

# Predicting the development and progression of primary open-angle glaucoma based on the determination of GST gene polymorphisms

S.O. Rykov<sup>1</sup>, Dr Sc (Med), Prof., A.V. Burdei<sup>1</sup>, Postgraduate, S.V. Ziablitsev<sup>2</sup>, Dr Sc (Med), Prof., S.Yu. Mogilevskyy<sup>1</sup>, Dr Sc (Med), Prof.

<sup>1</sup> Shupik National Medical Academy of Postgraduate Education;  
Kyiv (Ukraine)

<sup>2</sup> Bohomolets National Medical University  
Kyiv (Ukraine)

E-mail: sergey.mogilevskyy@gmail.com

**Background:** Epidemiological studies have demonstrated that primary open-angle glaucoma (POAG) is a complex multifactorial disease resulting from interaction of risk factors and gene polymorphisms. Previously, we have demonstrated the role of deletion polymorphisms of glutathione-S-transferase (GST) genes in the development and progression of POAG.

**Purpose:** To predict the speed of POAG development and the speed of POAG progression as early as at patient presentation based on determination of GST gene (GSTT1, GSTP1 and GSTM1) polymorphisms.

**Materials and Methods:** Data of 172 patients (344 eyes) diagnosed with stage 1, 2, 3 or 4 POAG, and of 98 control group volunteers (196 eyes) diagnosed with no POAG, were included in the analysis. Two indices, SDPOAG (speed of POAG development, expressed as stage/life year), and SPPOAG (speed of POAG progression, expressed as stage/disease year), were used as measures for assessing speed of the pathological process. Polymorphisms of GSTP1, GSTM1 and GSTT1 were investigated by real-time polymerase chain reaction (PCR) using TaqMan Mutation Detection Assays (Life Technologies, Carlsbad, CA) and Gene Amp® 7500 PCR System (Applied Biosystems, Foster City, CA).

**Results and Discussion:** We built an algorithm (computer program) for predicting the speed of the development and progression of POAG, with the potential to calculate these indices as early as at presentation. The GSTP1 Ile105Val polymorphism (with the presence of the mutant Val allele) and GSTM1 and GSTT1 deletion polymorphisms contributed more to prediction of dependent variables, SDPOAG and SPPOAG, than others. We found a significant effect of the combination of independent variables in the regression equation on the dependent variable SDPOAG, with a multiple correlation coefficient  $R$  of 0.801, a coefficient of determination  $R^2$  of 0.622, and an  $F$  value of 77.338 at  $p < 0.0001$ . In addition, the regression equation for prediction for the speed of the progression of POAG had a multiple correlation coefficient  $R$  of 0.857, a coefficient of determination  $R^2$  of 0.734, and an  $F$  value of 73.450 at  $p < 0.0001$ .

**Conclusion:** The objectives of the study (i.e., to predict the age of onset of clinical manifestations of POAG and the speed of the development of POAG, with the potential to calculate them as early as at presentation) have been met.

## Keywords:

primary open-angle glaucoma, GSTP1, GSTM1 and GSTT1 polymorphisms, speed of the development and progression

## Introduction

Primary open-angle glaucoma (POAG) accounts for around 80% of the total glaucoma cases and is a major cause of low vision and blindness among adults worldwide [1, 2]. By 2020, the number of people with open-angle glaucoma in the world is expected to increase to about 60 million [2]. The prevalence of POAG for 40-45 year olds is only about 0.7%, and it increases to 1.5-2% for 40-45 year olds, and to 10% for individuals aged  $\geq 75$  years [3, 4].

Risk factors for POAG can be divided into two categories: systemic and local. Systemic risk factors include

old age, a racial or genetic predisposition to the disease, vascular disorders, hypercholesterolemia, hypothyroidism, diabetes mellitus, and smoking. Local risk factors include anatomic features of the optic nerve, aqueous fluid drainage system, and ocular vascular structures, and the presence of concomitant eye disease [4, 5].

