PLISKA

STUDIA MATHEMATICA BULGARICA

ПЛИСКА

БЪЛГАРСКИ МАТЕМАТИЧЕСКИ СТУДИИ

Pliska Stud. Math. Bulgar. **22** (2013), 143–158

PLISKA
STUDIA MATHEMATICA
BULGARICA

MODELING DATA FOR COMPLICATIONS IN DIABETICS USING LOGISTIC REGRESSION*

K. Prodanova, S. Pashkunova

Abstract. A prospective study of the relationship between some clinical parameters, genetic markers and complications of the patients with diabetes is considered. About 200 patients (male and female) have been examined. The patients are classified into five groups subject to the type of the diabetes. Data obtained for each patient are related to the type of the complications – macro vascular, retina pathology, neuron pathology and nephrite pathology, 12 clinical parameters and 7 genetic markers. Data for the same genetic markers for 94 healthy persons (control group) are compared with those of the diabetics patients. The association of the genetic markers and the different types diabetes-related complications are investigated. A logistic regression to identify which factors are associated with the complications is performed. The associations between pathogenesis and gene genotypes are investigated for the first time for the population of diabetics in Bulgaria.

1. Introduction. Diabetes mellitus (DM) is a chronic (long lasting), metabolic disease which is characterized by high blood sugar (glucose). The human metabolism consists of the most simple breakdown of food to basic molecules that are used to provide energy and building material for body cells. The main energy substrate of humans is glucose. In a healthy person the level of blood glucose is regulated by several hormones, one of which is insulin. Insulin is produced by β-cells of the pancreas. High levels of blood glucose result from lack of production of

*This research is partially supported by pharmaceutical company Novo Nordisk®, Bulgaria.
2010 *Mathematics Subject Classification:* 62P10
Key words: diabetes-related complications, genotypes, logistic regression

the hormone insulin or lack of "response" of the cells to insulin. DM (also called diabetes only) is actually a lifelong disease that causes damage to blood vessels and nervous system after a longer period of time. These injuries are manifested as complaints from different organ systems such as the heart and circulatory system, retina and eye, kidney, peripheral nerves, and others [1, 2, 3, 4, 5, 6, 7].

A survey of four diabetics related complications in patients of the Department of Endocrinology of Military Medical Academy of Bulgaria, Sofia is presented. Some clinical parameters and genetic markers are considered to be prognostic factors of these complications. The patients are classified into five groups subject to the type of the diabetes. A logistic regression is performed in order to identify the clinical risk factors and risk gene genotypes which were associated with the complications. The software package used for statistical modeling of real data was STATISTICA 10 [8].

2. Measurements. During the study period 94 healthy subjects (control group) and 206 patients (male and female) have been prospectively studied. The patients were divided into 5 groups depending on the mode of inheritance, the amount of insulin secretion, treatment and course of illness:

- Insulin-dependent type 1 diabetes (juvenile diabetes) – 43 patients;
- Insulin- not dependent type 2 diabetes mellitus (DM Age) – 60 patients;
- Insulin-dependent type 2 diabetes mellitus – 42 patients;
- Diabetes type MODY – 29 patients;
- Diabetes mellitus type LADA – 32 patients.

Data obtained for each patient included 12 clinical parameters as: Age (in years), Sex, Glucosed Hemoglobin (HbAlc%), C-peptide, Creatinin, micro albumin, C-reactive Protein (CRP), Cholesterol, three-gliceridies (TRG), high density lipoprotein (HDL), low density lipoprotein (LDL), Duration (in years). The complications which are observed are retinopathy, neuropathy, nephropathy and macro vascular.

The primary data analysis was as follows: minimum age was 19 and maximum – 81 years, 35% of patients were men and 65% women, 23% was observed retinopathy, minimum of duration is 1 year and maximum was 40 years, 75% had neuropathy, in 21% nephropathy was observed and 18% was observed with macro vascular complications. The sum of percentages is more than 100 because there are patients with 2, 3 or 4 complications. In the study we also obtain data for

genotypes of several genetic markers – Calpain, Leptin, MnSod, eNos, eNosGly, LPL and AC in the patients and in the control group. The information about mean values and standard deviations of the clinical parameters are presented in Tables 2-7. More detailed information about the frequency distributions of gene markers are published in [19].

