Biocybernetics and Biomedical Engineering
2009, Volume 29, Number 2, pp. 59–71

Grade Approach to the Analysis of Questionnaires and Clinical Scales Data

ANNA WOLIŃSKA-WELCZ[1],*, HENRYK WELCZ[2]

[1] *Maria Curie-Skłodowska University, Institute of Mathematics, Lublin, Poland*
[2] *Chair and Department of Psychiatry, Medical University, Lublin, Poland*

Grade approach has been applied to visualize and reinvestigate a dataset from diagnostic questionnaires and scales used in psychological and psychiatric clinical practice. Grade exploration reveals important hidden variables strongly influencing other variables and splits questionnaire data into more regular monotone dependent blocks. Two sets of homogeneous and ordered clusters of patients, formed on the basis of first line variables and described also by second line variables, expose links between neurosis symptoms and Internal or External Locus of Control as well as prove useful for specifying and verifying diagnoses. All analyses were performed using the GradeStat program.

Keywords: data exploration, grade clustering, symptom check-list, neurotic disorders, Rotter scale

1. Introduction

The aim of our research was to analyze the data gathered from over 100 patients qualified for therapeutic treatment in the Department of Neuroses in the Psychiatric Clinic of Lublin (Poland). The patients' records contained the measurements of the "O" Symptom Check list of Aleksandrowicz [1], which consists of 13 subscales reflecting the intensity of various neurotic symptoms, the Rotter Internal-External Control Scale [2] and the Neurotic Disorders Scale of Bizoń. The remaining information concerns medical classifications and a few demographic variables.

The classic analysis of these data was performed in search of links between external or internal controllability and intensity of various neurotic symptoms in [3]. Now we present the results of the application of grade exploratory methods described in

* Correspondence to: Anna Wolińska-Welcz, Institute of Mathematics, Maria Curie-Skłodowska University, Plac M. Curie-Skłodowska 1, 20-031 Lublin, Poland, e-mail: welcz@hektor.umcs.lublin.pl
Received 29 September 2008; Accepted 05 February 2009

[4] to visualizations and clustering of this data set. The grade methods implemented in the GradeStat program have been recently applied to various medical data (cf.[5]) and used in numerous papers listed in [6].

The particular grade procedures are applied step by step according to data analyst's choices. The first choice met in case of the considered clinical data consisted in the splitting of the variables under observation into the first line variables (clinical scales and subscales) and the second line variables (demographic variables, medical classifications etc.).

The two way table with patients as rows and first line variables as columns is then transformed into a non-negative function on the unit square which could be formally treated as a density of bivariate distribution with uniform marginals, constant on rectangles at an intersections of vertical strips (corresponding to variables) and horizontal strips (corresponding to patients). The graph of this *grade density* can be presented as a map called *overrepresentation map*.

The strength of positive dependence between patients and variables is measured by well-known parameters: Spearman correlation, denoted ρ*, or Kendall tau, denoted τ. Particularly important among grade methods are Grade Correspondence Analyses denoted GCA, which serve to permute rows and columns in a two-way probability table to achieve the table with maximal values of ρ* or τ denoted $\rho^*_{\max}$ and $\tau_{\max}$. There is GCA_S maximizing ρ* and GCA_K maximizing τ (in most cases the post-GCA table is the same for GCA_S and GCA_K, or there are very slight differences). Details concerning GCA_S and GCA_K are summarized in [7].

In practical terms, GCA provides the strongest match between patients and variables: the patients placed by GCA in the upper rows tend to achieve large values of density in the columns placed at the left side of the map and small values in the columns at the right side, and this tendency gradually changes for next patients to become finally reversed for the patients placed by GCA in the bottom rows.

That tendency characterizes densities of positive dependent bivariate distributions. All such distributions have been tried to be ordered in probabilistic literature according to how regular this tendency is. The high level of regularity is acquired in a cases/variables table when, for any chosen ordered pair of rows accompanied by any chosen ordered pair of columns, determinants in the corresponding 2×2 subtable are all non-negative. Distributions with this property are called totally positive dependent of order 2 (briefly TP_2). The departure from the TP_2 property was used to construct the index of regularity basing on the fact that for any $m \times k$ cases/variables table transformed into a probability table $P_{m\times k} = [p_{ij}, i = 1, 2,\ldots, m, j = 1, 2,\ldots, k]$ τ can be expressed as a double sum of determinants in 2×2 subtables (cf. [4]):

$$\tau\left(\mathrm{P}_{\mathrm{mxk}}\right) = 2\sum_{r=2}^{m}\sum_{i=1}^{r-1}\sum_{s=2}^{k}\sum_{j=1}^{s-1}\left(p_{ij}p_{rs} - p_{is}p_{rj}\right). \qquad (1)$$

Thus the upper bound of τ, denoted τ_{abs}, is obtained when the determinants in all subtables are replaced by their absolute values. It follows that $\tau = \tau_{\mathrm{abs}}$ holds if and only if the table is TP_2. Therefore, the regularity level can be measured by τ/τ_{abs}. It is the regularity level of positive dependence if $\tau > 0$ and of negative dependence if $\tau < 0$, so it is often said to be the regularity level of monotone dependence.

