International Journal of Occupational Medicine and Environmental Health, Vol. 7, No 3, 245–255, 1994

SELECTED ASPECTS OF THE POPULATION HEALTH STATUS IN ECOLOGICAL HAZARD AREAS IN COMPARISON WITH ECOLOGICALLY "CLEAN" AREA. I. METHODOLOGY FOR ANALYSING HEALTH STATUS OF THE POPULATION LIVING IN ECOLOGICAL HAZARD AREAS BASED ON THE ROUTINE DATA COLLECTION SYSTEM

TADEUSZ DUTKIEWICZ, CZESŁAW ANDRYSZEK, JERZY KOŃCZALIK
and DARIUSZ RACHAŃSKI

Department of Environmental Health Hazards
The Nofer Institute of Occupational Medicine
Lodz, Poland

Key words: Spatial differentiation of health status – indices methods

Abstract. A proposal for an integrated system of analysis of complex health problems in ecological hazard areas in comparison with reference least polluted area is presented. The proposed procedure can be used in identifying priorities in health promotion.

A progressive degradation of the natural environment observed for many years, low birth rate of the eighties, continuous deprivation of social and living conditions, unfavourable occupational changes as well as increasing difficulties in satisfying health needs of the society give rise to an urgent need for assessing how far those negative factors have affected the human health.

A number of studies on certain health problems have been undertaken and performed. They proved, however, to be methodologically incoherent due to a lack of uniform methods, proper monitoring based on relevant parameters and analyses of complex health problems in regard to the spatial variability. Studies on the geography of health status reveal that the degenerated natural environment exercises a significant effect on the human health (1, 11).

This paper aims at presenting a methodology for building up an integrated system which would allow to analyse complex health problems with special reference to demographic factors and adverse health effects observed in ecological hazard

Address reprint requests to Prof. T. Dutkiewicz, Department of Environmental Health Hazards, The Nofer Institute of Occupational Medicine, P.O. Box 199, 90-950 Lodz, Poland.

areas in comparison with so called ecologically "clean" area. The selection of "clean" reference areas was based on the evaluation of routine data concerning dust and gas emissions, sludge of untreated municipal and industrial sewages, and the amount of collected industrial wastes in each evaluated area.

Health status is a term of broad and complex meaning. Fig. 1 shows the most general scheme of the health problem complexity.