An individual's genotype is the combination of alleles found in that individual at a given genetic locus [9]. Allele is a variant of the DNA level. In biological systems, each parent transmits one allele with germ cells. So each person has two alleles from their parents.

If there are two alleles in a population at locus A (denoted by $\boldsymbol{A}$ and $\boldsymbol{a}$), then the possible genotypes in that population are $\boldsymbol{AA}$, $\boldsymbol{Aa}$, and $\boldsymbol{aa}$. Individuals with genotypes $\boldsymbol{AA}$ and $\boldsymbol{aa}$ are homozygous (i.e., they have two copies of the same allele). Individuals with genotype $\boldsymbol{Aa}$ are heterozygous (i.e., they have two different alleles at the A locus). If the heterozygote is phenotypically identical to one of the homozygous, the allele found in that homozygote is said to be dominant, and the allele found in the other homozygote is recessive.

Allele demonstrates the power of gene mutation. If the allele is recessive, it will be associated with a phenotype – for example have or not a neurological complication only when is in a homozygous state – $\boldsymbol{aa}$ or $\boldsymbol{AA}$. So in these phenotypic characteristics, which are given the absence or presence of symptoms - nephrology, ophthalmology, etc., may be associated with dominant or recessive mark.

In the nomenclature there are several signs for denoting the different allels: "+" and "−" , $\boldsymbol{T}$ and $\boldsymbol{L}$, $\boldsymbol{T}$ and $\boldsymbol{G}$, $\boldsymbol{G}$ and $\boldsymbol{A}$, etc.

The comparison of the genotypes distribution of considered gene markers in our control group and in a healthy subjects from other countries shows that the distribution is similar to other European populations [19].

3. Statistical Methods and Models.

3.1. Hardy-Weinberg equilibrium model. The mathematical concept, called the Hardy-Weinberg equilibrium model or principle (HWP), is a crucial concept in the population genetics [10, 11]. It predicts how gene frequencies will be inherited from generation to generation given a specific set of assumptions.

The definition of HWP is the following [10, 11]: The stable frequency distribution of genotypes, $\boldsymbol{AA}$, $\boldsymbol{Aa}$, and $\boldsymbol{aa}$, in the proportions p^2, $2pq$, and q^2 respectively (where p and q are the frequencies of the alleles, $\boldsymbol{A}$ and $\boldsymbol{a}$, $p+q=1$) that is a consequence of random mating in the absence of mutation, migration, natural selection, or random drift.

This principle is important because it gives biologists a standard for mea-

suring changes in allele frequency in a population. When a population meets all of the Hardy-Weinberg conditions, it is said to be in Hardy-Weinberg equilibrium (HWE). Human populations do not meet all of the conditions of HWE exactly, and their allele frequencies will change from one generation to the next and the population will evolve. How far a population deviates from HWE can be measured using the χ^2 "goodness of fit" statistical test.

To test null hypothesis H_0: {The frequencies of the genotypes in the control group and in the group of patients with some type of diabetes are in Hardy-Weinberg equilibrium} we use the statistic χ^2 with 2 degrees of freedom at the 5%-level.

To test the null hypothesis H_0: {The frequencies of allels of given marker in the control group and in the group of patients with some type of diabetes are equally likely} we use Fisher's exact test for association [12, 13].

Table 1. Sample characteristics of clinical parameters for Type1 ($N = 43$)

	Mean	**Minimum**	**Maximum**	**Std.Dev.**
age	43.28	21.00	81.00	14.93
HbA1C	8.41	5.06	12.80	1.70
c-peptide	0.51	0.09	2.77	0.67
creatinin	4.98	0.80	41.10	7.96
micro albumin	85.58	51.00	128.00	17.37
CRP	68.51	0.00	626.00	135.48
Cholesterol	5.81	3.13	9.62	1.47
TRG	1.60	0.60	3.80	0.76
HDL	1.77	0.80	8.30	1.48
LDL	3.55	0.80	7.76	1.44

3.2. Dichotomous Response Models. Logistic regression model. The regression model is used to relate a categorical response (dependent variable $\boldsymbol{Y}$) to the explanatory variables (predictors) $\boldsymbol{x}_i (i = 1, \ldots, c)$. We are interested in the presence (coded by 1) or the absense (coded by 0) of the four categorical responses

- Retinopathy,
- Neuropathy,
- Nephropathy,
- Macro vascular complications.

