A C T A U N I V E R S I T A T I S L O D Z I E N S I S
FOLIA OECONOMICA 252, 2011

Paweł Kretowicz[*]

THE UTILIZATION OF CANONICAL ANALYSIS IN IDENTIFYING FACTORS INFLUENCING HOSPITALIZATION RATES ON THE REGIONAL SCALE

Abstract. The main goal of presented research is to show the possibilities of canonical analysis utilization in evaluations of intensity and spatial distribution of factors influencing hospital prevalence rates in regions. The setting of this study is the mountainous areas of Podkarpackie Province divided into municipalities, forming a total of 217 spatial units.

The major reason for utilizing canonical analyses is the large variation of hospital prevalence rates within selected groups of diseases (cardiovascular and respiratory) and between different age groups. As a consequence, hospitalization rates by age groups are included in the analysis as dependent variables, whereas the independent variables used are various socio-demographic as well as standard of living indicators. These independent variables are grouped into three sets (demographic, social and economic), which characterize municipalities under investigation. The mean for the years 2006-2008 for each variable is calculated.

The product of the canonical analysis, canonical variables with the highest correlation to independent variables, are subsequently presented on cartograms showing spatial distribution of interdependence between the examined phenomena.

The final part of this study includes a short evaluation of the method used for the explanation of spatial hospitalization rate inequalities, as well as some possibilities for enhancement of research into spatial variation of ill-health on the regional scale.

Keywords: canonical analysis, spatial disparities of ill-health, hospital prevalence rate, Podkarpackie Province.

1. INTRODUCTION

The main goal of spatial analyses is not only to find and depict peculiar or atypical spatial structures, but also to determine the causes and explain usually complex phenomena. Thus, the utilization of numerous data sources and multivariate methods is currently a must. In those spatial sciences based on the neopositivist approach, a very common problem is the lack of sufficient data necessary to fully determine and understand factors influencing different patterns of socio-economic events. This study concentrates on explaining the spatial distri-

[*]MA, Department of Population Geography, Settlement and Agriculture, Jagiellonian University.

bution of ill-health as measured by hospitalization rates on the regional scale. Essentially, an attempt is made to explain the spatial disparities of hospital prevalence rates by producing factors which are related to them, instead of showing what and how particular factors cause diseases. In other words, the author endeavors to show one of the possible ways explaining different spatial distributions of hospital admissions within one Polish region (Podkarpackie Province).

Explaining the spatial distribution of ill-health has traditionally fallen under the domain of health geography. While this tradition dates back to Snow's 1855 simple map of cholera hotspots in London, a widespread interest in more profound analyses has arisen since then, together with developments of multilevel statistical techniques and computer programs facilitating their utilization.

Geographers have struggled to distinguish contextual (place) and compositional (population) effects on health (Curtis and Rees Jones [1998], Shaw et. al. [2002], Smith and Easterlow [2005]). Nevertheless, a collective approach has recently become more popular as place effects combining context and composition become more appropriate (Macintyre et. al. [2002], Smyth [2008]). As a result, regional and local scales have come into prominence as far as health inequalities are concerned. Most studies done now utilize multivariate statistical techniques (Malmstrom et. al. [1999], Pickett and Pearl [2001]), yet it is contended that whole procedures incorporating a set of statistics could be more beneficial (Diez-Roux [2002]). Although such research papers are often produced to support state and regional health policies (Drummer [2008]), explaining inequalities in hospital prevalence rates in Polish regions does not seem to be a dominant issue in health planning. Both regional health plans and governmental documents are generally concentrated on the crude presentation of main health indicators, health risks and some possible causes with hardly any attempts spatially and statistically linking one to another (Wojtyniak and Goryński [2008], Regional Health Plans).

The author chooses canonical analysis as a primary method of explaining spatial disparities in health on the regional scale after considering two main reasons:

- a comparison of two sets of variables renders a multitude of explanatory factors influencing hospitalization rates (here: environmental, social and economic variables) and a complexity of different hospital morbidity spatial distributions (varying across age groups, types of diseases, and types of community),
- a lack of longitudinal health data for spatially-defined locations leaves cross-sectional approach utilizing community characteristics. Such an approach has frequently been used elsewhere (Chaix et al. [2005], Breckenhamp et. al. [2007], Li and Newcomb [2009]).

2. METHODS

This study aims to determine the factors influencing different distributions of hospitalization rates in the urban and rural municipalities of Podkarpackie Province. An additional, methodical goal is to evaluate the utilization of canonical analysis in obtaining spatial distribution of the factors influencing hospital prevalence rates on the regional scale. Research procedures include a set of grouping assumptions and statistical methods, which collectively help to conduct the canonical analysis.

Canonical analysis is the most generalized version of multiple regression and represents correlations between two sets of variables - predictor and criteria; however, this distinction is not always necessary. The main idea is to establish optimal canonical weights for the variables in a way to maximize the linear correlation between predictor (X) and criteria (Y) variables. As a result, new variables are created - canonical roots (U, V):

$$\begin{aligned} U &= \mathbf{L}'\mathbf{X} = l_1 x_1 + l_2 x_2 + \ldots + l_p x_p \\ V &= \mathbf{M}'\mathbf{Y} = m_1 y_1 + m_2 y_2 + \ldots + m_p y_p , \\ & corr\,(U, V) = \max \end{aligned} \tag{1}$$

where: U, V – canonical variables (roots), m, l – canonical weights.

