Bulgarian Chemical Communications, Volume 47, Special issue A (pp. 124–136) 2015

Multivariate statistical assessment of obesity patients' clinical parameters

R. T. Georgieva-Nikolova[1], P. A. Gateva[2], R. K. Hadjiolova[3], M. P. Slavova[4], M. M. Nikolova[5], V. D. Simeonov[*]

[1]*Department of Chemistry and Biochemistry, Faculty of Medicine, Medical University of Sofia, Sofia, Bulgaria*
[2]*Department of Pharmacology and Toxicology, Faculty of Medicine, Medical University of Sofia, Sofia, Bulgaria*
[3]*Department of of Pathophysiology, Faculty of Medicine, Medical University of Sofia, Sofia, Bulgaria*
[4]*Institute of Electrochemistry and Energy Systems, Bulgarian Academy of Sciences, Sofia, Bulgaria*
[5]*Middlesex University, School of Health and Education – London, United Kingdom*
[6]*Department of Analytical Chemistry, Faculty of Chemistry and Pharmacy, University of Sofia "St. Kl. Okhridski" Sofia, Bulgaria*

Received December 2, 2014, Revised March 22, 2015

The present study deals with multivariate statistical interpretation of clinical parameters of obesity patients. The goal of the study is to find relationship and similarity between the traditional obesity monitoring characteristics and to determine patterns of similarity between the patients participating in the investigation. Cluster analysis and principal components analysis were used as multivariate statistical methods in the data mining procedure in which 113 patients were included. It has been shown that the status of the patients is dominantly related to parameters characterizing the obesity problem (body mass index, fat mass, weight, degree of obesity etc.) and not so directly with other parameters characterizing mainly the general health status (cholesterol, triglycerides, glucose level etc.). This could help in optimizing the number of clinical variables necessary for monitoring obesity. Further, specific patterns of similarity between patients were defined and the parameters responsible for their formation were determined. In such a way a more individual treatment of the patients becomes possible. A distinctive separation between male and female patients was statistically proven.
It has to be stated that for the first time multivariate statistical analysis is applied for assessment of the health status of obesity patients.

Key words: Obesity, clinical parameters, multivariate statistics, health assessment

INTRODUCTION

Obesity is an issue of worldwide significance, affecting both adults and children. According to the World Health Organisation (WHO), over 400 million people in the world are suffering from it [1]. Obesity is a medical condition, in which the body fat levels are higher than normal and are considered harmful. It occurs as a result of imbalance between an individual's energy consumption through food and his energy expenditure [2].
It may also be triggered by medications or endocrine or psychiatric disorders. As with many other medical conditions, obesity results from the interplay between genetic and environmental factors.

Some medications can cause weight gain or changes in body structure [3]. Some physical and mental conditions and the medications used for their treatment can increase the risk of obesity. Although obesity in itself is not considered a psychiatric disorder, patients with such are more prone to becoming overweight or obese [4]. Polymorphism in genes controlling the appetite and metabolism, coupled with enough food energy, predisposes to obesity [5]. The percentage of genetic factor-related obesity in the study population varies between 6 and 85% [6, 7]. However, genetic factors only lead to obesity when coupled with environmental ones [8-10].
The dramatic increase in obesity cases worldwide cannot be explained with genetic factors alone [11]. Studies show that obesity is caused by a combination of different factors, rather than a high energy intake and a low expenditure [12].

Metabolic syndrome – a disruption in the body metabolism – results from the excess weight and obesity. Obesity and metabolic syndrome can cause diabetes mellitus type 2, obstructive sleep apnea, some types of cancer, osteoarthritis and osteoporosis, asthma, arterial hypertension, dyslipidemia, gout, liver steatosis, chronic gastroenterocolitis [13, 14].

Men with metabolic syndrome are marked with lower testosterone levels, i.e. sexual 'aging'.

* To whom all correspondence should be sent:
E-mail: vsimeonov@chem.uni-sofia.bg

The metabolic syndrome also raises significantly the risk of cancerous formations in the following organs: the prostate, the mammary glands, the endometrium, and the ovaries. The metabolic syndrome also injures the liver and leads to non-alcoholic steatohepatitis and cirrhosis.

Therefore, an assessment of the obesity as a serious medical problem needs large data sets of various indicators. The estimation and the useful information extraction from such big data set requires application of appropriate strategies most effective of which are the methods of the multivariate statistics like cluster analysis and principal components analysis.

The aim of the present study is to classify, model and interpret a clinical data set of obesity patients in order to detect relationships between the parameters or reveal specific patterns of obesity patients. It could be of use for optimization of the monitoring process and applying additional attention to the different groups of affected patients. This is the first ever attempt to interpret clinical data from obesity sufferers by the use multivariate statistical analysis.

MATERIAL AND METHODS

Data collection

Data from 113 patients (28 male and 85 female) with different degrees of obesity in University Hospital 'Alexandrovska', Sofia, Bulgaria have been used in this study. Totally 40 clinical parameters and sex differentiation were collected for the assessment procedure as follows [15-20]:

1. Sex (not a real parameter, just information);
2. Age, years;
3. Height, cm;
4. Weight, kg;
5. Fat Mass (FM), kg;
6. Fat, % – The percentage of the body fats is calculated in relation to the total patient's weight;
7. Fat-Free Mass (FFM), kg;
8. Muscle Mass (MM), kg;
9. Total body water (TBW), kg;
10. Total body water (TBW), % ;
11. Bone Mass (BM), kg – The obesity patients have lower bone density than this which corresponds to their age [15];
12. Basal Metabolic Rate (BMR), kJ. This is the energy which needs the body at resting to function effectively [16];
13. Basal Metabolic Rate (BMR), kcal;
14. Metabolic Age (MA), years – the age of the metabolism of the body;
15. Visceral Fat Rating – evaluation of the inner abdominal obesity [17];
16. Body Mass Index (BMI), кg/m²;
17. Ideal Body Weight (IBW), kg – it is the weight at which the individual has the chance to live longer;
18. Degree of obesity, %;
19. Hemoglobin (HGB), g/L
20. White blood cells (WBC), $x10^{-9}$/L;
21. Red Blood Cells (RBC), $x10^{-12}$/L
22. Hematocrit – HCT, it measures the volume of the erythrocytes in the blood;
23. Platelets (PLT), $x10^{-9}$/L
24. Mean Corpuscular Volume (MCV), or Mean Cell Volume, fL – MCR is a measure of the average volume of a red blood corpuscle (or red blood cell);
25. Mean Corpuscular Hemoglobin (MCH), pg;
26. Mean Corpuscular Hemoglobin Concentration (MCHC), g/L;
27. Red Blood Cell Distribution Width (RDW), % – is useful biomarker in the determining of cardiovascular risk;
28. Mean platelet volume (MPV), fL, – lower values of MPV are present in the aplastic anemia [18];
29. Alanine aminotransferase (ALAT)**,** U/L;
30. Aspartate aminotransferase (ASAT)**,** U/L – reveals fatty liver [19];
31. Creatinine, μkmol/L –a parameter of the kidney function;
32. Cholesterol, mmol/L – Total cholesterol is assessed for determination of the damage of the fat metabolism and estimation of the risk of cardiovascular diseases.
33. High-Density Lipoproteins (HDL), mmol/L – parameter for the risk of cardiovascular diseases.
34. Low-Density Lipopreteins (LDL), mmol/L ;
35. Triglycerides (TG), mmol/L ;
36. High-sensitivity C-reactive protein (hs CRP)**,** mg/L;