With regularity measured in this way it is evident that its value is maximal when $\tau = \tau_{max}$. It happens sometimes that the post-GCA table becomes TP_2, so that $\tau_{max}/\tau_{abs} = 1$, otherwise the table or tables with $\tau = \tau_{max}$ are “as close to the TP_2 table as possible”, at least under so defined regularity measure.

When the regularity of monotone dependence in the TOTAL table transformed by GCA is poor (as happened for the table with the first line variables attached to all patients, as visualized in Section 2), it is advisable to exclude those patients who spoil regularity and put them into a subset of outliers, called OUT; the remaining patients will form a subset called FIT. Then, the GCA procedure is applied to FIT and OUT separately. The regularity level in the post-GCA FIT increases from that in TOTAL, although the orderings of columns and rows remain the same as in TOTAL or are only slightly changed. In practice it often happens that the post-GCA OUT subset also acquires a certain plausible level of regularity. This is observed in the post-GCA FIT and OUT subsets visualized in Section 3. It is worth noting that the level of regularity used here depends on the table size (like most parameters including ρ^*, ρ^*_{max}, τ, τ_{max}) and, therefore, direct comparisons have to be made carefully.

The orderings of variables in FIT and OUT are usually different: in particular, other variables appear in some of the leading positions on the left and on the right side of the post-GCA maps. As illustrated in Section 4, the comparison between the orderings provides an important insight into the model of patients’/variables’ dependence.

However, it should be stressed that it is not easy to make a sensible division of the TOTAL set of patients into FIT and OUT. Usually, it is reached after many trials done according to data analyst’s choices. Each choice leads to different subsets FIT and OUT that are investigated whether they are regular enough and whether they can be sensibly interpreted.

2. Description and Visualization of First Line Clinical Data

The necessary data cleaning reduced the number of patients to 80 (the records of improperly qualified patients and those with a too small sum of neurotic symptoms were excluded).

The first line variables are designed for specifying the structure and the ordering of clusters. It means that the grade clusters are entirely formed on the basis of the data from clinical scales and questionnaires, while the second line variables provide an additional description.

The names of the first line variables are:

- *I-E Rotter scale* (points from 0 to 23, with low values indicating internal Locus of Control and high values indicating external Locus of Control);
- *Bizoń scale* (these values are linearly transformed to be positive, higher values are believed to reflect stronger neurotic disorders);
- 13 subscales in the Aleksandrowicz questionnaire: *fear, depression, anxiety, hypochondria, somatic symptoms, hysteria, neurasthenia, psychasthenia, derealization, compulsions, sexual disorders, sleep disorders, social difficulties.*

In the present study it was decided that the values of 13 subscales of the Aleksandrowicz questionnaire were left without normalization and, therefore, corresponding columns have different widths (proportional to weight of subscale and to sums of points gathered by all patients for this subscale); so e.g. the subscale of *somatic symptoms* gained a slightly higher importance. Furthermore, the properly chosen weights were ascribed to the Bizoń and Rotter scales. According to the intuition and experience of the researcher, widths of Bizoń and Rotter scales were chosen to be 0.15 and 0.2.

The algorithm of Grade Correspondence Analysis – implemented in the GradeStat program [6] – reordered rows and columns of the patients/variables matrix for 15 first line variables. The effect is presented in Fig. 1 in the *overrepresentation map*.

The map consists of rectangles characterized by different shades of grey varying from white to black. Intensity of grey corresponds to positive numbers from the interval (0.6, 1.7), according to the scale of grey given at the right side of the map. Number 1 would appear for an rectangle at the intersection of a row (patient) and a column (variable) when the value of this variable for this patient is equal to the product of the row and column sums, divided by the sum of all terms in the matrix. Numbers that are smaller (or resp. bigger) than 1 appear when the observed value is *underrepresented* (or resp. *overrepresented*), i.e. smaller (or resp. higher) than expected on the assumption of perfect proportionality to the row and column totals.
More precisely, the 80×15 patients/variables table is transformed into a probability table $P_{80\times 15} = [p_{ij}]$ and the value of grade density in point (u, v) from the unit square is determined as follows

$$f(u,v) = \frac{p_{ij}}{p_{i\cdot} p_{\cdot j}}, \quad (u,v) \in R_{ij}, \quad R_{ij} = [S_{i-1}, S_i) \times [T_{j-1}, T_j) \tag{2}$$

where $S_i = p_{1\cdot} + \ldots + p_{i\cdot}$, $T_j = p_{\cdot 1} + \ldots + p_{\cdot j}$, $S_0 = T_0 = 0$, $i = 1,\ldots,80$, $j = 1,\ldots,15$.