Canonical weights reflect the level and direction of interrelationships between two sets of variables. The importance of these relationships is evaluated by canonical correlations, thus the following correlation coefficients are always lower than the preceding ones, but not always significant. Although the mathematical basis and interpretations of canonical correlations are believed to be complex, readily understandable results are obtained by comparative analyses of factor structures - relationships between canonical variables and original variables represented by R or R^2 (Monmonier and Finn [1973], Runge [2007]). Only strong (decreasing variance) and significant (Chi square test, $p<0.05$) canonical correlations should be taken into account. Notably, such correlations are calculated for left and right sets, so the final interpretation is gained by a comparison of the two. Finally, redundancy scores show the overall variances explained by the particular root and, when summed up, the whole model.

As for the utilization of this method in spatial research, canonical scores are calculated and determine significant correlations between both datasets in the study area. Canonical scores identify the direction and degree of involvement of the observed values in the common patterns (Coshall and Potter [1987]). This enables regionalization of the examined interrelationships. In order to calculate canonical scores, raw data is multiplied by the appropriate canonical weights.

Canonical analysis was first utilized by Hotteling [1936], although the first geographical application is believed to be Berry's analysis of commodity movements in India (Berry [1966]). Many other studies dealt with socio-economic spatial analysis and utilized canonical correlations (Ray [1971], Norcliffe [1972]. Some Polish studies which made use of canonical correlations in socio-economic and spatial analyses include Ratajczak [1980], Zamelska [1980], Potrykowska [1983]. However, this technique remains among those less frequently used statistic methods. Some authors, though, have valued its use in time-space analyses (Kaczmarek and Parysek [1977], Runge [1991]).

3. DATA

Data used in this study was collected from two main sources: *The Center of Public Health of Podkarpackie Province in Rzeszów* and *The Central Statistical Office regional databank (www.stat.gov.pl*). All of the variables were assembled for the rural (143) and urban (45) municipalities of Podkarpackie Province. The health data included hospitalization rates calculated for circulatory system diseases and respiratory system diseases (averages for 2006-2008). Hospitalization rates (criteria variables) were age-standardized and divided in the following manner:

- children and young adolescents (age group 0-14),
- older adolescents and young adults (age group 15-34),
- older adults (age group 34-64),
- elderly (age group 65+).

These divisions were done because of the large differences within various health indices among different age groups (Townsend et. al. 1988).

A preliminary predictor dataset consisted of 54 population-standardized environmental (4), demographic (11) social (28) and economic (10) variables extracted from Central Statistical Office (average scores for 2006-2008) and 2002 Polish National Census. Preliminary data selection included calculating Pearson correlations between hospitalization rates for different age groups and each of the aforementioned variables. Insignificant coefficients ($p>0.05$) were excluded from further analyses.

Predictor variables were obtained through factor analysis (FA) conducted for the remaining preliminary variables. Extracted factors (predictor variables) are shown in Table 1.

Tab. 1. Factors (predictor variables) for circulatory/respiratory system diseases and rural/urban areas of Podkarpackie Province

DISEASES (AREA)	**VARIABLES**	**FACTORS**	**EXTRACTED VARIANCE**
CIRCULATORY SYSTEM (RURAL)	24	Sylvan areas, Social development and modern lifestyle, Unemployment, Social benefits, High fertility	67.9%
CIRCULATORY SYSTEM (URBAN)	31	Social development and modern lifestyle, Poor housing conditions, Ruralization, Unemployment, Favorable family relations	75.9%
RESPIRATORY SYSTEM (RURAL)	39	High social and economic development, Sylvan areas, Good housing conditions, Unemployment, Employment in services	65.9%
RESPIRATORY SYSTEM (URBAN)	24	Poor economic development, Social development, Sylvan areas, High mortality, Favorable family relations	72.2%

Source: developed by the author.

4. RESULTS

4.1. Circulatory system diseases in rural areas

First performed canonical analysis was of circulatory system hospitalization rates in rural areas of Podkarpackie Province. Although there are many factors influencing hospital prevalence rates, this study concentrates only on the potential factors underlying spatial differences. Hence, it is not always possible to find cause-effect relationships.

The canonical roots, correlations with original variables and redundancies for all of the rural municipalities in Podkarpackie Province are presented in Table 2.

Tab. 2. Results of the canonical analysis for circulatory diseases in rural areas of Podkarpackie Province

VARIABLES	**Can. Root 1**	**Can. Root 2**	**Can. Root 3**	**Can. Root 4**
Circulatory 0-14	0.08	-0.61	0.69	-0.38
Circulatory 15-34	0.78	-0.29	-0.22	-0.52
Circulatory 35-64	0.97	-0.10	0.21	0.07
Circulatory 65+	0.79	0.41	0.38	-0.26
CANONICAL CORRELATIONS	**0.46****	0.34*****	0.14	0.04
REDUNDANCY	**11.7%**	1.8%	0.3%	0.0%
TOTAL REDUNDANCY	**13.9%**			
Sylvan areas	0.76	-0.33	0.16	0.54
Social development and modern lifestyle	0.12	-0.76	-0.40	-0.51
Unemployment	0.57	0.37	0.30	-0.66
Social benefits	-0.03	-0.02	-0.08	0.05
High fertility	0.30	0.43	-0.85	0.10
** $p<0.001$; * $p<0.1$				

Source: developed by the author.

Only the first canonical correlation for circulatory system diseases in rural areas is significant, where a mildly strong correlation coefficient was produced (r=0.46). The whole model explains only 14% of the total variance within the criteria variables set, which indicates that many other factors affect health disparities across rural municipalities of Podkarpackie Province.