Parameters 32 – 36 estimate the risk of cardiovascular diseases.

37. Oral Glucose Tolerance Test (OGTT)**,** mmol/L – The most important diagnostic value has the fasting glucose, OGTT 0 [20];
38. OGTT 120 (2h after administration) with 75 g glucose;
39. Glycated Hemoglobin – Hb A_{1c}, %. A parameter for the long-term blood glucose;
40. Immuno-Reactive Insulin (IRI)**,** mU/L – obesity is associated with hyperinsulinism;
41. C-peptide**,** ng/mL – is a component of the proinsulin.

Multivariate statistics

In the present study Cluster analysis (CA) and Principal Components Analysis (PCA) were used. Both methods are well documented and used in many multivariate statistical studies for data

modeling, data projection and data interpretation procedures. They belong to the classical data mining approaches and represent a serious part of the intelligent data analysis strategies [21].

CA is well-known and widely used multivariate statistical approach In order to cluster objects characterized by a set of variables (e.g. patients by clinical parameters), one has to determine their similarity. A preliminary step of data scaling is necessary (e.g. autoscaling or z – transform) where normalized dimensionless numbers replaces the real raw data values. Thus, even serious differences in absolute values are scaled to similar ranges. Then, the similarity or the distance between the objects in the variable space can be determined usually by calculation of the Euclidean distance. There is a wide variability of clustering (linkage) algorithms but the typical ones include the single linkage, the average linkage or the Ward's method. The representation of the results of the cluster analysis is performed by a tree-like scheme called dendrogram.

PCA is a typical display method, which allows to estimate the internal relations in the data set. There are different variants of PCA but basically, their common feature is that they produce linear combination of the original columns in the data matrix (data set) responsible for the description of the variables characterizing the objects of observation. These linear combinations represent a type of abstract measurements (factors, principal components) being better descriptors of the data structure (data pattern) than the original (chemical or physical) measurements. Usually, the new abstract variables are called latent factors and they differ from the original ones named manifest variables. It is a common finding that just a few of the latent variables account for a large part of the data set variation. Thus, the data structure in a reduced space can be observed and studied.

RESULTS AND DISCUSSION

As already mentioned the data set consists of 113 cases (patients) and 40 clinical parameters [113x40]. The data set was analyzed by CA (z-transform of the raw data; squared Euclidean distances as similarity measures; Ward's method of linkage and Sneath's criterion for cluster significance) and by PCA (Varimax rotation mode). The main goals of the multivariate statistical data treatment were:

1. to find relationships between the clinical parameters and based on the relationships to determine significant indicators in the treatment of the problem;
2. to find patterns of similarity between the patients treated and to determine discriminant factors (clinical parameters) for each pattern;
3. to define the latent factors responsible for the data set structure and to relate them to the clinical parameters used.

In the first run of the statistical analysis all patients were involved. In Fig. 1 the clustering of clinical parameters for all patients is presented.

Two significant clusters are formed at level $2/3D_{max}$:

K1 (*Age, MetaAg, Crea, GlyHb, OGTT0, OGTT12, RDW, Trig, CHOL, LDL, TBW%, HDL, MCV, MCHC, MPV, MCH, ALAT, ASAT)* and

K2 (*Hei, IdBW, BMRc, BoneM, MusM, FFM. BMR, TBW, RBC, HCT, HGB, Wei, FatM, BMI, DegOb, VisFat, Fat, CRP, IRI, CPEP, WBC, PLT*)

In these two clusters some subclusters could be defined but, in general, there is a significant similarity between the clinical indicators for obesity. All indicators are generally divided in two big groups:

Obesity indicators (dominantly in cluster K2)

General health status indicators (dominantly in cluster K1)

There is an option to select smaller number of parameters when assessing the state of obesity and the general health status of the patients, which is related to the obesity syndrome.

In the next dendrogram (Fig. 2) the clustering of all 113 patients is shown. Two major clusters are formed (the number of patients is reduced for better readability of the graph but the clustering involved all patients). The separation is achieved by sex: cluster 1 (the smaller cluster) consists of totally 29 cases with 24 male patients and 5 female patients; cluster 2 (the bigger one) consists of totally 84 cases with 80 female and 4 male patients. Therefore, there is a significant separation between male and female obesity cases.

If the average values for each clinical parameters for the two clusters formed are compared (Table 1) and comparable values from

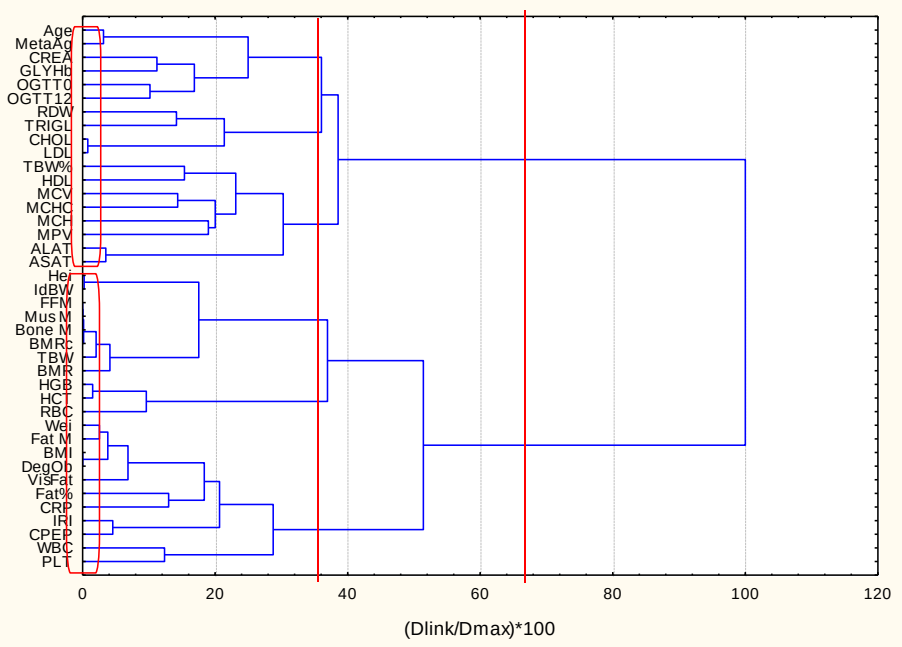

Fig. 1. Hierarchical dendrogram for clinical parameters (all patients)