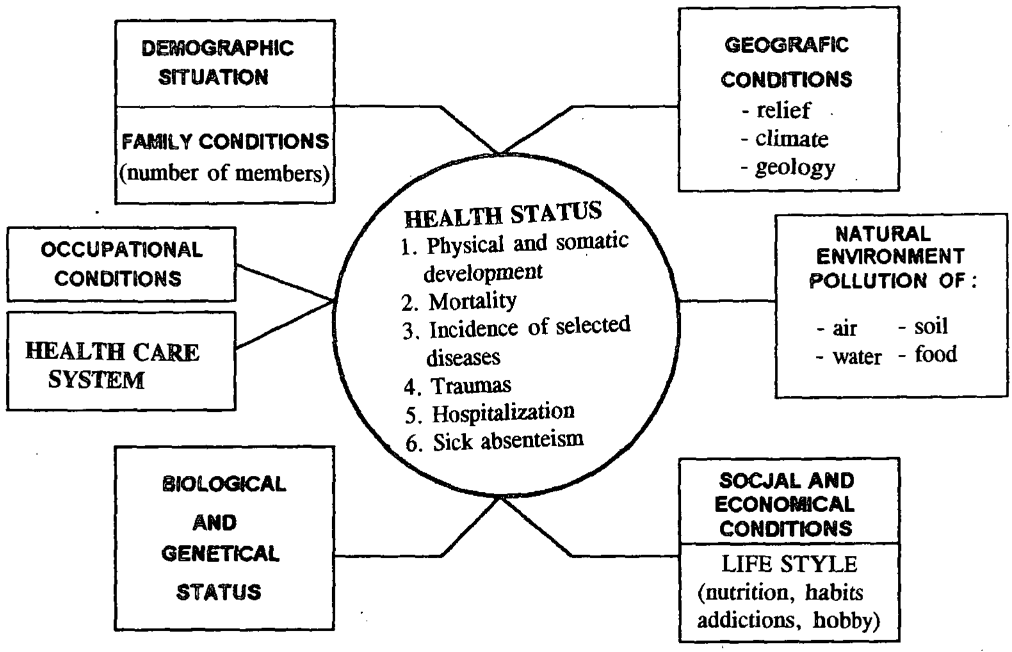

Fig. 1. Health status

This paper is focused on two basic elements presented in Fig. 1, namely: the health status and characteristics of the natural environment.

The health status as well as other complex health problems at spatial resolution are evaluated, first of all, on the basis of routine data collection system and immediate investigations of certain issues whenever necessary. Bearing in mind a progressive degradation of the natural environment, 27 ecological hazard (36 thousand km^2) areas were identified in Poland during the mid eighties (10). These areas are inhabited by one third of the population. Taking as a basis the variability of ecologically hazardous areas the attention was focused on the following:

1. Analysis and selection of relevant parameters of health status and characteristics of the natural environment;

2. Analysis of distribution of individual parameters in ecological hazard areas in comparison with reference area;

3. Development of combined indices for the health status evaluation as well as for other complex phenomena (demographic factors, natural environment) and determination of interrelationships between ecological hazard areas and reference area, based on identified parameters.

General mortality and mortality from major causes of death, infant mortality and an average life-span are most common parameters of health status of the population which are analysed on the basis of routine data collection system. It can be assumed after Kirschner (8) that spatial changes in these parameters caused by environmental hazardous factor occur at relatively distant time. Therefore, the methodology for identifying the impact of a progressive degradation of the natural environment is extremely difficult. One has to be very careful and critical in the analysis of changes and interpretation of data as there are many confounding factors. For example, smoking is one of such factors which may occur in the analysis of the effect of air pollution on the incidence of lung cancer.

Ecological hazard areas (Fig. 2) comprise urban and rural regions and the database of demographic and health factors are aggregated accordingly at spatial resolution.

In order to charactersie demographic factors, the following data were taken into consideration: population density, population structure according to sex, age, place of domicile, employment in basic branches of the national economy and essential parameters of population dynamics.

When assessing the health status at spatial resolution such negative parameters as birth rate of low body weight infants (below 2500 g), general infant mortality and that caused by low body weight at birth, early and late mortality of infants as well as selected causes of death were considered.

Mortality from diseases (general mortality rate excluding external causes of deaths – injuries and poisonings), mortality in selected age groups and mortality due to the most common environment-related causes were also analysed.

Spatial units were aggregated in the following subcategories:

1. Urban regions in ecological hazard areas,
2. Rural regions in ecological hazard areas,
3. Urban regions in the reference area,
4. Rural regions in the reference area.

The reference area was selected from voivodships* free from ecological hazard areas and where the analysis of the natural environment pollution proved to be most favourable. They comprise the following voivodships: Białystok, Biała Podlaska, Ciechanów, Krosno, Łomża, Siedlce i Zamość. For the purpose of spatial analysis rates of mortality from diseases in general, and from a particular cause were standardised using a direct method (1, 15).

Finally each spatial unit was described by the vector of relative parameters (indices, rates).

EVALUATION OF DISTRIBUTION OF MORTALITY AND OTHER CHANGES

As depicted above, an urban area or its district and a rural area which form larger ecological hazard areas or reference areas are considered as the smallest spatial units

The following methodology was adopted in the analysis of spatial distribution of mortality when comparing ecological hazard areas and the reference area.

*Region of the administrative division of the country.