However, the first canonical correlation shows that a population which resides in municipalities with a high share of sylvan areas and high unemployment is more likely than other populations to suffer from circulatory system diseases. This chiefly concerns that part of the population who are between 35 and 64 years old. Although sylvan areas do not seem to exacerbate circulatory system diseases, both of the produced factors (sylvan areas and unemployment) mirror one's economic status – municipalities lying in the mountainous, southern areas of Podkarpackie Province are generally less developed.

The overall fit of this model is low; thus, in this case, the canonical analysis requires more or different factors to be included in order to obtain higher redundancy scores.

4.2. Circulatory system diseases in urban areas

Hospitalization data and health studies almost always prove that overall health deteriorates the closer one lives to an urban environment (as compared with rural areas). The gap between Podkarpackie Province's urban and rural municipalities in hospitalizations for cardiovascular diseases increases with age.

Furthermore, there are many more hospital admissions for circulatory system diseases in the 65+ age group (the maximum in Sędziszów Małopolski town – 2724 per 10 thousand inhabitants in 2006-2008) than in the youngest (0-14 year) age group (the maximum in Iwonicz-Zdrój town - 35 per 10 thousand inhabitants in 2006-2008). The canonical analysis conducted for urban areas in Podkapackie Province clearly mirror age patterns (Table 3).

Two canonical roots are significantly correlated with original variables. The first canonical correlation is very strong (r=0.74) and explains about 16.5% variance in hospitalizations for circulatory system diseases (mostly among the population 0-14 years old).

Generally, the children residing in urban areas are more likely to be hospitalized, as worse social development and less modern lifestyles characterize town populations. Interestingly enough, children in cities are also very likely to be hospitalized, as high housing density, low education and higher divorce rates, together with low economic activity result in lower standard of living, thus, less care in children's upbringing.

Tab. 3. Results of the canonical analysis for circulatory diseases in urban areas of Podkarpackie Province

VARIABLES	Can. Root 1	Can. Root 2	Can. Root 3	Can. Root 4
Circulatory 0-14	0.97	0.05	0.14	0.17
Circulatory 15-34	0.35	-0.64	-0.66	-0.18
Circulatory 35-64	0.33	-0.93	0.05	0.13
Circulatory 65+	0.18	-0.63	-0.20	0.72
CANONICAL CORRELATIONS	**0.74****	**0.64***	0.36	0.11
REDUNDANCY	**16.4%**	**17.2%**	1.6%	0.2%
TOTAL REDUNDANCY	**35.4%**			
Social development and modern lifestyle	-0.67	0.32	-0.14	-0.63
Poor housing conditions	-0.40	-0.47	0.31	-0.08
Ruralization	-0.52	0.16	0.49	0.61
Unemployment	-0.28	-0.70	-0.52	0.17
Favorable family relations	-0.21	0.42	-0.61	0.44
** $p<0.0001$; * $p<0.001$				

Source: developed by the author.

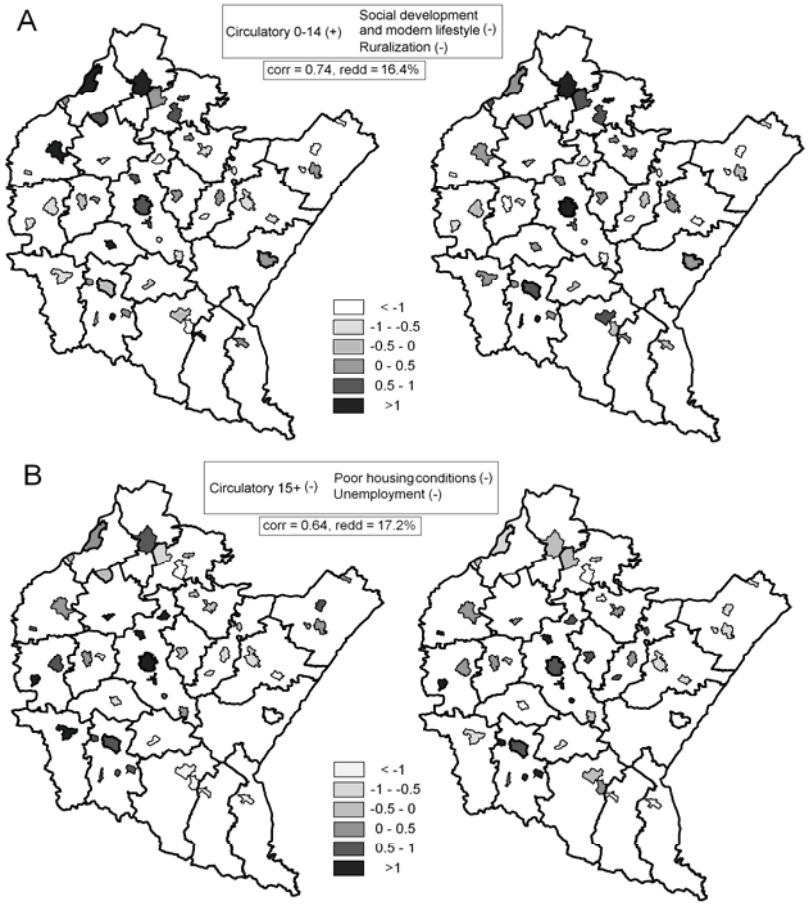
On the other hand, a decrease in the ruralization factor (a decrease in the share of farmlands and an increase in the share of forests), also contributes to higher hospitalization rates in the 0-14 age group. Again, this might suggest that more forested areas in a town negatively effects health but, in fact, it suggests that the worse economic status of small towns lying in the mountains is the real culprit.