The values $f(u, v)$ are mapped into a shade of gray according to the scale placed at the right side of the overrepresentation map.

The widths of rows and columns in Fig. 1 represent the values of the respective totals (marginal probabilities). The GCA reordering makes the intensity of grey follow the pattern of regular positive dependence. Consequently, intensity at the top of the map tends to be high for the variables from the left columns and small for the variables from the right columns. This tendency is reversed at the bottom of the map, and intensity proceeds from the top records to the bottom records as smoothly as possible.

We can observe on Fig. 1 that the ordering of the I-E Rotter values agrees with the GCA ordering which, however, has to compromise with the opposite ordering induced by the subscales of *the somatic symptoms, compulsions*, *hypochondria* and *derealization*. The records of patients with Locus of Control (LOC) classified as external (dark arrow-shaped markers) tend to concentrate at the bottom of the map, while the records of patients with LOC classified as internal (light rectangular markers) tend to concentrate in the upper part of the map. Apart from this regularity, the data seem to be rather irregular and their splitting is highly advisable.

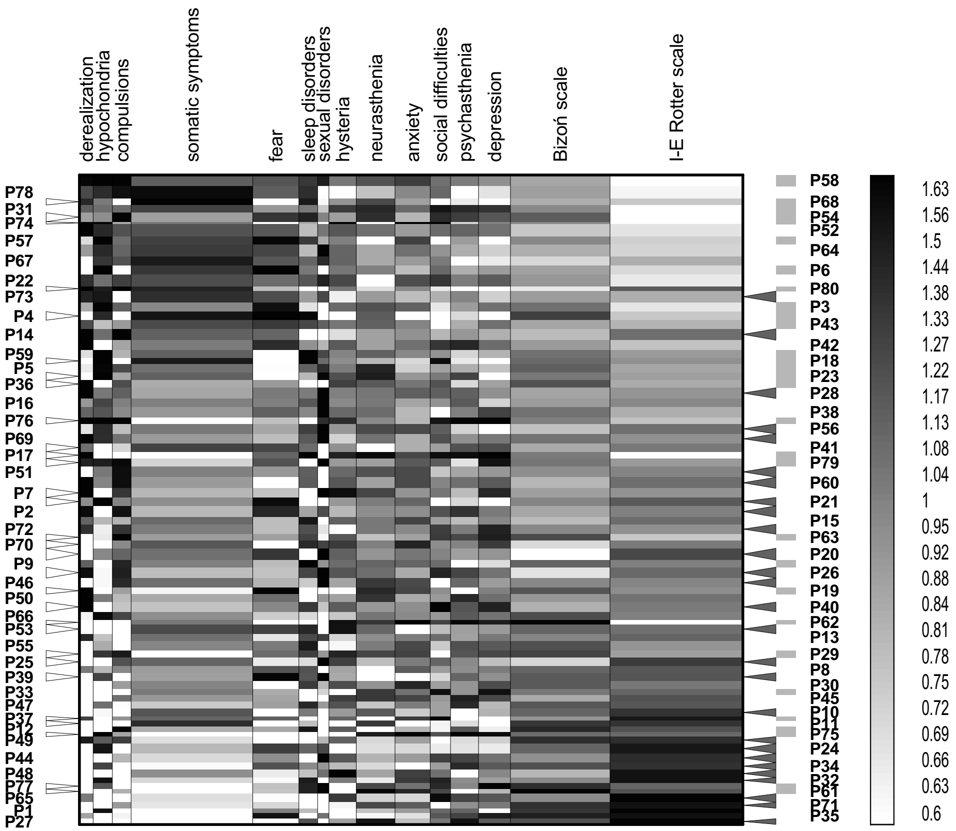

Fig. 1. The post-GCA overrepresentation map presenting patients/variables table for 15 first line variables and 80 patients (coded P1, P2, … P80 before GCA). White arrow-shaped markers at the left side of the map are attached to 30 patients most outlying from the positive dependence pattern. Additionally, at the right side of the map, dark arrow-shaped markers are attached to 27 patients with External Locus of Control (high values of the Rotter scale), and light rectangular markers are attached to 29 patients with Internal Locus of Control (low values of the Rotter scale). The values of grade density are determined according to the scale of grey given at the right side of the map. The values of ρ* and τ are 0.165 and 0.110, the regularity index τ_{max}/τ_{abs} is rather poor (0.44)