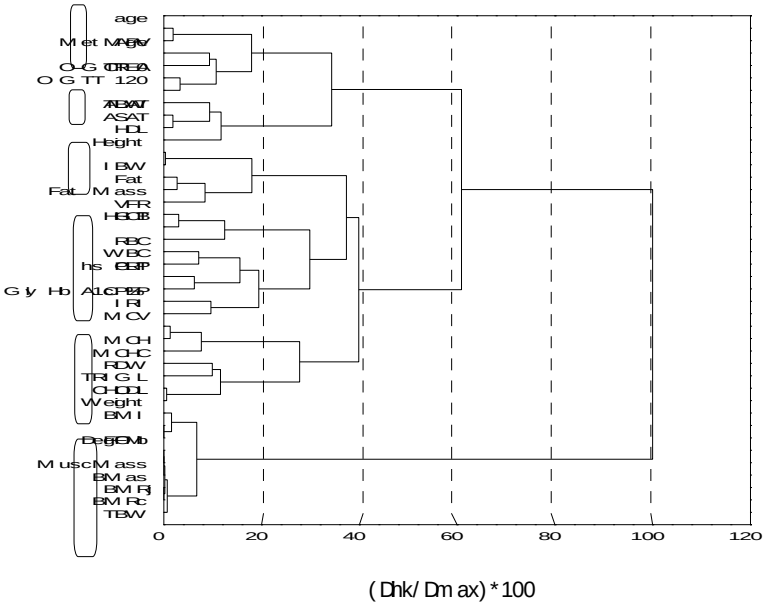

Fig. 2. Hierarchical dendrogram for clustering of 113 patients

both clusters differ around 50 % following conclusions could be mentioned:

There is no significant difference between lots of the clinical indicators for each one of the clusters formed. In general, the members of the "male"cluster show higher values for many of the "obesity" indicators like weight, fat mass, FFM, muscle mass, TWB, bone mass, BMR, visceral fat, BMI, ideal body weight, degree of obesity, HGB. Probably, it has to be expected due to objective reasons – men are physically stronger and more affected by obesity. The other clinical parameters related to the general health status (blood parameters, glucose parameters, liver parameters) are quite similar in both clusters.

It was interesting to separate the data set into "male" and "female" subsets and try to interpret separately both subsets.

Table. 1. Average values for the clinical parameters for clusters 1 and 2

Clinical parameter	Cluster 1 ("male")	Cluster 2 ("female")
Age	45.97	49.08
Height	176.79	161.40
Weight	**133.51**	88.42
Fat	39.77	41.13
Fat Mass	**54.00**	37.07
FFM	**79.53**	51.35
Muscle mass	**75.62**	48.78
TBW	**57.69**	36.69
TBW %	45.10	42.14
Bone Mass	**3.92**	2.60
BMR kJ	**10392.07**	6344.89
BMR ccal	**2483.76**	1580.42
Metabolic Age	59.41	60.39
Visceral Fat	**20.93**	10.86
BMI	42.92	33.85
Ideal Body	68.93	57.60
Degree of obesity %	**95.12**	54.05
HGB	152.17	132.96
WBC	8.38	7.17
RBC	5.08	4.59
HCT	0.45	0.40
PLT	248.55	279.93
MCV	89.14	87.42
MCH	30.01	32.24
MCHC	336.38	333.04
RDW	17.44	13.50
MPV	8.28	8.43
ALAT	28.79	20.40
ASAT	22.97	19.48
CREA	78.50	66.82
CHOL	5.35	5.55
HDL	1.27	1.45
LDL	3.26	3.42
TRIG	1.87	1.52
CRP	7.91	5.67
OGTT0	5.85	5.30
120OGTT	6.62	5.99
GlyHbA1	5.81	5.68
IRI	**23.21**	14.36
CPEP	5.04	3.76

The hierarchical dendrogram for linkage between clinical parameters for male patients is given (Fig. 3).

Six clusters are formed:

K1: *TBW BMRc BMR Bone Mass MuscMass CPR FFM DegOb BMI Weight VisFat*

K2: *LDL CHOL TRIG RDW MCHC MCH MCV*

K3: *IRI GLYHb CPEP PLT WBC RCB HCT HGB*

K4: *FatM Fat IBW Height*

K5: *HDL ASAT ALAT TBWc*

K6: *OGTT120 OGTT0 CREA MPV MetAg Age*

It is seen that the parameters are clustered in groups related to the obesity (K1, K4), blood indicators (K2, K3), liver parameters (K5) and glucose indicators and age (K6).

This separation is confirmed in principle by the application of principal components analysis. Six latent factors are responsible for explanation of nearly 70 % of the total variance of the system. Factor loadings are presented and the significant ones are marked by bold (Table 2).

The first latent factor (conditional name "obesity factor") indicates the close relationship between the indicators for obesity. It is interesting to note that the parameter "metabolic age" is negatively correlated to the other parameters with significant factor loadings. CPEP and IRI could be also included in this group of indicators although their factor loadings are lower tnan the required 0.70 level. The second principal component ("glucose level factor") stands for over 11 % of the total variance and

indicates logical relationship between glucose level and age. Several blood indicators are also included. The third hidden variable is related to PC1 since it includes other important obesity indicators ("fat indicators factor") and explains over 10 % of the total variance. The fourth principal component is related to the blood quality parameters ("blood parameters factor"). The last two latent factors indicate the role of several indicators for the general health status like cholesterol and triglycerides or blood quality (platelets) It is worth to mention that a certain number of clinical indicators do not contribute significantly to the description of the obesity syndrome (GlyHb, CRP, HDL, ASAT, ALAT). This conclusion offers an opportunity for experimentation aiming optimal selection of significant obesity indicators for male patients.