However, the spatial distribution of the first canonical root (Fig. 1A) shows that this considered relationship occurs more often in the biggest cities, specifically those located in the industrial area located to the north of Podkarpackie Province. Sylvan areas dominate in the northern part of the Podkarpackie Province (Sandomierska primeval forest). Forests were also planted in cities with industrial function as natural barriers against pollution. The latter relationships do not exist in towns located in mountainous areas.

The second canonical correlation represents the almost equally strong and significant relationship (r=0.64) between hospitalization rates for 15+ age groups and some predictor variables, affecting chiefly the older adult age group (35-64 years old). Not surprisingly, poor housing conditions and high unemployment in towns are positively correlated with circulatory hospitalization rates. Once more, low economic and material statuses seem to have a negative impact on health outcomes. Similarly, the unemployment factor is connected to a decrease in older adults' health statuses, especially in males, who are more likely to suffer from circulatory system diseases. The scores from second canonical root are presented in the Figure 1B.

Towns located in the middle and the south of Porkarpackie Province follow a relationship between hospitalization rates and factors concerning housing conditions and unemployment.

Fig. 1. Spatial distribution of canonical scores for circulatory system diseases in the urban areas of Podkarpackie Province.



Source: developed by the author.

This may be due to jobs created by Rzeszow and Krosno and the generally worse housing conditions in small towns around them. High canonical correlations provide a firm model, which explains 35.4% of the whole variation in hospitalization rates, thus both relationships should render valuable information about possible determinants of health inequalities.

4.3. Respiratory system diseases in rural areas

Respiratory system diseases which require hospitalizations are widely attributed to air quality and sanitary conditions. This mainly regards urban and industrial areas; however, behavioral and cultural factors also play a significant role.

The most important of these factors include smoking, participation in sports and obesity, especially among adults. According to the study carried out by Wróblewska [2002], women with poor financial positions are more likely to suffer from respiratory diseases. One the other hand, children are more susceptible than adults to adverse respiratory diseases as they breathe, drink, and eat more, thus absorbing more environmental agents. For example, an infant's respiratory rate is more than twice an adult's rate (Sullivan [2007]).

A canonical analysis is utilized to show how socio-economic and land-use factors render variations in hospitalization rates in rural areas of Podkarpackie Province. As many as three canonical correlations turned out to be strong and significant ($p<0.05$).

Tab. 4. Results of the canonical analysis for respiratory system diseases in rural areas of Podkarpackie Province

VARIABLES	Can. Root 1	Can. Root 2	Can. Root 3	Can. Root 4
Respiratory 0-14	0.96	0.07	0.00	-0.27
Respiratory 15-34	0.68	0.67	0.06	0.29
Respiratory 35-64	0.71	-0.16	0.17	0.67
Respiratory 65+	0.51	-0.05	0.79	0.33
CANONICAL CORRELATIONS	**0.68*****	**0.46*****	**0.38****	0.18*
REDUNDANCY	24.7%	2.6%	2.4%	0.6%
TOTAL REDUNDANCY	**30.3%**			
High socio-economic development	0.31	0.30	0.02	-0.79
Sylvan areas	0.94	-0.10	0.08	0.32
Good housing conditions	0.15	-0.15	-0.57	-0.39
Unemployment	0.02	0.36	0.72	-0.16
Employment in services	0.02	0.86	-0.39	0.31
*** $p<0.0001$; ** $p<0.001$,	* $p<0.1$			

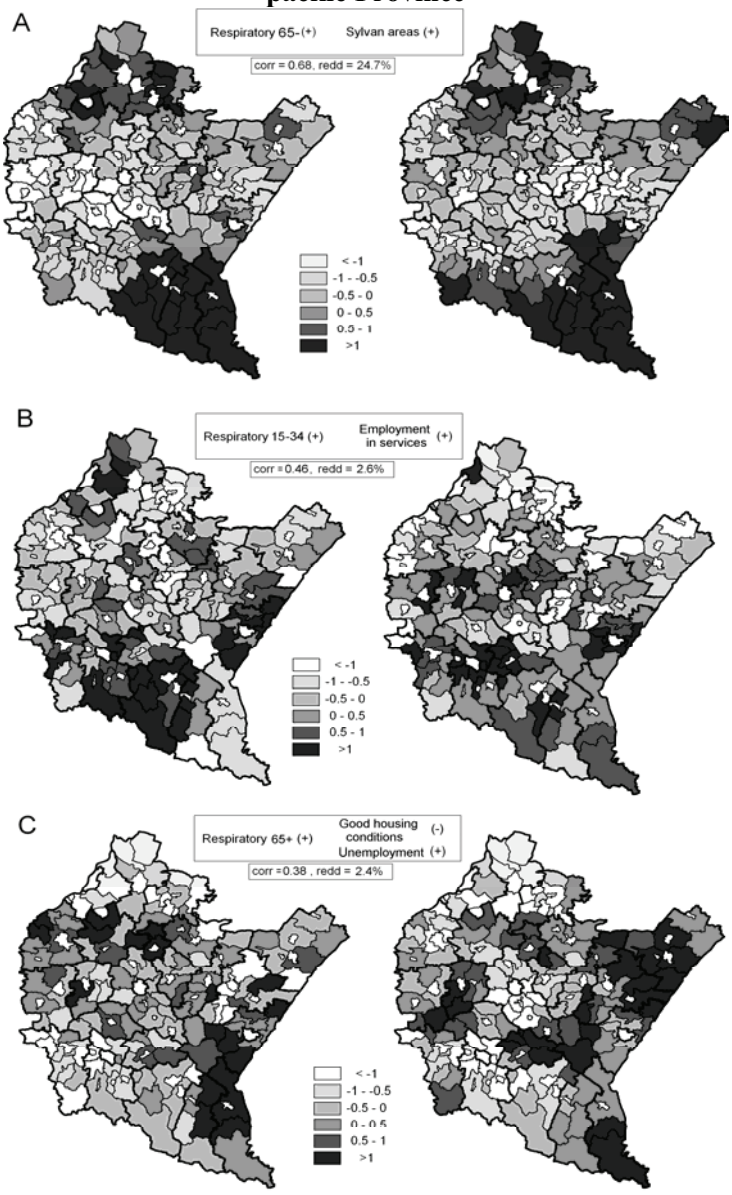
Source: developed by the author.