3. Division of First Line Questionnaire Data into More Regular Monotone Dependent Blocks of Patients' Records

The splitting algorithm implemented in GradeStat selected 30 rows which distinctly and most strongly departed from regularity under the GCA ordering. They are marked by white arrows shown at the left side of the map in Fig. 1. The TOTAL set of 80 patients was divided into a subset of size 50×15 called FIT (collecting the patients indicated as better fitted to the initial GCA ordering in Fig. 1) and a subset of size 30×15 called OUT (collecting the patients who were found worse fitted to this ordering). Then GCA was done anew in each subset. In FIT the new GCA orderings for patients and variables became close to those in Fig. 1. In OUT the new GCA orderings were strongly different from those in Fig. 1.

As mentioned in Section 1, the number of patients in FIT and OUT were chosen using trial and error method. The respective grade procedures ordered the patients in TOTAL from the most to the least outlying from the post-GCA regularity and the data analyst chose how many patients are included to OUT.

The post-GCA overrepresentation maps for FIT and OUT are shown in Figures 2 and 3 respectively. The respective maps for FIT and OUT are on rectangles which, put one over another, form a square with the size identical with that in Fig. 1 (this allows the widths of rows in Figures 2 and 3 to remain the same as in Fig. 1). The strength of dependence in FIT and OUT is slightly higher than in the TOTAL data (ρ^* and τ are 0.185 and 0.123 in the post-GCA FIT, 0.202 and 0.135 in the post-GCA OUT); regularity index τ_{max} / τ_{abs} is 0.55 for FIT and 0.45 for OUT as compared to 0.44 in TOTAL. Visual observation confirms that FIT is more positively regular than the whole set of patients, while OUT remains similarly regular, and that the strength of dependence between the patients and the variables increases in both subsets as related to 0.165 for ρ^* and 0.110 for τ in TOTAL.

Although the regularity of FIT and especially of OUT is not quite satisfactory, it seems worth to divide each of these subsets of patients using GCCA (Grade Cluster Correspondence Algorithm implemented in the GradeStat program) into adjacent disjoint clusters, four in FIT and three in OUT. The numbers of clusters were chosen by a few trials. It follows that clusters in FIT as well as in OUT are ordered according to the GCA orderings in the both subsets. The analogous clustering is also performed for the variables. The horizontal and vertical lines that separate the clusters are shown in Figures 2 and 3. Thus, the patients' records and the variables profiles can be visually compared with those for other patients or variables within the same cluster of patients or variables.

Visual presentation of FIT and OUT subsets emphasizes the aggregation of values of each first line variable in particular clusters of these subsets. The overrepresentation maps for the resulting tables of the aggregated values are shown in Figures 4 and 5.

Moderately regular monotone dependence may be noticed in both maps. Strength of the monotone dependence measured by ρ^* is 0.18 in both FIT and OUT. The regularity indices τ_{max}/τ_{abs} after aggregation become much higher (0.93 in FIT and 0.95 in OUT) than before (0.55 in FIT and 0.45 in OUT). So the aggregation increases the regularity and diminishes the differentiation.