Table 2. Sample characteristics of clinical parameters for Type2 ($N = 60$)

	Mean	**Minimum**	**Maximum**	**Std.Dev.**
age	62.47	48.00	78.00	7.79
HbA1C	7.79	5.50	11.56	1.29
c-peptide	2.89	1.10	9.61	1.23
creatinin	89.85	59.00	128.00	15.55
micro albumin	22.64	0.00	300.00	52.67
CRP	4.46	0.77	14.00	2.80
Cholesterol	5.98	1.90	10.20	1.51
TRG	2.31	0.80	6.14	1.13
HDL	1.26	0.60	4.20	0.60
LDL	4.65	1.42	9.00	1.69

Table 3. Sample characteristics of clinical parameters for Type2_insulun ($N = 33$)

	Mean	**Minimum**	**Maximum**	**Std.Dev.**
age	57.70	44.00	73.00	7.78
HbA1C	9.67	6.10	13.80	1.72
c-peptide	87.73	56.00	192.00	25.15
creatinin	88.19	57.00	110.00	13.59
micro albumin	24.94	0.00	285.00	50.31
CRP	5.51	1.10	37.90	8.50
Cholesterol	5.75	2.59	9.40	1.47
TRG	1.85	0.60	8.06	1.45
HDL	1.13	0.26	1.70	0.34
LDL	4.01	1.78	8.70	1.52

At the first part predictors are 12 clinical parameters and the considered models are 20 (4 complications in 5 groups of DM). At the second part the explanatory variables are genotypes of the patients for 7 markers.

About the Dichotomous Response Models [15, 16, 17]: Let us define the dummy random variable to indicate the two categories by $Y = 1$ for category A and $Y = 0$ for B. The probability density for Y given the parameter p is therefore point binomial

$$f(Y/p) = p^Y (1-p)^{(1-Y)}. \tag{1}$$

Table 4. Sample characteristics of clinical parameters for LADA ($N = 42$)

	Mean	**Minimum**	**Maximum**	**Std.Dev.**
age	65.62	50.00	80.00	8.32
HbA1C	9.01	6.30	12.40	1.44
c-peptide	2.13	0.11	6.95	1.36
creatinin	96.10	59.00	256.00	34.22
micro albumin	95.74	0.00	831.00	186.73
CRP	7.77	0.50	47.70	8.07
Cholesterol	5.65	3.53	8.36	1.22
TRG	2.29	0.42	6.04	1.28
HDL	1.23	0.65	5.20	0.68
LDL	3.76	1.10	7.80	1.69

Table 5. Sample characteristics of clinical parameters for MODY ($N = 31$)

	Mean	**Minimum**	**Maximum**	**Std.Dev.**
age	41.00	19.00	68.00	10.44
HbA1C	8.29	4.70	11.40	1.93
c-peptide	3.41	0.98	6.50	1.27
creatinin	88.19	57.00	110.00	13.59
micro albumin	152.09	0.00	3600.00	658.04
CRP	5.27	0.20	23.00	4.92
Cholesterol	5.65	2.10	10.60	1.67
TRG	2.30	0.76	5.36	1.22
HDL	1.17	0.60	1.90	0.31
LDL	3.82	1.03	9.10	1.60

Table 6. Distribution of the gender in the different groups

Type of Diabetes	female	male
Type1	28	15
Type2	47	13
Type2-insulin	29	13
MODY	14	17
LADA	18	15

We assume that the probability p depends on a linear function

$$d(x_1, \ldots, x_c) = \beta_0 + \sum_{i=1}^{c} \beta_i x_i \tag{2}$$

where $x_i (i = 1, \ldots, c)$ are the explanatory variables (the independent variables), β_i are the constants. So, the joint conditional density is

$$f(y_1, \ldots, y_n / p(d_1), \ldots, p(d_n)) = \Pi_j^n [p(d_j)]^{y_i} [1 - p(d_j)]^{(1-y_j)}, \tag{3}$$

where n is size of a random sample of data y for response variable Y.