Respiratory system hospitalization rates of the population below 65 years of age is highly and positively correlated with the sylvan areas factor ($r=0.68$). In conjunction with previous results, this may be because southern municipalities with a high share of forested areas in Podkarpackie Province incorporate many socio-economic characteristics which contribute to low standard of living. Besides, such a high canonical correlation for the rural areas can be explained by poor accessibility to medical care and other services.

Figure 2A clearly shows that southern, inaccessible and forested municipalities are strongly related to higher respiratory hospitalization rates. A relationship between a high share of sylvan areas and respiratory hospital admissions was also found high in the northern part of Podkarpackie Province – the former Central Industrial District. Nowadays, this area is considered an ecological hazard zone with high pollution emissions (e.g. CO_2 and dust, as determined by the Polish Space - Spatial Development Report 2007). Moreover, the relationship

evidenced by the first canonical correlation particularly regards children, who, as mentioned, are the most vulnerable age group.

Fig. 2. Spatial distribution of canonical scores (first A, second B and third C canonical roots) for respiratory system diseases in rural areas of Podkarpackie Province



Source: developed by the author.

The second canonical correlation reveals a positive relationship between respiratory system disease hospitalization rates of older adolescents and young adults and the employment in services factor (r=0.46). Although this association explains only 2.6% of the total variance of criteria variables, when visualized on the map (Figure 2B), an evident pattern is shown – higher hospitalization rates related to this factor mainly concern suburban areas of the biggest cities. Naturally, the service sector attracts young workers from adjoining municipalities (e.g. in Rzeszów, Przemyśl, Krosno, Dębica, Sanok and Łańcut surroundings). Hence, a simple conclusion can be drawn - young people living close to the city also commute to work or school there, thus they are exposed to the unfavorable environmental conditions. The canonical analysis does not reveal such a pattern in the northern part of Podkarpackie Province, as many cities there still possess a high share of industrial employment.

Finally, the third canonical correlation (r=0.38) illustrates that a rise in respiratory hospitalization rates for the population above 65 years old occurs along with worse housing conditions and higher unemployment. High unemployment portrays a relatively lower economic development. This sort of deprivation may contribute to a lack of necessary financial support given to the elderly residing in rural areas by both family and the local government. Actually, it is easier to find a cause-effect relationship between housing conditions and respiratory problems among the elderly, who would report worse health while living in crowded and tight apartments along with their relatives.

As shown in the Figure 2C, scores for the third canonical root are distributed mosaically across Podkarpackie Province. Only the eastern borderland area and some suburban municipalities located to the northeast follow the aforementioned pattern. Unfortunately, lower canonical correlation and redundancy bring up the need to conduct a more detailed analysis concerning the spatial disparities of the elderly and respiratory diseases. Notably, the fourth canonical correlation, although low and insignificant (p<0.1), also shows an interesting relationship – respiratory hospital admissions among older adults increase according to a decline in socio-economic development. The reliability the whole model is confirmed by a high total redundancy – 30.3%.

4.4. Respiratory system diseases in urban areas

Respiratory system diseases are generally exacerbated by environmental factors such as pollution (smog), cold and damp housing. For that reason, contextual effects seem to have more impact on hospitalization rates. Urban environment always contributes to the increase in respiratory hospital admissions, which particularly concerns industrial cities or these with big car traffic. The highest hospitalization rates in Podkarpackie Province occur in the cities and towns located in the south-eastern part. Besides, relatively high rates for 0-14 age group

are also observed in the industrial cities located to the north. The canonical analysis for urban areas shows an evident relationship between a rise in respiratory hospitalization rates along with the increase in the share of forested areas, but decrease in economic development (Table 5).

Tab. 5. Results of canonical analysis for respiratory system diseases in urban areas of Podkarpackie Province

VARIABLES	Can. Root 1	Can. Root 2	Can. Root 3	Can. Root 4
Respiratory 0-14	0.80	0.07	0.34	-0.48
Respiratory 15-34	0.65	-0.03	0.76	0.07
Respiratory 35-64	0.79	0.52	0.09	0.32
Respiratory 65+	0.27	0.94	0.18	0.05
CANONICAL CORRELATIONS	**0.79***	0.51	0.34	0.11
REDUNDANCY	**27.8%**	7.5%	2.2%	0.1%
TOTAL REDUNDANCY	**37.6%**			
Low economic development	-0.70	0.24	-0.48	-0.47
High social development	-0.17	-0.94	0.05	-0.28
Sylvan areas	0.61	0.01	-0.22	-0.61
High mortality	-0.13	0.24	0.80	-0.52
Favorable family relations	-0.31	-0.11	0.27	0.25
* $p<0.0001$				

Source: developed by the author.