For female patients (Fig. 4) the clinical indicators are generally divided into two major cluster (the first one included typical obesity parameters and some blood quality characteristics; the second one links glucose level, liver function, general health status parameters along with metabolic age and age). A closer look into the groups could reveal (cluster significance according Sneath of 1/3 D_{max}) five clusters of parameters:

K1: *BMR TBWc BMRc BoneMass MuscMass FFM*
K2: *CRP Fat VisFat DegOb BMI FatM Weight*
K3: *RDW PLT WBC IBW Height*
K4: *LDL CHOL MPV MCH MCHC MCV HDL TBW % RBC HCT HGB*
K5: *ASAT ALAT CPEP IRI OGTT120 OGTT0 TRIG GLyHb CREA MetA Age*

The "female clustering" resembles the "male" one revealing groups of similarity related to obesity indicators (K1, K2), blood, liver and glucose indicators (K3, K4, K5).

The factor loadings for this subset of patients are shown after carrying out principal components analysis (Table 3).

Five latent factors are responsible for the data structure in the female subset (explanation of nearly 60 % of the total variance). PC1 and PC2 are typical "obesity indicators factors", since PC3 and PC5 include "blood" and glucose level" indicators. PC 4 reveals a specific relationship for parameters defining general health status (age, ideal body weight, cholesterol, LDL). Even more indicators than those in the case with male patients remain insignificant for explanation of the data structure: HGB, WBC, PLT, MCH, MCHC, MPV, ALAT, ASAT, CREA, CHOL, HDL, TRIG, CRP, CPEP (Table 3).

There is a slight difference between the clustering of the clinical indicators for male and female patients – those for female patients are grouped more compact (less clusters) which is an indication for higher level of similarity between the indicators for the general health status (blood, liver, glucose).

In the next step of the statistical analysis it was of substantial interest to understand if there are specific patterns among the groups of male and female patients and to determine the discriminant indicators for these patterns.

The hierarchical dendrogram for 28 male obesity patients is shown (Fig. 5).

Two significant clusters could be determined:

K1: *6, 7, 8, 9, 10, 11, 12, 13, 14, 20, 24, 25, 26*
K2: *1, 2, 3, 4, 5, 15, 16, 17, 18, 19, 21, 22, 23, 27, 28*

The clustering of 85 female obesity patients is shown (Fig. 6).

Four clusters are found as follows:

K1: *67, 22, 79, 29, 12*
K2: *83, 46, 42, 21, 49, 85, 48, 43, 84, 82, 47, 72, 45, 74, 63, 44, 27, 18, 77, 14*
K3: *39, 66, 52, 37, 78, 81, 80, 76, 25, 35, 24, 34, 20, 65, 10, 75, 73, 61, 71, 69, 64, 26, 60, 53, 28, 50, 31, 11, 5, 13, 4*
K4: *16, 62, 9, 59, 58, 56, 38, 8, 41, 40, 17, 68, 30, 7, 54, 36, 6, 57, 51, 15, 3, 33, 19, 70, 55, 32, 23, 2,*

For identification of discriminant indicators the averages of each parameter for each cluster (both for male and female patients) were determined (Table 4, Table 5).

For male patients: two different patterns are identified among the group of totally 28 male obesity patients. The first pattern (cluster 1) consists of 13 patients characterized by high indication of most of the parameters (weight, fat content, fat mass, FFM, muscle mass, BMI, TBW, bone mass, degree of obesity or 23 out of all 40 parameters are with higher values). Obviously, this is the pattern of the *most affected patients with bad levels of obesity indicators.* Surprisingly, this is the group with the lower average age which proves the assumption that obesity starts recently in early age, even before 40. The second pattern (cluster 2) represents the rest of 15 patients with the relatively better levels of obesity indicators. They show only three higher levels of indicators forage, metabolic age (this is a logical relationship and ASAT but with very close value to that of values of cluster 1.