Epidemiological studies have demonstrated that POAG is a complex multifactorial disease resulting from interaction of risk factors and gene polymorphisms

[6, 7]. The percentage influence of genetic factors on the susceptibility to develop POAG has been reported to vary from 20% to 60% [8, 9]. Previously, we have demonstrated the role of deletion polymorphisms of glutathione-S-transferase (GST) genes in the development and progression of POAG.

Findings of molecular and genetic studies on the role of genetic factors in pathogenesis of a disease are utilized through models developed for predicting the progression of the disease or its complications. As individual genotype is invariable from life to death, accuracy of prediction using such models is rather high, and predictions will be relevant for long. Because our previous findings [10] have confirmed associations between polymorphic GST genotypes and POAG, we find it expedient to develop models describing the effect of genotype on disease progression and/or predicting the age at which the disease will progress to stage 2, 4 or 4.

**The purpose** of the study was to predict the speed of POAG development and the speed of POAG progression as early as at patient presentation based on determination of *GST* gene (*GSTT1*, *GSTP1* and *GSTM1*) polymorphisms.

#### Materials and Methods

Stepwise, multifactorial regression models were built using the General Regression Model (GRM) package. Data of 172 patients (344 eyes) diagnosed with stage 1, 2, 3 or 4 POAG, and of 98 control group volunteers (196 eyes) diagnosed with no POAG, were included in the analysis. Glaucoma was staged according to the classification of Nesterov and Bunin (1976) and the classification of perimetry changes related to the stages of glaucoma.

The study was conducted in accordance with the fundamental principles of the European Convention on Human Rights and Biomedicine, the Declaration of Helsinki, and the Order of Ministry of Health of Ukraine No 690 dated 23 September 2009. Patients with POAG and controls were familiarized with the study purpose and tasks.

Two indices, SDPOAG (speed of POAG development, expressed as stage/life year), and SPPOAG (speed of POAG progression, expressed as stage/disease year), were used as measures for assessing speed of the pathological process. These indices increased with increasing stage of the pathological process ( $p < 0.05$ ).

Polymorphisms of *GSTP1* were investigated by real-time polymerase chain reaction (PCR) using TaqMan Mutation Detection Assays (Life Technologies, Carlsbad, CA) and Gene Amp® 7500 PCR System (Applied Biosystems, Foster City, CA). Deletion polymorphisms of *GSTT1* and *GSTM1* were investigated by PCR with electrophoretic detection using AmpliSens® *GSTT1*/*GSTM1*-EPH kits (Interlabservis, Russia).

Statistical analyses were conducted using Statistica 10.0 (StatSoft, Tulsa, OK, USA) software.

#### Results and Discussion

In order to ensure that the developed models are used in practice, independent variables (predictors) were selected based on the following main criteria: exhibiting either no change or consistent pattern of changes during patient's life. In addition, patient's data relevant to predictors had to be easily available at presentation. Therefore, the selected predictors included the sex, age and the results of genotyping for the investigated *GST* polymorphisms having a proven effect on the course of the disease [10].

Predictor weights were taken into account when building regression models. Table 1 demonstrates the correspondence between either predictors or dependent variables, and categorical, continuous, or dichotomous values that were used to build regression equations.

The coefficients of regression equations for dependent variables (SDPOAG and SPPOAG) and the significance of difference from the null hypothesis are shown in Tables 2 and 3, respectively.

Analysis of beta coefficients of regressions found that, of all predictors considered, the *GSTT1*, *GSTP1*, and *GSTM1* contributed more to prediction of dependent variables, SDPOAG ( $0.0047 \pm 0.0013$ ,  $0.0038 \pm 0.0009$ , and  $0.0036 \pm 0.0012$ , respectively), and SPPOAG ( $0.4361 \pm 0.0937$ ,  $0.3212 \pm 0.0632$ , and  $0.2557 \pm 0.0854$ , respectively), than others. These predictors demonstrated direct associations with dependent variables.

Somewhat another situation was observed with the two other predictors. The contribution of the Sex variable both to SDPOAG and SPPOAG were greater than that of the Age variable ( $0.0007 \pm 0.0001$  vs  $0.0001 \pm 0.0003$  and  $0.0629 \pm 0.0199$  vs  $0.0014 \pm 0.0010$ , respectively), and these predictors demonstrated opposite associations with dependent variables.