In order to be able to relate the value of y to the value of d a most specific assumption about the form of $p(d)$ is required. In the so called logit or logistic model [5] the distribution function of logistic density is:

$$p(d) = \frac{e^d}{1 + e^d}. \tag{4}$$

The shape of $p(d)$ (logistic distribution) is quite similar to the shape for normal distribution. The odds of a dichotomous response is given by

$$o\,d\,d\,s = \left[\frac{p(d)}{1 - p(d)}\right]. \tag{5}$$

The logit transformation

$$\ln \left[\frac{p(d)}{1 - p(d)}\right] = d = \beta_0 + \sum_{i=1}^{c} \beta_i x_i \tag{6}$$

gives an important advantage of the model because Eq.(6) is a linear function of the explanatory variables. The Newton-Raphson iterative procedure is usually used to make maximal likelihood estimator $\hat{\beta}$ of coefficient vector β in the logistic model. The procedure is based on the preliminary estimator of β given by $\hat{\beta} = (X^T X)^{-1} X^T Y$ where Y is the vector of y_i response values $(i = 1, \ldots, n)$ and $X(n \times c)$ is the matrix of observations. The maximum likelihood is obtained by solving the system of $(c + 1)$ equations:

$$\sum_{i=1}^{n} p_i x_i = \sum_{i=1}^{n} y_i x_i \tag{7}$$

where

$$p_i = \exp(x_i^T \beta)/(1 + \exp(x_i^T \beta)). \tag{8}$$

The solutions to these equations given by $\hat{\beta}$ can be used to obtain the estimator $\hat{p}_i$ for each of n observations and hence the fitted sum $\sum_{i=1}^{n} \hat{p}_i x_i$ is equal to the observed sum in the right side of (7). In comparison to the multiple linear regression model, the coefficient vector $\hat{\beta}$ must be interpreted differently:

- The coefficients $\hat{\beta}$ were interpreted as estimates of *log odds* .
- A marginal one unit increase in x_j brings an increase in d (i.e., in *log odds*) of the amount $\hat{\beta}$.

Testing of hypothesis concerning the regression parameters can include a test of single parameter, a test involving several parameters from the same regression, and joint tests involving parameters from different regressions. In the polychotomous logistic regression, tests for contribution of one or more parameters from the same regression are usually constructed with a large sample Wald test, with test statistic

$$Q_W = \hat{\beta}^T [\mathrm{Var}(\hat{\beta})]^{-1} \hat{\beta}, \tag{9}$$

where $\mathrm{Var}(\hat{\beta})$ is the estimated covariance sub-matrix for the relevant parameters. This statistics is approximately distributed as a $\chi^2(r)$ random variable with r degrees of freedom under the null hypothesis that r-dimensional vector $\hat{\beta}$ is equal $\vec{0}$. When there is a single parameter of interest the test statistic is

$$Q_W = [\hat{\beta}_j / SE(\hat{\beta}_j)]^2 \tag{10}$$

($SE(\hat{\beta}_j)$ is the standard error of $\hat{\beta}_j$) and its distribution is $\chi^2(r = 1)$.

4. Application to the real data. Results and conclusions. To test the null hypothesis, that the frequencies of the genotypes in the control group and in all other groups of diabetics are in Hardy-Weinberg equilibrium, we use the package WinStat 4.3 [14]. In Table 8 the result of the test for the marker eNOS in the control group is given. The conclusion is that the differences between expected and observed frequencies is not statistically significant at the level of significance $\alpha = 0.05$, so the genotypes of marker eNOS in the control sample meets the Hardy-Weinberg equilibrium.

Table 7. Hardy-Weinberg equilibrium of genotypes for marker eNOS in the control group

Genotype eNOS Control group	Expected	Observed
Common homozygote (***TT***)	60	61
Heterozygote (***TL***)	30.99	29
Rare homozygote (***LL***)	4	5
p-allele ***T*** frequency	0.79	
P-level of $\chi^2(2)$	0.39	

The same tests for HWE about all other 7 markers are made for the control group and for the other five groups of diabetics. For all samples we derive to be in Hardy- Weinberg equilibrium [19].

Fisher's exact test for association between the control group and each of the other five groups of diabetics (for all 7 markers) have been carried out.

In Table 9 the result of Fisher's exact test for Insulin-dependent type 1 diabetes is showed.

Table 8. Fisher's exact test for marker eNOS in type 1 DM

	Column 1 Alel ***T***	Column 2 Alel ***L***	Row-Totals
Frequencies Control Group	151	39	190
Frequencies Type 1	65	21	86
Column totals	216	60	276
Fisher exact P, one-tailed		p= 0.282	
Fisher exact P, two-tailed		p= 0.529	

The two-tailed P value equals 0.529 and the conclusion is that the association between rows (groups) and columns (outcomes) is not statistically significant at 5%-level for marker eNOS in Type 1 diabetes.