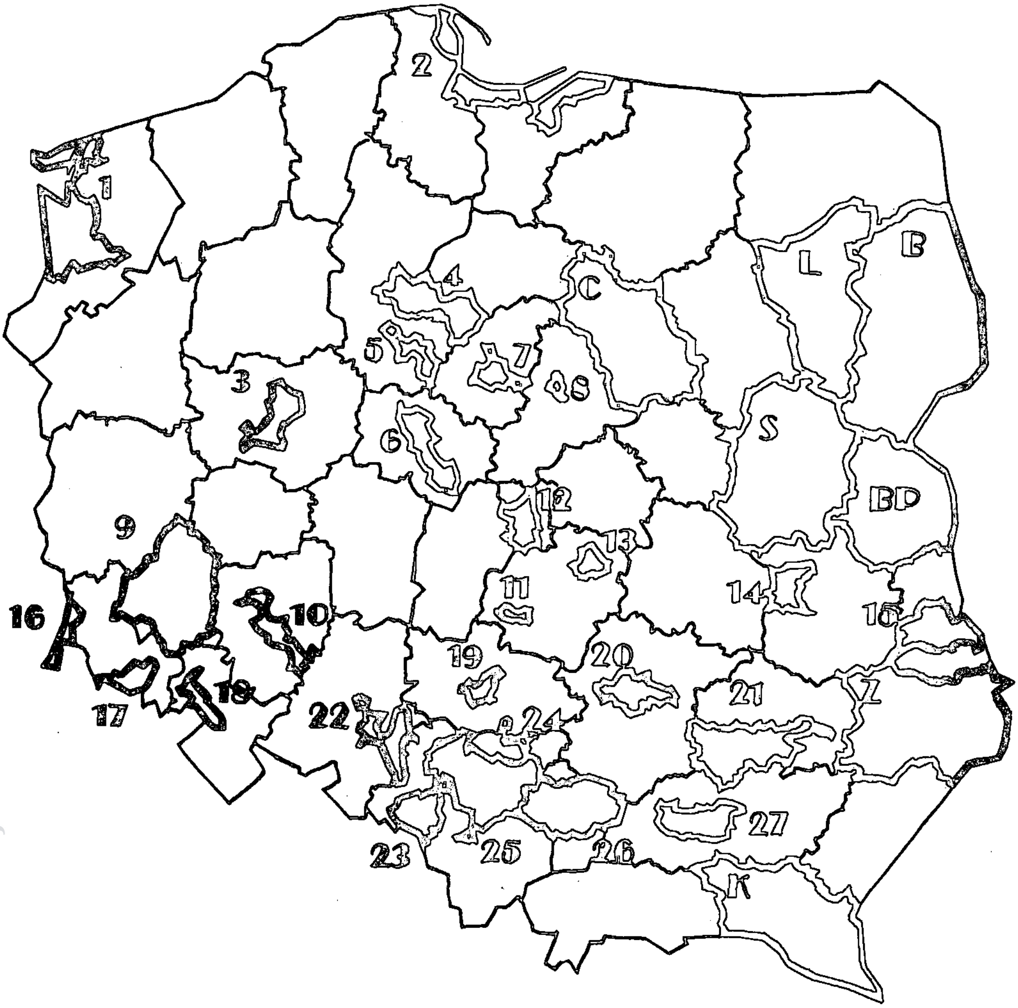

Fig. 2. Spatial distribution of ecological hazard areas and reference area in the map of Poland.

Reference area — voivodships:
B — Białystok
BP — Biała Podlaska
C — Ciechanów
K — Krosno
L — Łomża
S — Siedlce
Z — Zamość

Ecological hazard areas:
1. Szczecin
2. Gdańsk
3. Poznań
4. Bydgoszcz-Toruń
5. Inowrocław
6. Konin
7. Włocławek
8. Płock
9. Legnica—Głogów
10. Wrocław
11. Bełchatów
12. Łódź
13. Tomaszów
14. Puławy
15. Chełm
16. Turoszów
17. Jelenia Góra
18. Wałbrzych
19. Częstochowa
20. White Basine
21. Tarnobrzeg
22. Opole
23. Rybnik
24. Myszków—Zawiercie
25. Upper Silesia
26. Kraków
27. Tarnów