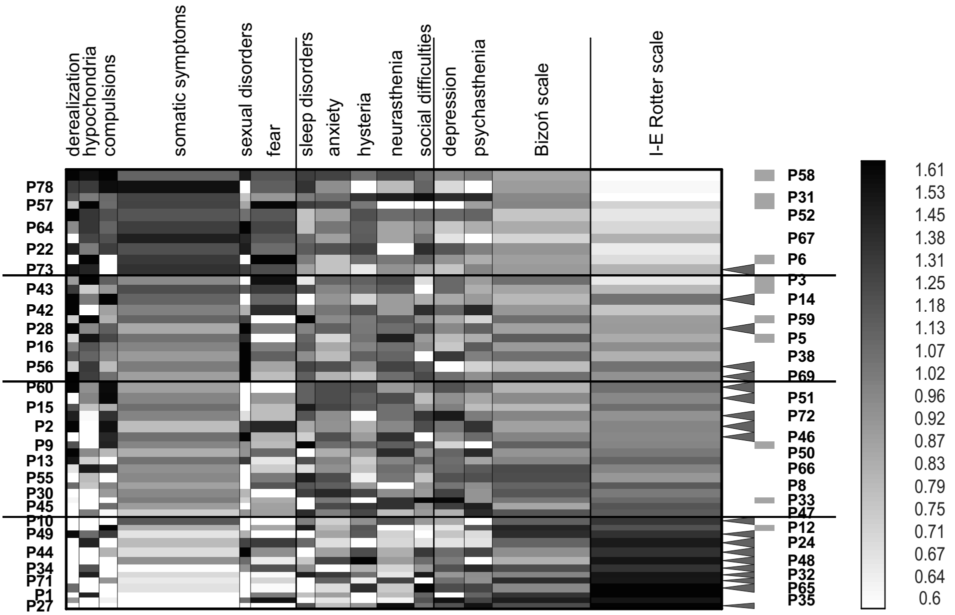

Fig. 2. The post-GCA overrepresentation map for the FIT subset (markers at right, the scale of gray and patients' labels have the same meaning as in Fig. 1). The ordered patients' records are divided into 4 clusters, first-line variables are also divided into 4 clusters. The values of ρ^* and τ are 0.185 and 0.123, the regularity index τ_{max}/τ_{abs} is 0.55

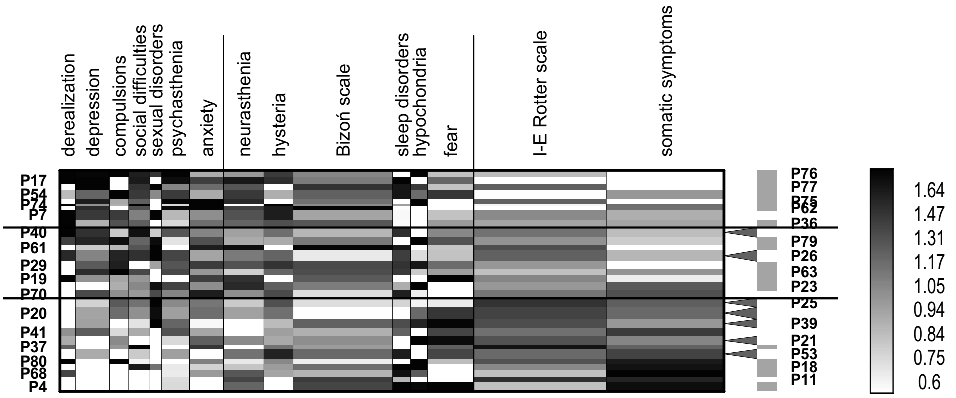

Fig. 3. The post-GCA overrepresentation map for the OUT subset (markers at right, the scale of gray and patients' labels have the same meaning as in Fig. 1). The ordered patients' records are divided into 4 clusters, the first-line variables are divided into 3 clusters. The values of ρ^* and τ are 0.202 and 0.135, the regularity index τ_{max}/τ_{abs} is 0.45

It is easy to find the sexual disorders subscale as an outlier among the variables, which is the same for FIT and OUT.

It should be reminded that the GCCA procedure provides an optimal division into clusters when the number of clusters is initially specified by an analyst. Hence, this is another individual decision made usually after some trials, while at each trial the GradeStat program supports the search with suitable computations.

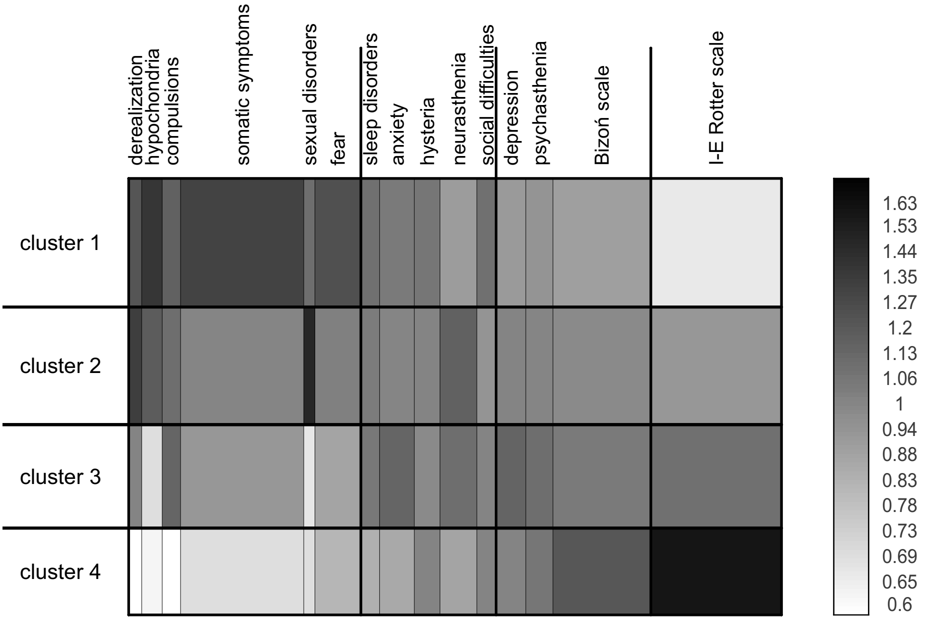

Fig. 4. The overrepresentation map for the first line variables with the patients' records aggregated in the particular clusters of subset FIT