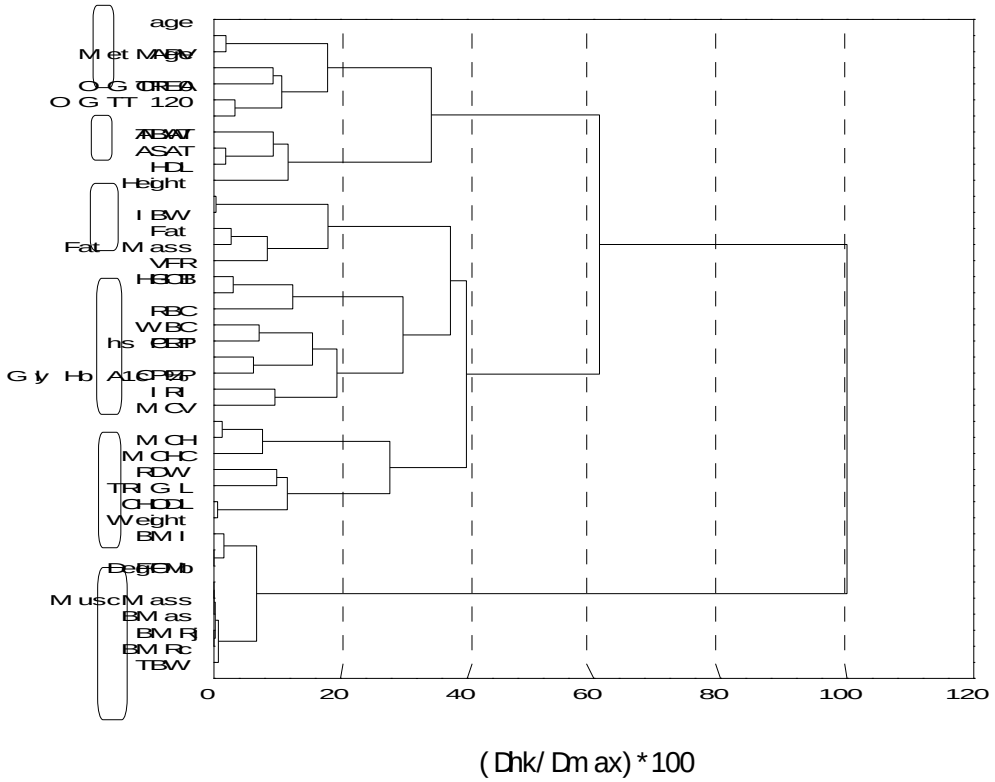

Fig. 3. Hierarchical dendrogram for clinical parameters for male patients

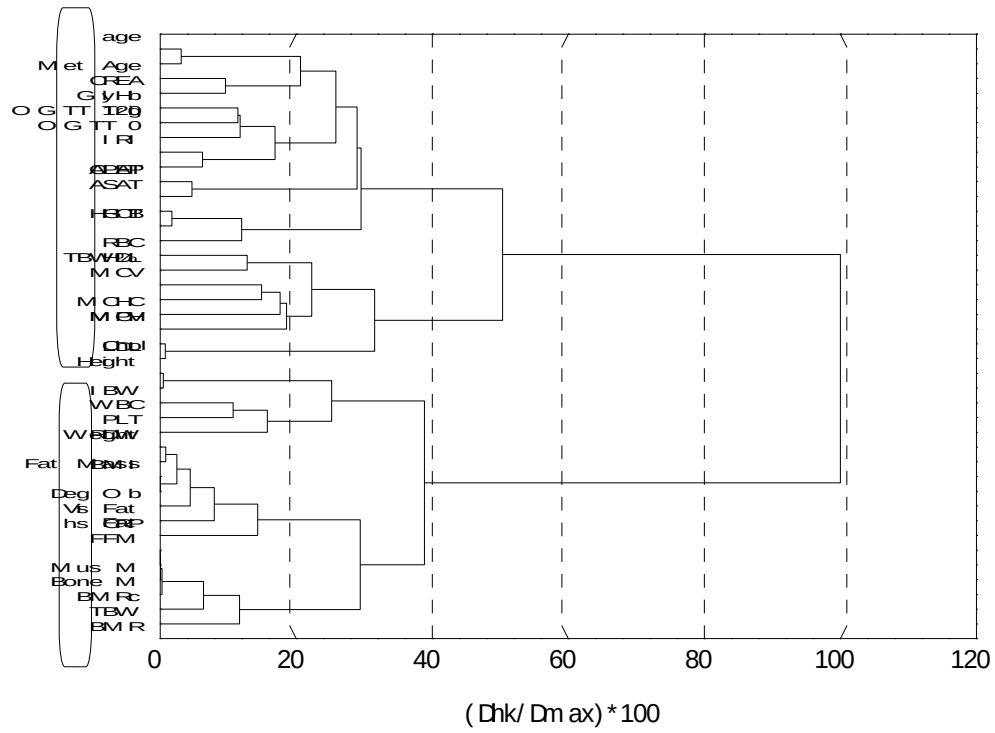

Fig. 4. Hierarchical dendrogram for clinical parameters for female patients

Table. 2. Factor loadings for male patients (significant loadings are marked by bold)

Variable	PC- 1	PC 2	PC- 4	PC 3	PC -6	PC- 5
age	-0,45	**0,703**	-0,171	-0,036	0,180	-0,197
Height	0,40	-0,113	0,035	**0,628**	-0,519	0,122
Weight	**0,872**	-0,004	0,024	0,478	0,000	0,007
Fat	-0,11	0,002	0,125	**0,926**	0,132	0,091
Fat Mass	0,44	-0,011	0,104	**0,863**	0,042	0,044
FFM	**0,982**	0,005	-0,060	-0,048	-0,039	-0,029
Muscle Mass	**0,985**	0,004	-0,061	-0,049	-0,039	-0,029
TBW, kg	**0,984**	0,004	-0,022	0,065	0,049	0,020
TBW, %	-0,001	0,141	-0,016	**-0,702**	0,051	0,030
Bone Mass	**0,98**	0,019	-0,043	-0,033	-0,045	-0,030
BMR, kJ	**0,99**	-0,002	-0,050	0,026	-0,017	-0,026
BMR, ccal	**0,99**	-0,002	-0,050	0,026	-0,017	-0,026
Meta Age	**-0,57**	0,497	-0,152	0,304	0,087	-0,181
Vis Fat	0,25	0,229	-0,111	**0,532**	0,469	0,019
BMI	**0,89**	0,063	0,028	0,277	0,268	-0,039
I B W	0,37	-0,094	0,026	**0,612**	-0,551	0,119
Deg obes	**0,89**	0,063	0,030	0,278	0,267	-0,039
HGB	-0,11	-0,489	**0,751**	0,172	0,131	-0,131
WBC	0,14	-0,238	0,042	0,089	**0,760**	0,284
RBC	-0,08	**-0,607**	-0,489	0,081	0,108	-0,269
HCT	-0,16	**-0,509**	0,371	0,166	0,304	-0,328
PLT	0,16	0,046	-0,076	-0,006	**0,745**	0,193
MCV	-0,07	0,281	**0,865**	0,065	0,129	0,052
MCH	-0,05	0,192	**0,927**	0,072	-0,004	0,173
MCHC	0,03	-0,048	**0,695**	0,035	-0,205	0,290
RDW	-0,10	-0,217	0,148	-0,130	-0,081	**0,599**
MPV	0,00	**0,701**	0,192	-0,053	0,025	-0,236
ALAT	*0,04*	*0,003*	*0,078*	*0,083*	*-0,147*	*0,085*
ASAT	*-0,10*	*0,289*	*0,142*	*0,105*	*-0,138*	*-0,024*
CREA	*-0,01*	*0,482*	*0,474*	*-0,176*	*-0,251*	*0,173*
cholesterol	-0,14	0,078	0,174	0,101	0,268	**0,846**
HDL	*-0,38*	*0,051*	*0,041*	*-0,149*	*0,156*	*-0,288*
LDL	-0,05	0,030	0,207	0,109	0,228	**0,826**
triglycerides	0,08	0,109	-0,048	0,181	0,061	**0,702**
hs CRP	*0,37*	*0,305*	*0,241*	*0,375*	*0,311*	*-0,022*
OGTT 0	0,16	**0,804**	0,109	-0,002	-0,058	0,145
OGTT 120	0,09	**0,876**	0,046	0,093	0,077	-0,054
Gly Hb A1c	*0,37*	*0,303*	*0,223*	*-0,330*	*0,253*	*0,027*
IRI	**0,55**	0,273	-0,003	0,290	-0,024	-0,193
CPEP	**0,59**	0,360	-0,100	0,234	0,110	-0,074
Expl.Var %	26.2	11.4	9.2	10.6	6.9	7.6

Table. 3. Factor loadings for female patients (significant loadings are marked by bold)