A computer analysis of the index variability provided interval limits of the distribution. Thus, empirical distributions acquired the form as presented below:

Intervals (about 10)	Ecological hazard areas		Reference area	
	n_j	P_j	n_j	P_j

where: n_j – number of town districts or towns or counties, and
P_j – empirical probabilities (fractions).
The empirical distribution was analysed by means of the set of statistical parameters and the goodness of fit empirical distribution to normal one was calculated.

Arithmetic mean ($\bar{x}$), geometric mean ($\bar{x}_g$), other parameters of central tendency such as mode (Mo), and median (Me) close to a geometric mean ($\bar{X}_g$) as well as quartiles (Q_1 and Q_3) were used in assessing the average mortality rate.

Standard deviation (SD), coefficient of variation $V\% = \frac{SD}{\bar{x}} \cdot 100$ and quartile deviation $Q = \frac{1}{2}(Q_3 - Q_1)$ were calculated in order to evaluate spatial variability of parameters in question.

The following was used in assessing the distribution:

1. Coefficient of asymmetry $As_1 = \frac{(\bar{x} - Mo)}{SD}$,

2. Coefficient of asymmetry based on quartiles:

$$As_2 = \frac{(Q_3 - Q_2)}{(Q_3 - Q_1)},$$

3. Exact coefficient of skewness based on the third central moment, and
4. Coefficient of kurtosis based on the fourth central moment.

In order to check whether the spatial distribution of mortality can be considered as the normal one:

1. Factions and theoretical numbers were calculated assuming the hypothesis on the normal distribution, and

2. Goodness of fit of empirical distribution to theoretical distribution was analysed applying various criteria:
– the Chi^2 test,
– the Komogorow-Smirnow λ test.

As the statistics of the Chi^2 test depends directly on the sample size it often happens that when analysing distributions in large samples (range of thousands) empirical and theoretical distributions are very similar, however, the Chi^2 test rejects the hypothesis on their agreement. Therefore, it is useful to use certain parameters based on relative numbers in order to assess the distribution agrrement.

Bearing this in mid the following parameters were additionally considered:
1. Standard deviation of differences of relative numbers (S),
2. Maximum difference in relative numbers (V_{max}),
3. Coefficient of determination according to the formula:

$$R = 1 - \frac{\sum_{j=1}^{k} (f_j - g_j)^2}{\sum_{j=1}^{k} \left(f_j - \frac{1}{k}\right)^2},$$

where: f_j — empirical relative number, and
g_i — theoretical relative number

If S and V_{max} values are below 0.05 and R^2 value exceeds 0.95 then the compared distributions can be considered as highly compatible.

Following the comparative analyses of distribution and their parameters, ecological hazard areas and reference area were arranged in the list, and urban and rural regions of particularly high values of standardised mortality rates, were identified.

This phase of the study was concluded with the assessment of mortality in towns considered as a potential source of factors responsible for ecological hazard.

The question arises how almost direct contact with hazardous factors is reflected in late health effects expressed in mortality rates.

THE DEVELOPMENT OF COMPREHENSIVE PARAMETERS OF THE HEALTH STATUS, THE CONDITION OF NATURAL ENVIRONMENT AND THE EVALUATION OF RELATIONSHIPS BASED ON THOSE PARAMETERS IN COMPARISON WITH ECOLOGICAL HAZARD AREAS AND REFERENCE AREA

The collected material in database was used for drafting a list of parameters for a comprehensive evaluation of identified complex phenomena.