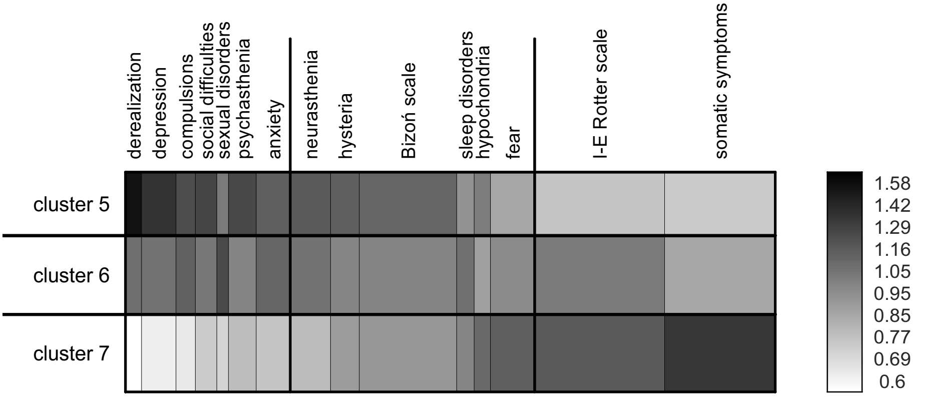

Fig. 5. The overrepresentation map for the first line variables with the patients' records aggregated in the particular clusters of subset OUT

4. Grade Clusters in FIT and OUT; Their Description and Interpretation

Figure 4 is supplemented by Table 1, the left part of which contains mean values for all first line variables in the aggregated clusters of patients in FIT. Similarly, Fig. 5 is supplemented by Table 2. The right parts of Table 1 and 2 describe additionally each cluster, basing on respective values of the second line variables.

Table 1. Mean values obtained in the clusters of patients in FIT for all first-line variables and an additional description of the clusters by the second-line variables

Cluster of patients in FIT	First – line variables															Second – line variables										number of patients
	subscales of „O" Symptom Check list of Aleksandrowicz																				No of ICD-10 diagnoses					
	derealization	hypochondria	compulsions	somatic symptoms	sexual disorders	fear	sleep disorders	anxiety	hysteria	neurasthenia	social difficulties	depression	psychasthenia	Bizoń scale	I-E Rotter scale	sum of „O" subscales	% married	% women	age	level of education	F41	F43.2	F48.0	F48.9	F60.8	
Cluster 1	18	32	24	183	14	65	23	41	31	38	24	31	29	79	9.6	550	50	40	38	3	1	4	3	0	3	10
Cluster 2	18	25	21	128	17	49	20	36	27	44	19	31	28	79	12.1	462	56	55	36	5	6	2	1	1	4	11
Cluster 3	12	13	19	105	7	37	18	36	23	37	18	31	27	74	12.4	382	63	63	30	3	3	2	2	8	6	16
Cluster 4	3	10	8	66	6	29	12	23	20	25	15	23	22	71	15	261	70	77	35	4	3	1	3	5	2	13
FIT	12	19	18	116	10	43	18	34	25	35	19	29	26	75	12.4	402	60	60	34	3.5	13	9	9	14	15	50

Table 2. Mean values obtained in the clusters of patients in OUT for all first-line variables, with an additional description of the clusters by the second line-variables

Cluster of patients in OUT	First – line variables															Second – line variables										number of patients
	subscales of „O" Symptom Check list of Aleksandrowicz																				No of ICD-10 diagnoses					
	derealization	depression	compulsions	social difficulties	sexual disorders	psychasthenia	anxiety	neurasthenia	hysteria	Bizoń scale	sleep disorders	hypochondria	fear	I-E Rotter scale	somatic symptoms	sum of „O" subscales	% married	% women	age	level of education	F41	F43.2	F48.0	F48.9	F60.8	
Cluster 5	17	32	17	19	8	24	27	32	23	62	11	11	26	6	57	304	22	33	27	3	4	2	0	2	2	9
Cluster 6	15	31	19	19	13	24	32	37	24	69	16	12	36	11	81	358	22	56	32	4	2	1	0	5	2	9
Cluster 7	6	10	11	13	7	18	22	27	22	63	15	15	43	12	125	341	42	83	32	3	3	3	1	4	5	12
OUT	12	26	15	17	9	22	26	31	23	64	14	13	36	10	91	335	30	60	30	3.5	9	6	1	11	9	30