In addition, analysis of the significance of difference from the null hypothesis for beta coefficients of predictors found that, due to its low significance ( $t = -1.559$ ;  $p = 0.060$ ), the Age variable may be not considered when calculating the regression equation for SDPOAG.

Pareto charts in Figs. 1 and 2 demonstrate statistical significances of regression beta coefficients for dependent variables, SDPOAG and SPPOAG, respectively, with the predictors arranged in descending order for Student test  $t$  values.

Therefore, a formula for calculation of SDPOAG can be presented as:

$$SD_{POAG} = -1.28134 + 0.00067 * S + 0.00382 * P1 + 0.00356 * M1 + 0.00472 * T1 \quad (1),$$

where  $S$  is a dichotomous value assigned to the Sex variable,

$P1$  is a dichotomous value assigned to the *GSTP1* variable,

$M1$  is a dichotomous value assigned to the *GSTM1* variable,

$T1$  is a dichotomous value assigned to the *GSTT1* variable.

At large, we found a significant effect of the combination of independent variables on the dependent variable SDPOAG ( $R^2=0.622$  at  $F=77.338$ ,  $p<0.0001$ ). The  $R$  was 0.801, indicating that approximately 80.1% of the variance in SDPOAG was accounted for by the combination of independent variables in the model.

A regression formula for calculation of SPPOAG can be presented as:

$$SP_{POAG} = -95.5044 - 0.0629 * S + 0.014A + 0.3212 * P1 + 0.2557 * M1 + 0.4361 * T1 \quad (2),$$

where  $S$  is a dichotomous value assigned to the Sex variable,

$A$  is a continuous value assigned to the Age variable,

$P1$  is a dichotomous value assigned to the *GSTP1* variable,

$M1$  is a dichotomous value assigned to the *GSTM1* variable,

$T1$  is a dichotomous value assigned to the *GSTT1* variable.

The regression equation for prediction for the speed of the progression of POAG had a multiple correlation coefficient  $R$  of 0.857, a coefficient of determination  $R^2$  of 0.734, and an  $F$  value of 73.450 at  $p<0.0001$ , pointing to achievement of acceptable results for the model built for predicting SPPOAG.

In addition, table 4 presents the results of the analysis of differences between actual and calculated values for dependent variables SDPOAG and SPPOAG. Moreover, Figs 3 and 4 graphically demonstrate differences between actual and calculated values for SDPOAG and SPPOAG, respectively.

Therefore, the objectives of the study (i.e., to predict the age of onset of clinical manifestations of POAG and the speed of the development of POAG, with the potential to calculate them as early as at presentation) have been met.

## Conclusion

We used regression analysis to build an algorithm (computer program) for predicting the speed of the development and progression of POAG, with the potential to calculate them as early as at presentation. Of all predictors

considered, the *GSTT1*, *GSTP1*, and *GSTM1* contributed more to prediction of dependent variables, SDPOAG ( $0.0047 \pm 0.0013$ ,  $0.0038 \pm 0.0009$ , and  $0.0036 \pm 0.0012$ , respectively), and SPPOAG ( $0.4361 \pm 0.0937$ ,  $0.3212 \pm 0.0632$ , and  $0.2557 \pm 0.0854$ , respectively), than others. In addition, we found a significant effect of the combination of independent variables in the regression equation on the dependent variable SDPOAG, with a multiple correlation coefficient  $R$  of 0.801, a coefficient of determination  $R^2$  of 0.622, and an  $F$  value of 77.338 at  $p<0.0001$ . Furthermore, the regression equation for prediction for the speed of the progression of POAG had a multiple correlation coefficient  $R$  of 0.857, a coefficient of determination  $R^2$  of 0.734, and an  $F$  value of 73.450 at  $p<0.0001$ .