The result of Fisher's exact test is that the association is statistically significant for the gene marker MnSod in patients of dibetes type LADA:

Table 9. Fisher's exact test for marker MnSod in LADA DM

	Allele (+)	Allele (−)	Total	P-value, two-tailed
LADA	34 (60.7%)	22 (39.3%)	56	$P = 0.042$
Control group	84 (44.7%)	104 (55.3%)	188	

Our result for MnSod in LADA DM is similar to the one in [2]. The results of all other Fisher's exact tests have shown that the association is not statistically significant at 5%-level as it is in [4, 6, 7]. Our results for eNOS are different from [1] and [5], where the association is discovered to be significant.

Four logistic regressions for each of 5 groups DM have been considered, to identify which of 12 clinical parameters (factors) were associated with the complications (macro vascular, retinopathy, neuropathy and nephropathy). The resulting models for retinopathy (for the different types of diabetics) are presented in Table 11.

Table 10. The statistical significant clinical factors associated with the complication retinopathy identify by the logistic regression

Logistic regression models: The complication Retinopathy as a function of clinical factors						
Type Diabetics	Factor	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
TYPE1	creatinin	1.487	−0.166	0.046	0.84	0.69–1.09
	duration		−0.133	0.009	0.88	0.81–0.97
TYPE 2 – insulin	c-peptide	1.117	−0.767	0.045	0.047	0.21–1.03
MODY	cholesterol	−8.135	0.822	0.022	2.28	0.91–5.72
	LDL		1.134	0.033	3.11	0.83–11.63
LADA	duration	−4.062	0.341	0.014	1.41	1.06–1.86

From these results and according to the remarks in Subsection 3.3 we can infer that the probability for retinopathy is higher when:

- The predictors "Creatinin" and "Duration" are lower in the patient with type 1 DM;
- The predictor "C-peptide" is lower the patient with type 2-insulin DM;
- The predictor "Cholesterol" is higher in the patient with type MODY DM;
- The predictor "Duration" is higher in the patient with type LADA DM.

For the factor "Duration", using (see [15]) minimum and maximum values in the corresponding samples, we estimate that for type 1 DM (min = 1 and max = 40 years) with increase of years, the probability of complication "retinopathy" decreased from 0.96 to 0.11, as for type LADA (min = 1 and max = 20 years) this probability increases from 0.02 to 0.93.

The same type of analysis for macro vascular complications, neuropathy and nephropathy have been carried out and published in Tables 11-13. Similar results are published in [1], [2], [4], [18].

Table 11. The statistical significant clinical factors associated with the complication neuropathy identify by the logistic regression

Logistic regression models: The complication neuropathy as a function of clinical factors						
Type Diabetics	Factor	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
TYPE1	duration	0.173	−0.071	0.049	0.93	0.85–1.01
TYPE2	gender	−1.584	1.429	0.035	4.17	1.07–16.24
	creatinin		0.041	0.048	1.04	0.99–1.09
TYPE 2 – insulin	cholesterol	−4.781	2.301	0.039	9.99	1.04–19.00
	TRG		0.761	0.048	2.13	0.98–4.66
MODY	age	−7.517	0.181	0.015	1.19	1.02–1.39
	duration		0.428	0.005	1.53	1.12–2.01
LADA	duration	0.533	−0.449	0.043	0.64	0.41–1.04

Table 12. The statistical significant clinical factors associated with the macro vascular complications identify by the logistic regression

Logistic regression models: Macro vascular complications as a function of clinical factors						
Type Diabetics	Factor	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
TYPE1	duration	−4.552	0.118	0.033	1.12	1.01–1.27
	age		0.091	0.026	1.09	1.01–1.19
	creatinin		1.121	0.023	1.13	1.02–1.26
TYPE 2 – insulin	LDL	−2.011	0.485	0.032	1.62	1.03–2.57

The main part of study is to investigate the association of the genetic markers with the different types diabetes-related complications. For this purpose we make two types of code clustering:

Table 13. The statistical significant clinical factors associated with the complication nephropathy identify by the logistic regression