The mean values in the clusters 1 to 4 of the subset FIT diminish in almost all sub-scales of Aleksandrowicz, firstly very fast for variables on the left, then gradually more slowly. The only exception from this rule concerns neurasthenia, since its maximal mean value appears in the cluster 2 (it is surprising that there is only one patient with neurasthenia diagnosed in this cluster, so these diagnoses should be verified). The means in the cluster 4 are everywhere distinctly smaller. The means of the Bizoń scale slowly diminish too. Conversely, the means of the Rotter scale increase, reaching even 15 in the cluster 4.

A similar analysis of the left side of Table 2 reveals that the cluster means of the first line variables tend to be the highest in the clusters 5 and 6 except for *the somatic symptoms* that attains the high mean 125 in the cluster 7 and also except for the Rotter scale, for which the means increase (rapidly from the cluster 5 to the cluster 6).

Let us turn to the second line variables, starting with the sum of all 13 subscales of the Aleksandrowicz "O" Symptom Check list. The mean of this sum rapidly decreases in the clusters in FIT (from 550 in the cluster 1 to 261 in the cluster 4).

The next group of second line variables concerns marriage states, gender, age and education. Each cluster was described with the percent of married patients (label *"% married*"), percent of women (label *"% women"),* mean age (in years) and mean level of education (according to the usual standards on the scale 1-2-3-4-5-6). In FIT as well as in OUT, percents of *married* and *women* increase in the successive clusters.

The last group of variables concerns medical ICD-10 classification. ICD-10 diagnoses were made irrespectively of the questionnaires and the scales. In Tables 1-2 only the most frequent diagnoses were noted, such as:

- F 41 – anxious and depressive disorders,
- F 43.2 – adjustment disorders,
- F 48.0 – neurasthenia,
- F 48.9 – neurotic disorders,
- F 60.8 – personality disorders.

Some of them (especially personality disorders) appeared together with other diagnoses.

The informal summary description of the clusters 1–7, using the information from first and the second line variables, is as follows:

cluster 1: on average, the oldest patients (mostly men), usually with Internal Locus of Control; very high mean level of neurotic symptoms (especially hypochondria, somatic symptoms, fear), the highest mean sum of "O"; most frequent diagnosis: adjustment disorders;

cluster 2: the best educated patients, the highest mean level of derealisation, sexual disorders, neurasthenia; most frequent diagnoses: anxious and depressive disorders with personality disorders;

cluster 3: the patients (in major part married, mostly women) with high level of anxiety and depression, diagnosis: neurotic disorders and personality disorders;

cluster 4: the married women with External Locus of Control and with neurotic disorders diagnosed in majority of cases; the lowest mean sum of neurotic symptoms; almost absent derealisation and sexual disorders;

cluster 5: on average, the youngest single men, mostly with Internal Locus of Control, with the lowest mean level of the somatic symptoms, very low mean sum of "O" symptoms, high derealisation and depression; dominating diagnosis: anxious and depressive disorders;

cluster 6: neurotic patients, single, educated, higher level of sexual disorders, anxiety and neurasthenia (but no one with neurasthenia diagnosis);

cluster 7: the women, with low mean level of education, rather single, mostly with the somatic and fear symptoms; diagnoses in majority: personality disorders and neurotic disorders.

Each aggregated cluster in FIT and OUT can be compared with the aggregations performed for the whole FIT and OUT subsets, shown in additional rows at the bottom of Table 1 and Table 2. It is also important to compare two aggregated subsets: FIT with OUT. First of all, they differ in the order of variables. The most spectacular differences refer to *the somatic symptoms* (the fourth position in FIT, the last – fifteenth-in OUT) and *the depression* (the twelfth place in FIT, the second in OUT). This will be commented in Section 5. As far as the mean values are concerned, they are much higher in FIT than in OUT for the majority of variables (*somatic* in FIT 116, in OUT 91; the Bizoń scale in FIT 75, in OUT 64; the Rotter scale in FIT 12.4, in OUT 10). The mean age is 34 years in FIT and 30 in OUT. The percent of married patients is 60 in FIT, 30 in OUT, while the mean sum of “O” subscales amounts to 402 in FIT and 335 in OUT. Almost all neurasthenics are in FIT, only one case belongs to OUT and their cluster localizations and clinic diagnoses in terms of ICD-10 are different. In both subsets there is only partial agreement between the grade taxonomy obtained from the scales and questionnaires and clinic diagnoses in terms of the ICD-10 classification.