Variable	**PC- 1**	**PC - 2**	**PC - 3**	**PC 5**	**PC-4**
age	-0,261	0,402	0,219	0,025	**0,602**
Height	0,388	0,021	-0,311	0,091	**-0,621**
Weight	**0,862**	0,446	0,114	0,047	-0,166
Fat	0,243	**0,898**	0,029	0,046	-0,125
Fat Mass	**0,697**	0,656	0,099	0,051	-0,179
FFM	**0,959**	0,013	0,117	0,031	-0,113
Muscle Mass	**0,959**	0,008	0,111	0,029	-0,116
TBW	**0,810**	-0,092	0,020	0,014	-0,083
TBW, %	-0,126	**-0,921**	-0,014	-0,045	0,051
Bone Mass	**0,958**	0,027	0,114	0,049	-0,112
BMR, kJ	**0,615**	-0,132	0,258	0,175	-0,080
BMR, ccal	**0,960**	0,104	0,113	0,032	-0,159
Metabolic Age	-0,132	**0,728**	0,154	0,016	0,435
Visceral Fat Rating	0,482	**0,733**	0,303	0,009	0,058
BMI	**0,798**	0,487	0,219	0,003	0,070
I B W	0,353	0,017	-0,289	0,099	**-0,634**
Deg of obesity	**0,794**	0,489	0,223	0,004	0,066
HGB	0,062	0,066	0,472	**-0,800**	0,081
WBC	*0,396*	*0,050*	*0,280*	*-0,042*	*-0,255*
RBC	0,184	-0,005	**0,692**	0,120	0,001
HCT	0,110	0,094	0,557	**-0,655**	0,123
PLT	*0,241*	*-0,010*	*-0,042*	*0,204*	*-0,283*
MCV	*-0,116*	*0,091*	*-0,179*	*-0,783*	*0,130*
MCH	*0,068*	*-0,087*	*-0,090*	*-0,196*	*-0,021*
MCHC	*-0,063*	*-0,029*	*-0,176*	*-0,441*	*-0,064*
RDW	0,261	0,010	0,089	**0,723**	-0,064
MPV	*-0,213*	*0,243*	*-0,031*	*-0,077*	*-0,108*
ALAT	*0,307*	*-0,045*	*0,344*	*-0,007*	*0,166*
ASAT	*0,100*	*-0,070*	*0,338*	*0,083*	*0,233*
CREA	*-0,060*	*0,341*	*0,167*	*-0,169*	*0,156*
cholesterol	0,032	-0,006	-0,128	-0,011	**0,836**
HDL	*-0,096*	*-0,308*	*-0,498*	*-0,012*	*0,308*
LDL	0,041	0,012	-0,065	-0,036	**0,760**
triglycerides	*0,122*	*0,375*	*0,391*	*0,101*	*0,287*
hs CRP	*0,435*	*0,355*	*0,032*	*0,062*	*0,089*
OGTT 0	0,158	0,096	**0,651**	-0,072	-0,005
OGTT 120	-0,225	0,140	**0,645**	0,140	0,115
Gly Hb A1	*0,127*	*0,093*	*0,467*	*0,001*	*0,103*
IRI	0,469	0,183	**0,607**	0,139	-0,251
CPEP	*0,190*	*0,365*	*0,448*	*0,157*	*-0,144*
Expl.Var %	22.4	12.1	10.2	6.7	8.5