Finally matrices of parameters were developed as follows:

$$Y_r = \begin{bmatrix} Y_{11}, Y_{12} \ldots\ldots\ldots\ldots\ldots Y_{1m} \\ Y_{21}, Y_{22} \ldots\ldots\ldots\ldots\ldots Y_{2m} \\ Y_{27,1}, Y_{27.2} \ldots\ldots Y_{27.m} \\ Y_{28.1}, Y_{28.2} \ldots\ldots Y_{28m} \\ Y_{36.1}, Y_{36.2} \ldots\ldots Y_{36.m} \end{bmatrix} \tag{1}$$

where: m — means number of parameters, number of rows represents 27 ecological hazard areas and the reference area considered as ecologically "clean" one.

In the literature one can find different approaches to aggregation of complex phenomena. Among taxometric measures proposed by Cieślak (5), Bartosiewicz (2), Pluta (14), Dittman (6), Borys (4), Hellwing (9), Strahl (17), the two latest seem to be most suitable for evaluating complex phenomena of the health status (7, 8). The majority of proposed parameters is based on the so called "development model". Hellwig's and Strahl's measures identify three types of variables in the matrix of variables: stimulants, destimulants and nominants.

A stimulant is a variable whose higher values are attributerd to better condition of an object in view of the measured phenomenon. In the evaluation of health status, indices of resources are examples of stimulants, e.g. number of doctors per 10 000 population, number of hospital beds per 10 000 population, etc. In the Polish national economy, the national income is a basic stimulant.

A destimulant is a variable whose lower values mean better condition of an object in view of the measured phenomenon. Such rates as mortality, infant mortality, incidence, etc. belong to the category of destimulants.

A nominant is a variable for which a certain value considered to be optimal (neither higher nor lower) defines the best condition. Females to males ratio, natural increase rate, bed occupancy rate, etc. are good examples of nominants.

In the process of developing a combined index the standardisation of variables occurring in the data matrix creates a certain problem. Various approaches to standardisation are suggested by Borys (3).

Generally the method of stndardisation of variables is most common in the statistical analysis. According to Hellwig's measure (9), standardisation of variables should follow a traditional procedure namely an arithmetic mean should be deducted from a variable value and then divided by a standard deviation:

$$Y'_{ij} = \frac{(Y_{ij} - \bar{Y}_j)}{S_j} \tag{2}$$

where: i – spatial unit index,
j – variable index.

A traditional method for the standardisation reduces all parameters to the interval (−3, +3) Strahl (17) proposes the following method for variable standardisation:

$$Y'_{ij} = \frac{Y_{ij}}{\max_k Y_{ij}} \tag{3}$$

when Y_{ij} is a stimulant,

$$Y'_{ij} = \frac{\min_k Y_{ij}}{Y_{ij}}$$

when Y_{ij} is a destimulant,

$$Y'_{ij} = \frac{Y_{ij}}{\mathrm{nom}_k Y_{ij}}$$

when Y_{ij} is a nominant, and $Y_{ij} \geqslant$ nom Y_{ij}

$$Y'_{ij} = \frac{\mathrm{nom}_k Y_{ij}}{Y_{ij}}$$

when Y_{ij} is a nominant and $Y_{ij} \geqslant$ nom Y_{ij}, then Y_{ij} variables are in the interval 0–1.

On the basis of the register of urban and rural regions which constitute ecological hazard areas and the reference area mean coefficients for those areas are calculated according to the following formula:

$$\bar{Y}_k = \frac{\Sigma m_i}{\Sigma L_i} \times C \text{ or } \bar{Y}_k = \frac{\Sigma^{Li} Y^i}{\Sigma^{Li}} \times C \tag{4}$$

where: $\bar{Y}_k$ – mean rate "k" ecological area,
m_i – absolute number of health events in "i" urban or rural region, eg. number of deaths,
L_1 – number of the population in "i" urban or rural region,

Y_i – crude rate if „i" urban or rural region,
C – constant used for calculating of the rate (in the analysis of mortality C = 10 000).

To sum up in selected urban and rural regions which form a given "k" area.

A calculated matrix of rates averaged for given areas was standardised taking into account mean age structure in areas under study.