A very strong canonical correlation (r=0.79), as well as a high redundancy score (27.8%) were found for first canonical root. Predictor variables that correlate with respiratory hospitalization rates for patients under 65 inform that more sylvan areas within town borders worsen health outcomes. This reflects spatial configuration - higher hospitalization characterizes towns located in the industrial north and, to some extent, poor areas located in the southern-east part of the Podkarpackie Province, where there are also many forests. Surprisingly, hospitalization rates increase also along with improved economic development. This finding can be explained by the size of the town – larger cities are generally more developed and, in the end, more polluted, which may be a cause of higher respiratory hospital admissions. Second canonical root displays a strong but insignificant correlation, which shows a sensible relationship – a decrease in social development results in higher hospitalization rates for patients aged 65+. In general, this model explains more than 37% of the total variance in hospitalization rates, but requires more detailed factors to disaggregate sylvan areas factor.

5. DISCUSSION

A technique of canonical analysis was utilized to find environmental, demographic and socio-economic health determinants. A set of produced factors showed various degrees of relationship with hospitalization rates calculated for

four age groups. Canonical analysis was performed four times simultaneously in order to distinguish between circulatory/respiratory system diseases and urban/rural areas of Podkarpackie Province. As it was not always possible to find cause-effect associations, some results suggested strong relationships with other "hidden factors" (ecological fallacy) or showed a need to add more variables to the models. Generally, an important asset of canonical analysis was evidenced by sensible correlations as no reverse cause-effect outcomes were found.

First and the foremost, both social and economic development seem to have the biggest influence upon hospitalization rates in Podkarpackie Province. In almost all of the performed analyses unemployment was connected with worse health outcomes especially among people in productive age. Several other studies also reported this association (Payne et. al. [1994], Farr et. al. [2000], Beckfield [2004]). Low economic status was more often related to higher hospitalization rates in the rural areas and population in older age groups. However, this analysis lacks wealth indicators such as average household income what might have a sizable impact upon health. Foreign studies widely reported such a relationship for example in case of respiratory system diseases (Stelianides et al., [1999]), but some found contradict results (Morris et. al. [1994]). Financial situation could be caught by sylvan areas factor, which usually mirrored mountainous, forested and poorly-developed areas. Both unemployment and poverty indicators decide about community's social capital (Hancook [2001]) – this feature is also believed to elevate hospitalization rates in Podkarpackie Province. On the other hand, an accessibility to healthcare and other services would be another factors necessary for more detailed studies including mountainous areas.

A study conducted by Morris et. al. [1994] in the United States found a clear relationship between respiratory hospital admissions and dwelling crowdedness. Also in Podkarpackie Province housing conditions factor was discovered to exert on the higher number of respiratory hospital admissions in rural areas. Social development factor, which included marital status, household size and education, turned out to be correlated with a in decrease hospital admissions. Interestingly, correlations produced by social development factor, whether significant or not, always regarded children or the elderly. A good review of socioeconomic factors influencing health may be found in Judge et. al. [1998] and Haas [2008]. Differently, family relations factor did not draw a parallel with hospitalization rates.

While one of the biggest limitations of this study is a lack of cross-sectional data about lifestyle, aforementioned social factors might have included some variations of everyday life. For example, a high alcohol use might be a significant health determinant especially in the southern and southern-east part of Podkarpackie Province. Low economic status and an easy access to cheap liquor (Ukraine) may cause frequent drinking to aggravate health outcomes. Similarly,

such environmental factor as Chernobyl disaster might have undermined population health in the eastern borderland municipalities. A lack of data about population past experiences is a common problem called latency period (Schærström [1999]).

Finally, an interesting finding was a relationship between respiratory hospitalization rates for people in 15-34 age group and employment in services. This confirms a previous finding that determinants of hospital admissions are significantly different in suburban areas of Podkarpackie Province from the cities and rural areas (Kretowicz [2010]).

Canonical analysis turned out to be a powerful tool, which revealed lots of possible health determinants on the regional scale. One of the biggest advantage of this method was numerous relationships found between the examined variables simultaneously. This is particularly important in very complex phenomena like health, which varies not only spatially but along with age, gender and types of diseases. The data used to explain health differences are often even more compound, thus a canonical analysis managed to simplify multilevel interpretation and highlighted only the most important associations. However, factor analysis method was also utilized in this study in order to reduce the number of variables – a loss in extracted variance limited further findings. Such problem is not seen in case of canonical analysis. Additionally, factor analysis helped to keep factors from correlating with each other. Multicollinearity was widely reported as a serious problem in canonical analysis (Clark [1975]). In this research, such condition was not fully fulfilled in criteria variables set, thus sometimes more than one of them correlated with predictor variables.

Although it is not always possible to find cause-effect relationships, a complete set of data would hopefully improve total redundancy. A lack of predictor data is the biggest problem to fully understand health disparities, which brings up a need to search through many sources. Moreover, complex calculations require computer programs. Fortunately, canonical analysis is available in most statistical software.