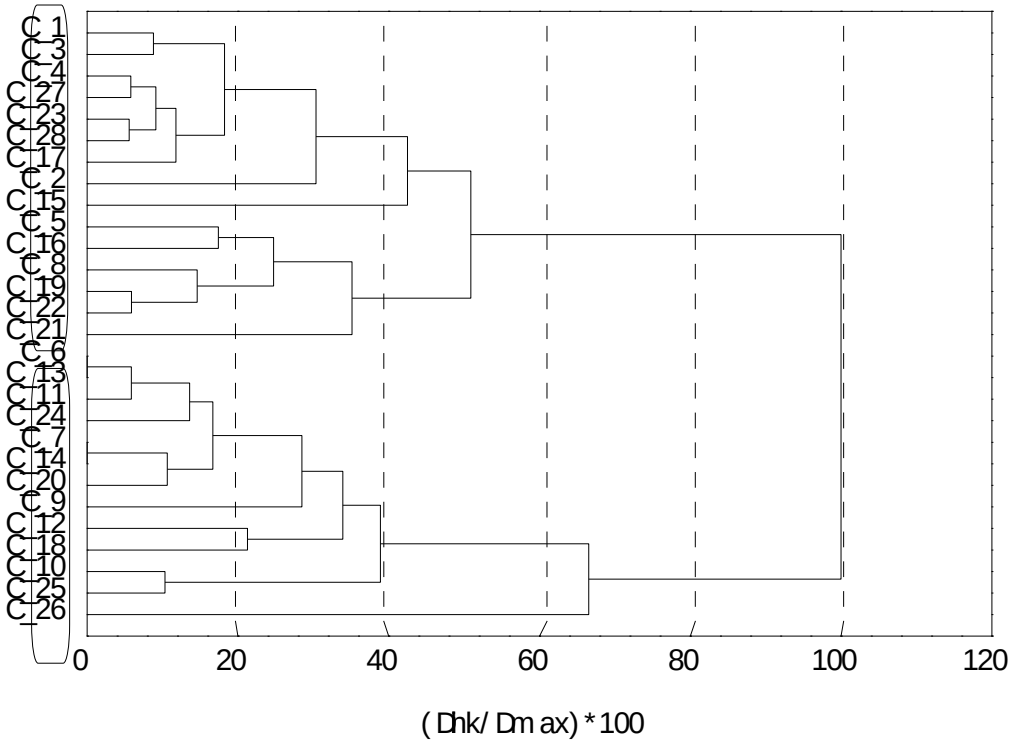

Fig. 5. Hierarchical dendrogram for 28 male obesity patients

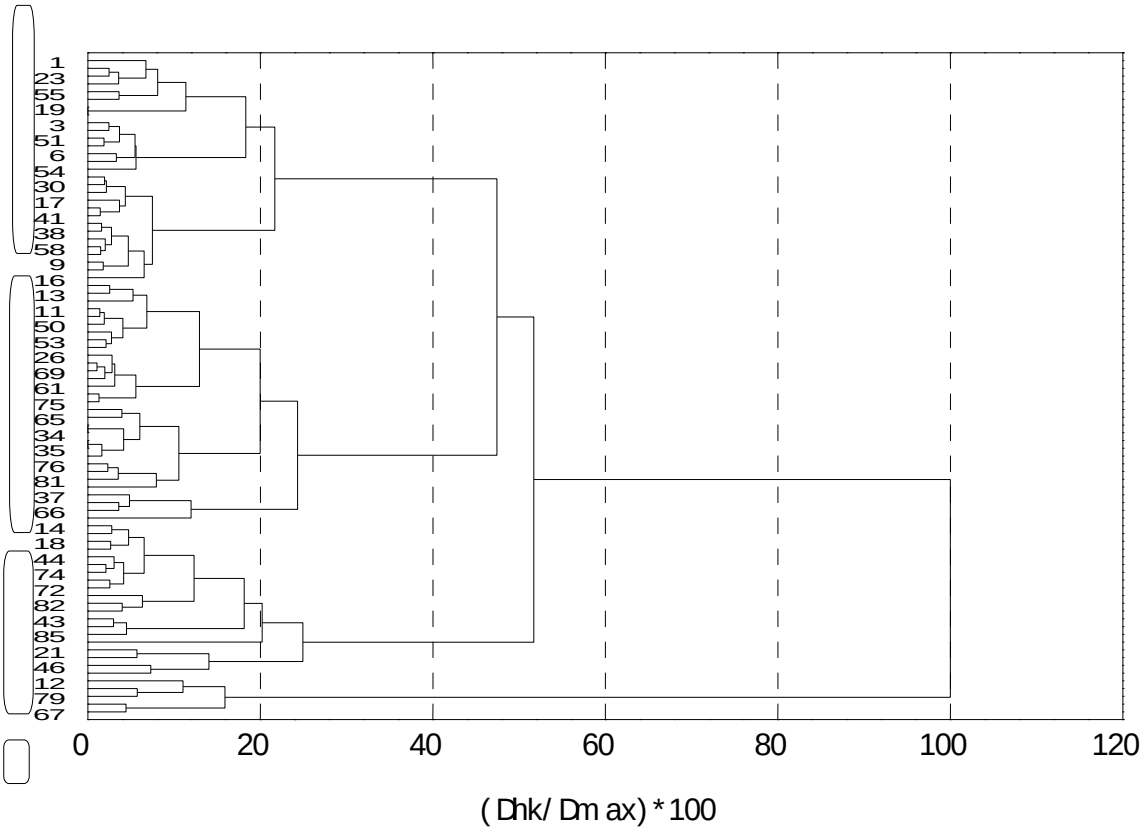

Fig. 6. Hierarchical dendrogram for female patients

So, this pattern could be conditionally named *patients with acceptable and controlled obesity problem.* It is important to note that for 14 parameters (out of all 40) the average levels are almost equal for the patients of both patterns among them cholesterol, LDL, HDL, several blood parameters, glucose level, ALAT.

Thus, they do not have important discriminating effect for the group of male patients. It could be recommended to use mainly obesity indicators for establishing the obesity status of the patients and to separate them into different patterns needing respective medical care and treatment.

The situation with the female patients is slightly different. Four groups of similarity are formed.

Cluster 1 (Table 5) includes only 5 cases (out of 85) having highest obesity indicators values − weight, fat, degree of obesity, BMI etc. This is definitely the pattern of *most affected female patients with bad levels of obesity indicators.* The group is of relatively young age (although not the lowest average age) and it is a troubling symptom. Cluster 2 with 20 cases resembles group of relatively young patients with better obesity indicators. This corresponds entirely to the concept of the statistical recognition as pattern of *patients with acceptable and controlled obesity problem.* Cluster 3 in the case with female patients with lowest average age covers the pattern of the *patients with starting obesity problem*.

Table. 4. Average values for clinical parameters for clusters of male patients

Parameter	Cluster 1	Cluster 2
Age	40.62	53.53
Height	181.38	173.80
Weight	145.77	106.81
Fat	41.15	33.97
Fat Mass	59.85	36.46
FFM	85.92	70.35
Muscle mass	81.72	66.87
TBW	64.57	51.95
TBW %	44.35	49.59
Bone Mass	4.20	3.47
BMR kJ	11399.85	8862.27.0
BMR ccal	2724.62	2128.13
Metabolic Age	54.46	65.47
Visceral Fat	22.31	18.60
BMI	44.26	35.02
I B W	72.48	67.09
Deg. obesity %	101.22	59.19
HGB	155.08	149.2
WBC	8.94	7.25
RBC	5.14	5.08
HCT	0.46	0.45
PLT	259.46	232.40
MCV	89.15	88.45
MCH	30.27	29.57
MCHC	339.23	333.67
RDW	21.83	13.63
MPV	8.06	8.49
ALAT	27.54	30.47
ASAT	19.69	26.60
CREA	73.53	81.09
CHOL	5.55	5.17
HDL	1.18	1.47
LDL	3.40	3.00
TRIG	2.19	1.57
CRP	7.82	4.18
OGTT0	5.74	5.85
120OGTT	6.22	6.89
GlyHbA1	5.81	5.75
IRI	26.26	15.98
CPEP	5.92	4.16

Table. 5. Average values for clinical parameters for clusters of female patients

Parameter	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Age	43.6	43.3	41.1	61.1
Height	165.4	164.7	162.4	157.6
Weight	151.3	105.3	78.4	86.7
Fat	50.2	45.5	36.7	43.7
Fat Mass	76.0	48.2	29.3	38.1
FFM	75.4	57.1	49.0	48.6
Muscle mass	71.6	54.2	46.6	46.2
TBW	51.5	41.6	35.6	33.2
TBW %	37.8	39.6	45.1	39.9
Bone Mass	3.8	2.9	2.5	2.5
BMR kJ	10142.0	6337.8	6271.1	6260.1
BMR ccal	2424.0	1783.4	1498.8	1496.2
Metabolic Age	58.6	58.3	50.2	72.2
Visceral Fat	22.8	12.4	7.2	12.6
BMI	55.4	38.9	29.7	34.9
I B W	60.3	59.7	58.4	54.7
Deg. obesity	151.9	76.7	35.4	58.6
HGB	145.4	129.1	133.0	134.0
WBC	10.3	7.9	7.0	6.7
RBC	5.1	4.7	4.4	4.5
HCT	0.4	0.4	0.4	0.4
PLT	290.4	316.8	284.5	250.0
MCV	85.8	83.6	89.2	88.9
MCH	28.8	40.7	29.9	29.7
MCHC	329.0	329.5	335.8	333.8
RDW	14.5	14.4	13.2	13.2
MPV	8.1	8.6	8.2	8.6
ALAT	23.8	24.5	18.2	19.1
ASAT	18.2	22.7	17.9	18.6
CREA	70.9	61.0	64.9	74.0
CHOL	5.3	5.2	5.3	6.1
HDL	1.2	1.4	1.5	1.5
LDL	3.3	3.1	3.3	3.8
TRIG	1.8	1.5	1.2	1.8
CRP	14.8	6.3	4.6	6.9
OGTT0	5.9	5.3	5.1	5.4
120OGTT	6.8	5.2	5.5	6.9
GlyHbA1	6.0	5.6	5.6	5.8
IRI	32.0	18.0	12.9	12.8
CPEP	5.3	3.9	3.6	3.6

The number of cases is 31 out of 85 i.e. the biggest group of female patients. The forth cluster of 29 female patients reveals the pattern of *patients with chronic obesity problem.* This is cluster having relatively high average age but with levels of obesity close to cluster 1. This is proof that obesity is more spread among female patients and already in young age. As in the situation with the male patients the blood, liver and glucose indicators for all obesity patterns do not differ significantly.