## References

1. World Health Organization. Blindness and visual impairment (accessed Oct 2017). <http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
2. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. JAMA. 2014; 311:1901-11.
3. Nesterov AP. [Glaucoma]. Moscow: Meditsinskoye Informatsionnoye Aгенstvo; 2008. Russian
4. The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration. The AGIS Investigators. Am J Ophthalmol. 2000 Oct;130(4):429-40
5. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. Br J Ophthalmol. 2006; 90:262-7.
6. Ojha P, Wiggs JL, Pasquale LR. The genetics of intraocular pressure. Semin Ophthalmol. 2013 Sep-Nov;28(5-6):301-5.
7. Huang W, Wang W, Zhou M, et al. Association of glutathione S-transferase polymorphisms (GSTM1 and GSTT1) with primary open-angle glaucoma: an evidence-based meta-analysis. Gene. 2013 Sep 10;526(2):80-6
8. Surkova VK, Safina AZ, Orenburkina OI. [On the role of genetic predisposition to the development of primary open-angle glaucoma]. Glaukoma. 2007;4:21-2. Russian
9. Liu Y, Allingham RR. Molecular genetics in glaucoma. Exp Eye Res. 2011 Oct;93(4):331-9
10. Rykov SO, Burdei AV. [Association of GST deletion polymorphisms with primary open-angle glaucoma]. Archive of Ophthalmology of Ukraine. 2017;5(3):9:61-7. Ukrainian

**Table 1.** Correspondence between either predictors or dependent variables, and categorical, continuous, or dichotomous values that were used to build regression equation

Characteristics	Categorical values	Dichotomous values	Regression variable names	Used for calculations
Sex	M	101	S	dichotomous value
	W	102		
Age	-	-	A	continuous value
GSTP1	Ile/Ile	101	P1	dichotomous value
	Ile/Val	102		
	Val/Val	103		
GSTM1	+	101	M1	dichotomous value
	- (null)	102		
GSTT1	+	101	T1	dichotomous value
	- (null)	102		
SD <sub>POAG</sub>	-	-	SD <sub>POAG</sub>	continuous value
SP <sub>POAG</sub>	-	-	SP <sub>POAG</sub>	continuous value

**Table 2.** Regression equation coefficients and their statistical significance for dependent variable SD<sub>POAG</sub>

Characteristics	$\beta \pm SE$	t	p
Sex	0.0007 $\pm$ 0.0001	2.4130	1.650E-02
Age	-0.0001 $\pm$ 0.0003	-1.5590	0.0601
GSTP1	0.0038 $\pm$ 0.0009	4.4032	1.549E-05
GSTM1	0.0036 $\pm$ 0.0012	3.0759	2.318E-03
GSTT1	0.0047 $\pm$ 0.0013	3.6330	3.363E-04
Free term	-1.2813 $\pm$ 0.2514	-5.1249	5.743E-07

**Notes:**  $\beta \pm SE$ , beta coefficient with its standard deviation; t, Student test t value; p, significance of difference from the null hypothesis (the level of significance  $p \leq 0.05$  was assumed)

**Table 3.** Regression equation coefficients and their statistical significance for dependent variable  $SP_{POAG}$ 

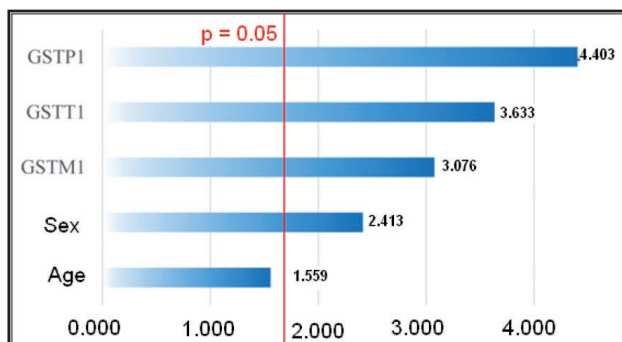
Characteristics	$\beta \pm SE$	t	p
Sex	-0.0629 $\pm$ 0.0199	-2.6721	8.050E-03
Age	0.0014 $\pm$ 0.0010	2.32134	2.110E-02
GSTP1	0.3212 $\pm$ 0.0632	5.09458	7.038E-07
GSTM1	0.2557 $\pm$ 0.0854	2.99565	3.023E-03
GSTT1	0.4361 $\pm$ 0.0937	4.65381	5.370E-06
Free term	-95.5044 $\pm$ 18.5515	-5.1686	4.938E-07