A positive feature of canonical analysis is a possibility to clearly visualize interdependence which makes it feasible to distinguish regions of different factors influencing various health outcomes In spite of this, disadvantageous interpretability of canonical scores may be meaningless for non-specialists. There is a difficulty with canonical analysis for the low number of cases while a very minimum condition is 3 cases per one variable. Such condition was nearly met in case of urban areas of Podkarpackie Province. Finally, a good idea would be to use canonical analysis before utilizing some local statistics such as Geographically Weighted Regression (GWR) (Fotheringham et al. [2002]) - a comparative studies might reveal basic differences between these two methods.

6. CONCLUSIONS

Canonical analysis technique with spatial approach may be a reliable statistical tool to be used in regional health planning. Apart from detecting some spatial disparities in hospital morbidity, which is a frequent practice in public health documents, a preliminary recognition of possible determinants for health inequalities come to the fore. As previously stated, complete datasets should be collected with a conscious and precise selection of variables used in the analysis. Canonical analysis would not offer accurate explanations of health disparities as cause-effect results are not always obtainable. Nevertheless, this technique should be valued for its simplification and clarity in presenting both time and spatial interrelationships. This requires a great deal of generalization together with disaggregation of data to some, selected extent. Agreeably with Laessig and Duckett [1979] such multivariate approach provide insights to environmental health with a possible further utilization in epidemiological research. As a consequence, canonical analysis may be performed as a prerequisite, carried out before more scrutinized investigation of contextual and compositional health determinants. Therefore, regional scale ought to be the smallest area to be examined in such studies since canonical analysis often hides local area effects. To conclude, more attention should be paid to multivariate spatial analysis in public health, which undoubtedly would help to outline main health issues to be solved or devoted to more detailed examination by proper governmental institutions.

REFERENCES

Beckfield, J., [2004], *Does income inequality harm health? New cross-national evidence*, Journal of Health and Social Behavior, 45, 3, pp.231–248.

Berry B.J.L., [1966]*, Essays on commodity flow and the spatial structure of the Indian economy*, University of Chicago, Department of Geography, Research Paper 111.

Breckenhamp J., Mielck A., Razum O., [2007], *Health inequalities in Germany: do regional-level variables explain differentials in cardiovascular risk?*, BMC Public Health 7, 132.

Chaix B., Merlo J., Chauvin P., [2005], *Comparison of a spatial approach with the multilevel approach for investigating place effects on health: the example of healthcare utilization in France*, Journal of Epidemiology & Community Health, 59, pp.517-526.

Clark D., [1975], *Understanding canonical analysis*, CatMog 3, Concepts and techniques in modern geography 3.

Coshall J.T., Potter R.B., [1987], *A nonparametric approach to canonical correlation in geography*, Geografska Annaler, 69 B, pp.127-132.

Curtis, S., Rees Jones, I., [1998], *Is there a place for geography in the analysis of health inequality?* [in:] Bartley M., Blane D., Davey Smith G. (eds.), *The sociology of health inequalities*. Oxford: Blackwell Publishers.

Diez-Roux, A., [2002], *A glossary for multilevel analysis*, Journal of Epidemiological and Community Health, 56, pp. 588–594.

Drummer T. J. B., 2008, *Health geography: supporting public health policy and planning,* CMAJ, 178, 9, p. 1177-1180.

Farr B.M., Bartlett C.L.R., Wadsworth J., Miller, D.L., [2000], *Risk factors for community-acquired pneumonia diagnosed upon hospital admission*, Respiratory Medicine, 94, pp.954–963.

Fotheringham A.S, Charlton M., Brunsdon M., [2002], *Geographically Weighted Regression*, Wiley & Sons Ltd., West Sussex.

Judge K., Mulligan J., Benzeval M., [1998], *Income inequality and population health,* Social Science and Medicine*,* 46, pp.567-579,

Haas S., 2008, *Trajectories of functional health: The 'long arm' of childhood health and socioeconomic factors*, Social Science and Medicine, 66, pp.849-861.

Hancook T., [2001], *People, partnerships and human progress: building community capital*, Health Promotion International 16, 3, pp.275-280.

Hotelling H., [1936], *Relations between two sets of variants*, *Biometrika*, 28, pp.321-377.

Laessig R.E. Duckett E.J., *Canonical correlation analysis: potential for environmental health planning,* American Journal of Public Health, 69, 4, pp.353-359.

Li J., Newcomb P., [2009], *Disparities in childhood asthma hospitalizations: A spatial analysis of contextual effects*, Transportation Research, Part D, 14, pp.317–325.

Kaczmarek Z., Parysek J.J., [1977], *Zastosowanie analizy wielowymiarowej w badaniach geograficzno-ekonomicznych*, [in:] Ciszewska B. (ed.), *Metody i modele ilościowe w geografii*, Państwowe Wydawnictwo Naukowe, Warszawa, pp.94-127.

Kretowicz P., [2010], *The influence of socio-economic factors upon public health on the example of Podkarpackie Province*, Bulletin of Geography, Socio-Economic series (in press).

Macintyre S., Ellaway A., Cummins S., [2002], *Place effects on health: how can we conceptualise, operationalise and measure them?*, Social Science and Medicine 55, pp.125-139.

Malmstrom M., Sundquist J., Johansson S.E., [1999], *Neighborhood environment and self-reported health status: A multilevel analysis*, American Journal of Public Health, 89, 8, pp.1181–1186.

Monmonier M. S., Finn F. E., [1973], *Improving the interpretation of geographical canonical correlation models*, The Professional Geographer, 25, 2, pp.140-142.