5. Final Remarks

Let us summarize the main results of the study. General trends became visible when the variables’ values in the patients clusters were aggregated (Figures 4–5) and supplemented with the suitable aggregation of second line variables shown in Tables 1–2. Two sets of the ordered clusters, one set for FIT and one for OUT, were formed and thoroughly described. These clusters are highly diversified in an ordered way relative to an axis with two poles. In the post-GCA FIT one pole is formed by *the somatic symptoms* accompanied by *derealisation, hypochondria, compulsions,* while the other one consists of the Rotter scale accompanied by the Bizoń scale. The two opposite poles in the post-GCA OUT are: *depression* accompanied by *derealisation,* and the Rotter scale accompanied by *somatic symptoms*. We observe in the post-GCA OUT the same directions of changes in the Rotter scale and the adjacent *somatic symptoms* while in the post-GCA FIT the direction of changes and the allocation of these variables are opposite. Thus, the relations between *the somatic symptoms* and the Rotter scale form the main difference between FIT and OUT.

The presented figures and tables are used to describe the model specified for the whole dataset of size 80×25. It should be mentioned that according to the classification considered in [8], the performed exploration study can be recognized as modeling with hidden variables but using specific grade methods presented in [4]. The important hidden variables (the Rotter scale*, the somatic symptoms, the depression*) appearing in the present study order the set of patients in FIT and in OUT according to the optimal GCA-reordering.

It is worth noting that the *fear* and *depression* columns in all figures and tables are situated far from each other, which indicates their dissimilarity, in spite of the fact that they are put together in the ICD-10 categories (as in F.41.2). Therefore, it seems that in clinical practice depression disorders should rather be treated differently from fear disorders.

Furthermore, analyzing Tables 1 and 2 it may be noticed that there is a tendency to exterior-controllability for neurotic married women and to intra-controllability for single young men.

Finally let us stress that rank correlations between some neurotic symptoms and the Rotter scale are small in FIT and TOTAL, while they are rather high in OUT. Those relevant rank correlations appearing in OUT are collected in Table 3. Data analysis based only on the correlation table of the whole set would overlook this phenomenon revealed after the division of TOTAL into FIT and OUT.

Table 3. Rank correlations of I-E Rotter scale with some neurotic symptoms in FIT, OUT and in TOTAL

NEUROTIC SYMPTOMS	RANK CORRELATIONS WITH I-E ROTTER		
	FIT	OUT	TOTAL
Somatic symptoms	–0.16	**0.53**	0.19
Fear	–0.11	**0.52**	0.24
Hysteria	0.0	**0.48**	0.28
Sexual disorders	0.12	**0.48**	0.28
Compulsions	–0.09	**0.31**	0.12
Depression	0.07	**0.25**	0.18
Sum of "O" subscales	–0.04	**0.61**	0.27

Further analysis of the dataset obtained after the completion of the clinical therapy is planned, basing on the clusters introduced in the present paper.

Acknowledgment

The authors would like to express great gratitude to their Master and Friend Prof. Elżbieta Pleszczyńska for her invaluable help.

References

1. Aleksandrowicz J.W. et al.: Symptoms Check List “S” and “O” – Tools Used for Diagnosis and Description of Neurotic Disorders (in Polish), Psychoterapia, 1981, 37, 11–27.
2. Rotter J.B.: Generalized Expectancies for Internal Versus External Control of Reinforcement, Psychological Monographs, 1966, 80, 1.
3. Welcz H.: The Exterior- and Intra-Controllability and Neurotic Symptoms (in Polish). PhD thesis, Medical Academy in Lublin, 2002.
4. Kowalczyk T., Pleszczyńska E., Ruland F. (Eds.): Grade Models and Methods for Data Analysis. With Applications for the Analysis of Data Populations. Berlin, Springer Verlag, 2004.
5. Pleszczyńska E.: Application of Grade Methods to Medical Data: New Examples, Biocybernetics and Biomedical Engineering, 2007, 27, 3.
6. http://gradestat.ipipan.waw.pl/ [Accessed 2009 Jan 10]
7. Kowalczyk T. On derivation of maximal Spearman rho and maximal Kendall tau for bivariate distributions, ICS PAS Report No. 1011, 2008.
8. Hand D., Mannila H., Smyth P.: Principles of Data Mining, MIT Press, Cambridge, 2001.