For both groups of patients (male and female) the major separation is a result of differences between the obesity indicators, so that they are the only discrimination parameters for the various clusters (patterns of patients).

CONCLUSION

For the first time in the medical practice multivariate statistical analysis was applied for interpretation of clinical data of obesity patients. It was found that after carrying out cluster analysis and principal components analysis specific relationships between the clinical parameters and between the obesity patients could be assessed and modelled. A clear difference between male and female patients is proven. The clinical parameters are definitively divided into two major groups (clusters) combining, on one hand, obesity specific parameters and parameters characterizing the general health status, on another. This general result could help in optimization of the monitoring procedures for obesity sufferers.

Several specific patterns among the female and male patients could be also assessed. In principle, these patterns indicate various levels of obesity and could be used for more detailed treatment of the problem with respect to the patterns identified.

REFERENCES

1. B. Caballero, *Epidemiol. Rev.*, **29**, 1 (2007).
2. R. Sturm, *Public Health*, **121**, 492 (2007).
3. D.W. Haslam, W.P. James, *Lancet*, **366**, 1197 (2005).
4. R.F. Kushner, , D.H. Bessesen, Treatment of the Obese Patient. Contemporary Endocrinology. Totowa, NJ, Humana Press, 2007.
5. J.S. Flier, Cell, **116**, 337 (2004).
6. W. Yang, T. Kelly, He, J., *Epidemiol. Rev.*, **29**, 49 (2007).
7. S. Farooqi, S. O'Rahilly, *Endocr. Rev.*, **27**, 710 (2006).
8. P. Poirier, *Arterioscler. Thromb. Vasc. Biol.*, **26**, 968 (2006).
9. G. Kolata, Rethinking thing: The new science of weight loss– and the myths and realities of dieting. New York: Picador, 2007.
10. A.J. Walley, J.E. Asher, P. Froguel, *Nat. Rev. Genet.***,** **10**, 431 (2009).
11. D. Yach, D. Stuckler, K.D. Brownell, *Nat. Med.***, 12**, 62 (2006).
12. L.A. Barness, J.M. Opitz, E.Gilbert-Barness, *Am. J. Med. Genet. A*, **143A**, 3016 (2007)
13. G.A. Bray, *J. Clin. Endocrinol. Metab.,* **89**, 2583 (2004).
14. H.N. Sweeting, *Nutr. J.*, **6**, 32 (2007).
15. E.A. Greco, R. Fornari, F. Rossi, V. Santiemma, G. Prossomariti, C. Annoscia, A. Aversa, M. Brama, M. Marini, L.M. Donini, G. Spera, A. Lenzi, C. Lubrano, S. Migliaccio, *Int. J. Clin. Pract.***, 64**, 817 (2010).
16. Bosy-Westphal, W. Braun, B. Schautz, M.J. Müller, *Front. Physiol.*, **25**, 4 (2013).
17. K. Rabe, M. Lehrke, K.G. Parhofer, U.C., *Mol. Med*, **14**, 741 (2008).
18. S. Liu, J. Ren, G. Han, G. Wang, G. Gu, Q. Xia, J. Li, *Eur. J. Med. Res.*, **17**, 3 (2012).
19. Y. Elitsur, Z. Lawrence, *Med. J.*, **100**, 67 (2004).
20. V.M. Cambuli, M. Incani, S. Pilia, T. Congui, M.G. Cavallo, E. Cossu, F. Sentinelli, S.Mariotti, S. Loche, M.G. Baroni, *Diabetes Metab. Res. Rev.*, **25** 528, (2009).

МНОГОВАРИАЦИОННА СТАТИСТИЧЕСКА ОЦЕНКА НА КЛИНИЧНИ ПАРАМЕТРИ НА ПАЦИЕНТИ СЪС ЗАТЛЪСТЯВАНЕ

Р. Т. Георгиева-Николова[1], П. А. Гатева[2], Р. К. Хаджийолова[3], М. П. Славова[4], М. М. Николова[5], В. Д. Симеонов[*]

[1]Катедра по химия и биохимия, Медицински факултет, Медицински университет, София, България
[2]Катедра по фармакология и токсикология, Медицински университет, София, България
[3]Катедра по патофизиология, Медицински университет, София, България
[4]Институт по електрохимия и енергийни системи, Българска академия на науките, София, България
[5] Университет Мидълсекс, училище по здравеопазване и образование, Лондон, Великобритания
[6]Катедра по аналитична химия, Факултет по химия и фармация, Софийски университет "Св Кл. Охридски "София, България

Постъпила на 2 декември, 2014 г., коригирана на 22 март, 2015 г.

(Резюме)

Настоящото проучване се отнася за многовариационна статистическа интерпретация на клинични параметри на пациенти със затлъстяване. Целта на проучването е да открие връзки и подобие между традиционно наблюдаваните характеристики при затлъстяване и да определи модели на подобие между пациентите, участващи в изследването. Кластерен анализ и анализ на главни компоненти бяха използвани като многовариационни статистически методи за изследване на данните от проучването, в което бяха включени 113 пациенти. Беше установено, че състоянието на пациентите основно се определя от параметрите, характеризиращи затлъстяването (фактор на телесното тегло, мастна тъкан, телесна маса, степен на затлъстяване и др.) и много по-слабо зависи от параметрите, характеризиращи тяхното общо здравно състояние (общ холестерол, триглицериди, ниво на глюкоза и др.). Това може да помогне за оптимизиране на броя на клиничните променливи, които са необходими за контрол на затлъстяването. Освен това, бяха определени специфични модели на подобие между пациентите и параметрите, отговорни за тяхното формиране. Това дава възможност за по-индивидуално лечение на пациентите. Статистически беше доказано характерното разделяне на мъжете и жените пациенти.

Може да се каже, че за първи път многовариационен статистически анализ е приложен за оценка на здравното състояние на пациенти със затлъстяване.