Logistic regression models: The complication nephropathy as a function of clinical factors						
Type Diabetics	Factor	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
TYPE1	c-peptide	1.443	−1342	0.034	0.26	0.07–0.94
TYPE 2	c-peptide	−28.236	5.284	0.045	3.82	1.03–1.42
	micro albumin		0.123	0.016	1.05	1.01–1.09
MODY	cholesterol	−8.293	0.823	0.047	2.28	0.92–5.73
	LDL		1.134	0.048	3.11	0.83–11.63
LADA	micro albumin	−3.749	0.069	0.016	1.07	1.01–1.14

- genotypes ***AA*** and ***Aa*** (code1) versus ***aa*** (code2),
- genotypes ***aa*** and ***Aa*** (code1) versus ***AA*** (code2),

and than the respective 40 (4 complications in 5 groups and 2 types code clusters) dichotomous response models with 7 predictors (markers) are built.

It turns out (see [4, 5]) that if we find differences in these two groups, we will know if allele ***A*** or allele ***a*** is dominant.

The association between pathogenesis and gene genotypes of all types of diabetes are analysed using logistic regression. The results, when we assumed a recessive model of inheritance (i.e., ***AA*** and ***Aa*** versus ***aa***) for the group of diabetes type 2, are given in Table 12. Lack of an increased risk of complication posed by the ***A*** allele is not dominantly expressed and that the increased risk is confined to ***aa*** homozygotes for:

- Calpain of neuropathy and macro vascular complications,
- LPL of nephropathy,
- MnSod of neuropathy,

Table 14. Models of inheritance ***AA*** and ***Aa*** versus ***aa*** for the group of type 2 DM for diabetes-related complications

Diabetics type 2: Genotypes ***AA*** and ***Aa*** (code1) versus ***aa*** (code2):						
Complication	Gene marker	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
NEPHRO	Calpain	−5.128	2.014	0.041	7.5	1.1–53.7
	LPL		2.1	0.04	8.1	1.01-66.9
NEURO	MnSod	−1.386	1.098	0.044	10.2	2.1–51.1
MACRO	Calpain	−4.678	1.992	0.019	7.3	1.3–40.6
	eNos		2.322	0.035	10.2	1.1–93.1
RETINO	eNosGly	−4.388	1.471	0.011	4.3	1.2–27.1
	AC		2.14	0.041	4.1	1.1–16.2

Table 15. Models of inheritance ***aa*** and ***Aa*** versus ***AA*** for the group of type 2 DM for diabetes-related complications

Diabetics type 2: Genotypes ***aa*** and ***Aa*** (code1) versus ***AA*** (code2):						
Complication	Gene marker	$\hat{\beta}_0$	$\hat{\beta}_i$	p-level (Wald's χ^2)	odds ratio (unit change)	odds rat. 95% conf. interval
NEPHRO	AC	0.209	−1.868	0.014	1.8	1.1–3.9
RETINO	eNOS	−4.738	1.742	0.047	1.6	1.0–7.7

- eNosGly of retinopathy
- AC of retinopaty and nephropathy.

By contrast, the increased risk of retinopathy (Table 8) is confined to eNOS ***AA*** homozygotes for code cluster "***aa*** and ***Aa*** versus ***AA***".

Our results imply that homozygosity for the eNos may be involved in predisposition to retinopathy and macro vascular complications.

The results of association between pathogenesis and gene genotypes in other groups of DM patients are published in [19] and are similar as in [1, 2, 3, 4, 5, 6, 7].

Using logistic regression the association between pathogenesis and gene genotypes of all other types of DM are analysed and published in [19]. To our knowl-

edge, this is the first study in Bulgaria to implicate polymorphisms of Calpain, Leptin, MnSod, eNOS, eNosGly, LPL and AC as a genetic risk factor for retinopathy, neuropathy, nephropathy and macro vascular complications.

In conclusion, we believe that our findings could contribute to any future analyses validating these relationship.

Acknowledgements The authors would like to thank the referee for the valuable comments on the paper.

REFERENCES

[1] E. Noiri, H. Satoh, J. Taguchi, S. V. Brodsky, A. Nakao, Y. Ogawa, S. Nishijima, T. Yokomizo, K.Tokunaga, T. Fujita. Association of eNOS Glu298Asp Polymorphism with End-Stage Renal Disease. *Hypertension* **40** (2002), 535–540.