The vector of values of the so called "development model" was identified for matrix of comparable rates.

$$W = (Y_{01}\, Y_{02} \ldots\ldots Y_{om}) \tag{5}$$

In developing the synthetic parameters according to Hellwig or Strahl it is suggested to adopt the following for the model values:

$Y_{oj} = \max_k Y_{kj}$ if the variable is a stimulant, (6)
$Y_{oj} = \min_{kj}$ if the variable is a destimulant,
$Y_{oj} = \mathrm{nom}_k\, Y_{kj}$ an optimal value adopted a *priori* if the variable is a nominant.

Categorisation has been done before. Estimation of a nominal value for nominant variables is most difficult; one can adopt a target value formulated in longterm programmes (e.g. the National Health Programme, formerly called the "Health Programme – Poland 2000"). If there are no other proposals, a national mean value can be adopted as a nominal value.

Both, the information matrix and the "development model" value are standardised according to the formula (2) or (3).

For the modified information matrix a synthetic parameter is calculated according to the formula:

$$C_{ko} = \left[\sum_{j=1}^{m} (Y'_{ij} - Y'_{oj})^2 \right]^{\frac{1}{2}} \tag{7}$$

The formula: $H_{ko} = 1 - \dfrac{C_{ko}}{C_o}$ (8)

is finally adopted as a synthetic measure according to Hellwig where:

$$C_o = C + 2S_o \tag{9}$$

$$C = \frac{1}{n} \sum_{k=1}^{n} C_{ko} \tag{9.1}$$

$$C = \frac{1}{n} \left[\sum_{k=1}^{n} (C_{ko} - C)^2 \right]^{\frac{1}{2}} \tag{9.2}$$

Provided that C_{ko} index has a normal distribution, H_{ko} parameter can be, with small probability (approx. 5%), outside the interval (0 – 1).

Values close to "0" provide evidence that the situation in a given area is infavourable in regard to a complex phenomenon presented by the set of coefficients used in the evaluation. The higher the value, the more favourable situation in the area presented by the set of variables – on average, values of individual variables are closer to values adopted as the so called "development model".

A proposed modification of a synthetic parameter according to Hellwig depends on the diversification of the importance of variables used in the analysis. Such diversification would mean introduction of weights of individual variables. In the evaluation of health status, infant mortality represents a different weight (higher) than mortality (or incidence) from a relatively rare disease. Thus, C_{ko} value will be calculated according to the following formula:

$$C_{ko} = \left[\sum_{j=1}^{m} W_j (Y'_{kj} - Y'_{oj})^2 \right]^{\frac{1}{2}} \tag{10}$$

where: W_j — means weight of individual variables.

For a region under consideration a synthetic parameter according to Strahl is an arithmetic mean of variable values stnadardised according to the formula:

$$S_k = \frac{1}{n} \sum_{k=1}^{n} Y'_{kj}$$

Finally Hellwig's or Strahl's parameters for a comprehensive evaluation of the health status will be expressed in the form of an index. Mean values H or S and a standard deviation S_H or S_S will be calculated for a set of synthetic parameters of all ecological hazard areas. The index values will be assigned in the following way:

for $H > H + \frac{1}{2} S_H$ index 1 – the most favourable,

for $H - \frac{1}{2} S_H < H < + \frac{1}{2} S_H$ index 2 – intermediate health status,

for $H < H - \frac{1}{2} S_H$ index 3 – unfavourable health status.

Similar indexing will be used for calculation of Strahl's measure.

ESTIMATION OF EXCESS MORTALITY RELATED TO ENVIRONMENTAL HAZARDS

Calculation of age adjusted mortality rates in urban and rural regions which belong to ecological hazard areas and in the selected reference area on the basis of routine information system is a starting point. For each "i" area and "t" time two figures of deaths are calculated (12, 13).

