data=indicator&menu_group_id=25.

Oyana TJ, Rivers PA. Geographic variations of childhood asthma hospitalization and outpatient visits and proximity to ambient pollution sources at a U.S.-Canada border crossing. International Journal of Health Geographics 2005; 4(1):14.

Pearce N et al. Asthma epidemiology: principles and methods. New York: Oxford University Press, 1998.

Planinšek, A (ed.). Onesnaženost zraka v Sloveniji v letu 1996. Ljubljana: Ministrstvo za okolje in prostor, Hidrometeorološki zavod RS, 1997.

Richardson S, Thomson A, Best N, Elliott P. Interpreting Posterior Relative Risk Estimates in Disease-Mapping Studies. Environ Health Perspect 2004; 112: 1016-1025.

Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990; 1: 43-46.

Rothman KJ, Greenland S. Modem Epidemiology (2nd ed.). Philadelphia: Lippincout-Raven, 1998

Sahsuvaroglu T, Jerrett M, Sears MR, McConnell R, Finkelstein N, Arain A, Newbold B, Burnett R. Spatial analysis of air pollution and childhood asthma in Hamilton, Canada: comparing exposure methods in sensitive subgroups. Environ Health 2009; 8:14.

Šuškovič S, Košnik M, Šorli J. Astma, Golnik: Klinika za pljučne bolezni in alergijo, 2000.

Šuškovič S. Etiologija, epidemiologija in diagnostika astme in KOPB. Farm Vestn 2005; 56: 147-150.

Terčelj-Zorman M. 2006. Kriteriji za diagnozo astme, astma ob naporu in bronhomotorna preodzivnost. Delo in zdravje 2006; 29, 2: 30.

Tiefelsdorf M. Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I. Berlin: Springer-Verlag, 2000.

Tobler W. A computer movie simulating urban growth in the Detroit region. Econ Geogr 1970; 46: 234-240.

Tobler W. 1979. Smooth pycnophylactic interpolation for geographical regions (with discussion). J Am Stat Assoc 1979; 74: 519-536.

Tukey JW. Exploratory Data Analysis. Reading, MA: Addison-Wesley; 1977.

Uršič A, Krulec A, Zagajšek A, Tominšek L, Samardžič S. Preliminarni monitoring onesnaženosti zraka v MO Koper-poročilo o rezultatih meritev. Strokovno poročilo. Celje: ZZV Celje, 2000.

Waller L, Gotway C. Applied Spatial Statistics for Public Health Data. New Jersey: John Wiley & Sons, 2004.

Walter SD. The analysis of regional patterns in health data. II. The power to detect environmental effects. Am J Epidemiol 1992; 136: 742-759.

Weiss ST, Wagener DK. Geographic variations in US asthma mortality: small-area analyses of excess mortality 1981–1985. Am J Epidemiol 1990; 132: 107-115.

Weiss ST. Environmental risk factors in childhood asthma. Clin Exp Allergy 1998; 28: 29-34.

WHO. Disease mapping and risk assessment for public health decision-making. Copenhagen:
WHO regional office for Europe, 1999.

Wise S, Haining R, Signoretta P. Scientific visualization and the exploratory analysis of area data. Environment Plann A 1999; 31 (10): 1825-1838.

Zeger SL. et al. Exposure measurement error in time-series studies of air pollution: concepts and consequences. Environ Health Perspect 2000; 108: 419-426.