Morris R.D., Munasinghe R.L., [1994], *Geographical variability in hospital admission rates for respiratory disease among the elderly in the United States*, Chest, 106, 4, pp.1172-1181.

Norcliffe G.B., [1972], *Canonical analysis of the relations between certain aspects of the demographic and urban systems of the Republic of Ireland*, Irish Geography, 6, pp.411-427.

Payne J. N., Coy J., Milner P. C., Patterson S., [1993], *Are deprivation indicators a proxy for morbidity? A comparison of the prevalence of arthritis, depression, dyspepsia, obesity and respiratory symptoms with unemployment rates and Jarman scores*, Journal of Public Health Medicine, 15, 2, pp.161–170.

Pickett K.E., Pearl M., [2001], *Multi-level analyses of neighborhood socioeconomic context and health outcomes: A critical review*. Journal of Epidemiology Community Health, 55, 2, pp.111–122.

Potrykowska A., [1983], *Współzależności między dojazdami do pracy a strukturą społeczną i demograficzną regionu miejskiego Warszawy w latach 1950-1973*, Dokumentacja Geograficzna IGiPZ PAN, z. 2.

Ministry of Construction, [2007], *Polish Space – Country's Spatial Development Report*, Warsaw.

Ratajczak W., [1980], *Analiza i modele wpływu czynników społeczno-gospodarczych na kształtowanie się sieci transportowej*, PWN, Poznań.

Ray D.M., [1971], *From factorial to canonical ecology, the spatial interrelationships of economic and cultural differences in Canada*, Economic Geography, 47, pp.344-355.

Runge J., [2007], *Metody badań w geografii społeczno-ekonomicznej – elementy metodologii, wybrane narzędzia badawcze*, Wydawnictwo Uniwersytetu Śląskiego, Katowice.

Schærström A., [1999], *Apparent and actual disease landscapes. Some reflections on the geographical definition of health and disease*, Geografska Annaler 81, 4, pp.235-242.

Sullivan M, [2007], *Child with Asthma Living in a Moisture-Damaged Home*, [in:] Robson M.G., Toscano W.A. (eds.), *Risk assessment for environmental health*, John Wiley and Sons, pp. 534.

Runge J., [1991], *Dojazdy do pracy w przestrzennej strukturze powiązań miast województwa katowickiego*, Uniwersytet Śląski, Katowice.

Shaw M., Dorling D., Mitchell R., [2002], *Health, place and society*, Pearson, London, p.126.

Smith S., Easterlow D., [2005], *The strange geography of health inequalities*, Transactions of the Institute of British Geographers, 30, 2, pp.173-190.

Smyth F., 2008, *Medical geography: understanding health inequalities*, Progress in Human Geography, 32, pp.119-127.

Snow J., [1855], *On the mode of communication of cholera*, 2nd ed. London, Churchill.

Stelianides S., Golmard J.L., Carbon C., Fantin B., [1999], *Influence of socioeconomic status on features and outcome of community-acquired pneumonia*, European Journal of Clinical Microbiology and Infectious Diseases, 18, 10, pp.704–708.

Townsend P., Davidson N., Whitehead M., [1988], *Inequalities in Health*, London, Penguin.

Wojtyniak B., Goryński P. (eds.), [2008], *Sytuacja zdrowotna ludności Polski*, Narodowy Instytut Zdrowia Publicznego - Państwowy Zakład Higieny, Warszawa.

Wróblewska W., [2002], *Women's health status in Poland in the transition to a market economy*, Social Science and Medicine 54, pp.707–726.

Zamelska M., [1980], *Wpływ uprzemysłowienia na procesy urbanizacyjne w regionie bydgoskim*, Dokumentacja Geograficzna IGiPZ PAN, z. 5.

WYKORZYSTANIE ANALIZY KANONICZNEJ W BADANIACH PRZESTRZENNEGO ZRÓŻNICOWANIA PRZYCZYN HOSPITALIZACJI NA POZIOMIE REGIONALNYM

Celem referatu jest pokazanie możliwości zastosowania analizy kanonicznej do oceny natężenia i przestrzennego zróżnicowania przyczyn wpływających na różne wartości współczynników hospitalizacji ludności w regionie. Na obszar badań wybrano województwo podkarpackie w podziale na gminy (w sumie 188 jednostek przestrzennych).

Główną przesłanką wyboru metody jest znaczne zróżnicowanie chorobowości szpitalnej wybranych grup chorób (układu krążenia i oddechowego) w zależności od wieku ludności w poszczególnych obszarach. W konsekwencji współczynniki hospitalizacji dla poszczególnych grup wieku przyjęto za zmienne zależne, natomiast zmienne objaśniające stanowią różnego rodzaju wskaźniki struktury demograficznej, poziomu i jakości życia ludności oraz niektóre wskaźniki ekonomiczne. Zmienne objaśniające zebrano w trzy grupy uśredniając wartości za lata 2006-2008: demograficzne, społeczne i ekonomiczne, które charakteryzują badane gminy. Obliczone zmienne kanoniczne o największym stopniu skorelowania z grupą zmiennych niezależnych przedstawiono na kartogramach uzyskując przestrzenny rozkład współzależności pomiędzy badanymi zjawiskami. Końcową część opracowania stanowi ocena wad i zalet analizy kanonicznej dla wyjaśniania przestrzennych nierówności chorobowości szpitalnej, a także możliwości pogłębienia analizy zróżnicowań poziomu zdrowia na poziomie regionalnym.