**Notes:**  $\beta \pm SE$ , beta coefficient with its standard deviation; t, Student test t value; p, significance of difference from the null hypothesis (the level of significance  $p \leq 0.05$  was assumed)

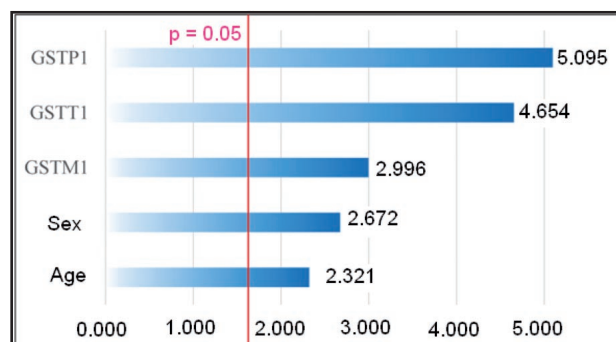
**Table 4.** Differences between actual and calculated values of  $SD_{POAG}$  and  $SP_{POAG}$ 

Characteristics	Me (Q1;Q3)	T	$p_{(T)}$	$\rho$	$p_{(\rho)}$
Actual value for $SD_{POAG}$	0.019 (0.016; 0.022)	347.0	2.16E-27	0.651	0.033
Predicted value for $SD_{POAG}$	0.014 (0.010; 0.014)				
Actual value for $SP_{POAG}$	1.000 (1.000;1.500)	990.0	1.85E-18	0.720	0.021
Predicted value for $SP_{POAG}$	0.866 (0.593; 0.642)				

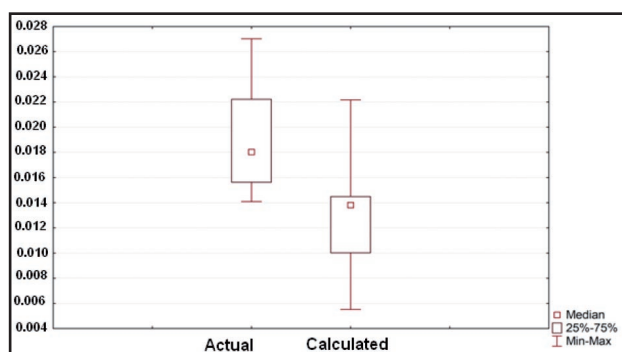
**Notes:** Me (Q1; Q3), median value, the 1st and 3rd quartiles; t, t value for Wilcoxon matched pairs test;  $\rho$ , Spearman rank correlation coefficient;  $p(t)$  and  $p(\rho)$ , statistical significance of differences for the relevant tests (the level of significance  $p \leq 0.05$  was assumed)



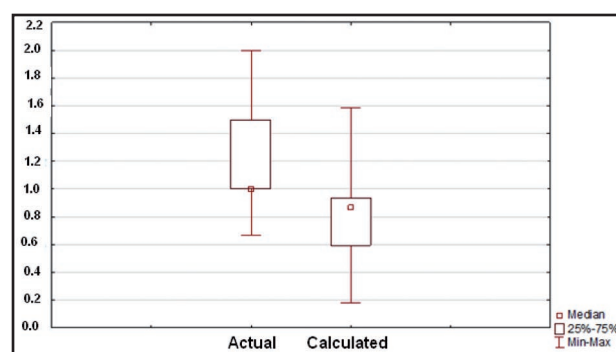
**Fig. 1.** Pareto chart for regression coefficients of the regression equation for dependent variable SDPOAG. The values on the x-axis are the Student test t values, obtained in the statistical interpretation of each coefficient of the regression analysis. The factors whose bars pass the line (p value equals to 0.05) indicate significance on the response value.



**Fig. 2.** Pareto chart for regression coefficients of the regression equation for dependent variable SPPOAG. The values on the x-axis are the Student test t values, obtained in the statistical interpretation of each coefficient of the regression analysis. The factors whose bars pass the line (p value equals to 0.05) indicate significance on the response value.



**Fig.3.** Actual and calculated values for SD<sub>POAG</sub>



**Fig.4.** Actual and calculated value for SP<sub>POAG</sub>