Z_{ob} – number of deaths observed provided that the age structure of the population in ecological hazard areas and in the reference area is the same and correspondents with a standard structure (mean structure for the whole country).

$$1.\ Z_{ob} = L_1 \times P_{obi} \tag{11}$$

where: Z_{obi} – number of deaths according to standardised mortality rate,
L_1 – number of population in "i" area,
P_{ob1} – age adjusted mortality rate in "i" area.
Z_{exp} – number of deaths expected provided that mortality in "i" ecological hazard area will be the same as the age adjusted mortality in the reference area.

$$2.\ Z_{exp\,i} = L_1 \times P_c \tag{12}$$

where: P_c – average age adjusted rate in the reference area.

The difference between $Z_{ob\,1}$ and $Z_{exp\,1}$ is an estimated additional number of deaths in "i" area related to exposure to harmful factors present in the area in question.

A relative number $I_j = \frac{Z_{ob.1}}{Z_{exp.1}}$ could be adopąted as a ratio of excess mortality (relative mortality risk) resulting from exposure to harmful factors present in a given area in relation to environmental factors observed in the reference area.

CONCLUSIONS

An ecological approach to the evaluation of spatial variables of the health status assumes that the distribution of incidence of many diseases depends, to a great extent, on environmental factors. Thus far, a methodological coherence has not yet been reached in spatial analyses of the complex health condition.

This paper presents a method for a comprehensive analysis of these phenomena present in ecological hazard areas in comparison with the area considered to be ecologically "clean".

The following was proposed and presented:

1. The method for the analysis of spatial distributions of individual parameters.
2. Identification of spatial units (urban areas, their districts and rural areas) of particularly unfavourable (high) values of health status parameters under study.
3. The method for the development of synthetic measure for evaluating health status of the population living in ecological hazard areas and in the reference area.
4. A proposal for index assessment of the health status in ecological hazard areas and in the reference area.
5. A proposal for estimating the environment-related excess mortality in areas of ecological hazard (relative mortality risk).

The procedures proposed will provide a uniform approach to evaluating the environment and identifying health effects in the population living in given areas. In opinion of the authors the presented methodology seems to be appropriate and useful for setting up priorities in the field of health and environment promotion.

REFERENCES

1. Andryszek C. Spatial variability of the population health status. Theory and practice of the health care organization. Bulletin of theory and practice of care health organisation: IMP, 1991 (in Polish).
2. Bartosiewicz S. A proposal for developing synthetic variables. Scientific papers of the Academy of Economics: Wroclaw, 1976 (in Polish).

3. Borys T. Standardization methods in statistical comparative studies. Przegląd Statystyczny 25: 227–232, 1978 (in Polish).
4. Borys T. A proposal for an aggregative measurement of object developments. Przegląd Statystyczny 25: 371–382, 1978 (in Polish).
5. Cieślak M, Jasiński E. A global index of health status. Studia Demograficzne 45: 91–96, 1976 (in Polish).
6. Dittman P, Pisz Z. A method for dynamic investigation of spatial variability of social and economic phenomena. Wiadomości Statystyczne 20: 27–36, 1975 (in Polish).
7. Hellwig Z. Application of the taxonometric method for the typological division of the country in view of its development, resources and qualified personnel. Przegląd Statystyczny 15; 307–327, 1968 (in Polish).
8. Kirschner H, Tyszko P. Monitoring of health status in the population living in regions of ecological hazard. Methodology. Bulletin of the Commission on Environmental Impact Assessment: 1991 (in Polish).
9. Pluta W. A taxometric procedure of performing synthetic comparative studies using a modified parameter of the economic development. Przegląd Statystyczny 22; 511–516, 1976 (in Polish).
10. Sawicki F. Statistics for Physicians. PZWL; Warsaw, 1982 (in Polish).
11. Strahl D. Models for complex phenomena. Models of social infrastructure. A dissertation. Scientific Papers of the Academy of Economics; Wroclaw 1980 (in Polish).

Received for publication June 23, 1993
Approved for publication: June 19, 1994