[2] M. Tarek, E. Masry, A.M. Mona, Z. Abou, M.Mohamed, E.Tawil, A.Rawhia, A.Khalifa. Manganese Superoxide Dismutase Alanine to Valine Polymorphism and Risk of Neuropathy and Nephropathy in Egyptian Type 1 Diabetic Patients. *Review of Diabetic Studies* **2** (2005), 70–74.

[3] F. Thameem, S. Puppala, N. H.Arar, M. P.Stern, J. Blangero, R. Duggirala, H. E. Abboud. Endothelian nitric oxide synthase gene polymorphisms and their association with type 2 diabetes-related traits in Mexican Americans. *Diabetes and Vascular Disease Research* **5** (2008), 109–113.

[4] A. Zanchi, D. K. Moczulski, L. S. Hanna, M. Wantman, J. H. Warram, A. S. Krolewski. Risk of advanced diabetic nephropathy in type 1 diabetes is associated with endothelial nitric oxide synthase (eNOS) gene polymorphism. *Kidney International* **57** (2000), 405–413.

[5] K. Hibi, T. Ishigami, K. Tamura, S. Mizushima, K. Kimura, M. Ishii, S. Umemura, T. Fujita, H. Ochiai, M. Kosuge, Y. Watanabe. Infarction Endothelial Nitric Oxide Synthase (eNOS) Gene Polymorphism and Acute Myocardial. *Hypertension* **32** (1998), 521–526.

[6] L. Haddad, J. C. Evans, N. Gharani, C. Robertson, K. Rush, S. Wiltshire, T. Frayling, T. Wilkin, A. Demaine, A. Millward,

A. Hattersley, G. Conway, N. Cox, G. Bell, S. Franks, M. Mccarthy. Variation within the Type 2 Diabetes Susceptibility Gene Calpain10 and Polycystic Ovary Syndrome. *The Journal of Clinical Endocrinology & Metabolism* **87**(2002), No 6, 2606–2610.

[7] J. Vaškù, A. Vaškù, Z. Dostálová, P.Bienert. Association of Leptin genetic polymorphism with gestational diabetes mellitus. *Genes & nutrition* **1** (2006), No 2, 117–124.

[8] STATISTICA 10. USA, StatSoft Inc., 2010.

[9] F. Griffiths, J. F. Anthony An introduction to genetic analysis. 6th edn. New York, W. H. Freeman, 1998.

[10] G. H. Hardy. Mendelian Proportions in a Mixed Population. *Science, New Series* **28**(706) (1908), 49–50.

[11] W. Weinberg. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte des Vereins für vaterländische Naturkunde in Württemberg* **64** (1908), 368–382.

[12] R. A. Fisher. A New Test for 2×2 Tables. *Nature* **156** (1945), 783–784.

[13] R. M. Cyrus, P. Senchaudhuri. Conditional versus Unconditional Exact Tests for Comparing Two Binomials, (2003). Retrieved 20 November 2009 from `http://www.cytel.com/Papers/twobinomials.pdf`.

[14] H. M. Stauss. Doing Statistics with WinStat. Univ. of Iowa, 2008, 1–75.

[15] J. D. Jobson. Applied Multivariate Data Analysis, vol.II, Springer Verlag, 1991, 278–303.

[16] D. Atanasov, K. Prodanova, D.Tersiiski. Risk factors of central venous catheters related ifections in intensive care patients. *Pliska Stud. Math. Bulgar.* **14** (2003), 15–22.

[17] K. Prodanova, I. Stoinov, D.Terziiski. Logistic regression in modelling data for CVC related infection. *Pliska Stud. Math. Bulgar.* **17** (2005), 221–228.

[18] I. Larson, M.M. Hoffmann, J.M. Ordovas, E.J. Schaefer, W. Marz, J. Kreuzer. The lipoprotein lipase HindIII polymorphism: association with total cholesterol and LDL-choleterol, but not with HDL and triglycerides in 342 females. *Clin. Chem.* **45** (1999), 963–968.

[19] S. Pashkunova. Associative investigations of genetic markers determing type of dibetes and chronic diabetes-related complications. PhD thesis, Military Medical Academy of Bulgaria, Sofia, 2012, pp.1–35 (in Bulgarian).

K. Prodanova
Technical University of Sofia
e-mail: kprod@tu-sofia.bg

S. Pashkunova
Military Medical Academy of Bulgaria
e-mail: pashkunovasylvia@yahoo.com