COMENIUS UNIVERSITY IN BRATISLAVA

FACULTY OF MEDICINE

ANALYSIS OF PREDICTIVE FACTORS OF PRE-ECLAMPSIA

Diploma Thesis

BRATISLAVA 2018 **RNDr. Iva Mišigová MSc**

COMENIUS UNIVERSITY IN BRATISLAVA
FACULTY OF MEDICINE

ANALYSIS OF PREDICTIVE FACTORS OF PRE-ECLAMPSIA

Diploma Thesis

Study program:	General Medicine
Study branch:	7.1.1. General Medicine
Department:	Department of Biology, Genetics and Clinical Genetics
Supervisor:	Ing. Helena Gbelcová PhD

BRATISLAVA 2018 **RNDr. Iva Mišigová MSc**

Comenius University in Bratislava

Faculty of Medicine

THESIS ASSIGNEMENT

Name and Surname:	Iva Mišigová
Study programme:	General Medicine
Field of Study:	General Medicine
Type of Thesis:	Diploma Thesis
Language of Thesis:	English
Title:	**Analysis of predictive factors of pre-eclampsia** ***Analýza prediktívnych faktorov pre-eklampsie***
Supervisor:	Ing. Helena Gbelcová, PhD
Institutes:	Institute of Medical Biology, Genetics and Clinical Genetics
Assigned:	2 November 2016
Approved:	2 November 2016

Doc. MUDr. Daniel Böhmer, PhD
Head of Department / Clinic

....

student supervisor

Univerzita Komenského v Bratislave
Lekárska fakulta

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Iva Mišigová
Študijný program: všeobecné lekárstvo (Jednoodborové štúdium, doktorské I.II. st., denná forma)
Študijný odbor: všeobecné lekárstvo
Typ záverečnej práce: diplomová práca
Jazyk záverečnej práce: anglický
Sekundárny jazyk: slovenský

Názov: Analysis of predictive factors of preeclampsia.
Analýza prediktívnych faktorov preeklampsie.

Vedúci: Ing. Helena Gbelcová, PhD.
Ústav: LF.ÚLBG - Ústav lekárskej biológie, genetiky a klinickej genetiky LF UK a UN Bratislava

Dátum zadania: 02.11.2016

Dátum schválenia: 02.11.2016

doc. MUDr. Daniel Böhmer, PhD.
vedúci ústavu/kliniky

.....
študent

.....
vedúci práce

DECLARATION

I hereby declare that this diploma thesis is my original work and it has been written by me in its entirety. I have acknowledged all the sources of information which have been used in this diploma thesis.

Date Signature

ABSTRACT

Pre-eclampsia is one of the most important causes of pregnant women's mortality and morbidity in Europe and globally. The study provides an overview of known risk factors and markers and concludes there is no reliable predictor for prediction and stratification of pre-eclampsia risk. It is proposed that mutations of the ERVW-1 gene are investigated as potential predictors to be used in clinical practice either standalone or in multivariable models. A methodology for such investigation including the proposal to develop a multivariable model based on the logistic regression is provided. The methodology is demonstrated and tested on simulation data. The simulation shows a limitation of small-scale studies with a sample smaller than 100 to identify risk factors with low prevalence. The study argues that it is still reasonable to conduct small scale studies because they should be able to identify risk factors with high risk that could improve the accuracy of predictive models.

Keywords: pre-eclampsia, ERVW-1 gene polymorphism, predictive model

ABSTRAKT

Pre-eklampsia je jednou z najvýznamnejších príčin mortality a morbidity tehotných žien v Európe a celosvetovo. Táto štúdia poskytuje prehľad známych rizikových faktorov a markerov pre-eklampsie a konštatuje, že neexistuje žiaden spoľahlivý prediktívny marker pre predpovedanie a stratifikáciu rizika pre-eklampsie. Štúdia navrhuje, aby boli preskúmané mutácie génu ERVW-1 ako potenciálne prediktívne markery, ktoré by mohli byť použité v klinickej praxi samostatne alebo v rámci modelov s viacerými premennými. Práca zároveň poskytuje metodológiu pre takýto výskum, vrátane modelu s viacerými premennými na základe logistickej regresie. Táto metodológia je testovaná na simulačných dátach. Simulácia poukazuje na obmedzenú schopnosť malých štúdií, s počtom vzoriek menším ako sto, identifikovať rizikové faktory s nízkou prevalenciou. Štúdia napriek tomu argumentuje, že má zmysel vykonávať malé štúdie, lebo tieto sú schopné identifikovať rizikové faktory s veľkým rizikom, ktoré by mohli zvýšiť presnosť prediktívnych modelov.

Kľúčové slová: pre-eklampsia, ERVW-1 génový polymorfizmus, prediktívny model

FOREWORD

The human genome research is a fast moving science area that opens new opportunities to address health challenges. It is important that this research is linked to clinical practice so that it provides tangible benefits. In creating such a link multifaceted and multidisciplinary collaboration and integration is necessary to discover the way how genetic research can make the work of clinical practitioners more effective in delivering better health care.

This study is an attempt to link the knowledge in several fields – genetics, microbiology gynecology and obstetrics, epidemiology and statistics – to find the way how to better manage pre-eclampsia, the disease that remains to be a mystery, and continues to claim lives and seriously affects the health of pregnant women and their children. The objective is modest -- it only seeks to improve prediction of pre-eclampsia – but the benefits can be substantial in improved outcomes of the prophylaxis and treatment. If we improve prediction we will also be closer to our understanding of the root cause of pre-eclampsia.

This integrated approach was possible thanks to the enabling role of the Institute of Medical Biology, Genetics and Clinical Genetics. This project evolved in the ecosystem created by the Institute that includes participating hospitals and laboratories but more importantly the people – a network of collaborating scientists and medical practitioners. The author would like to thank all collaborators and in particular to the supervisor Ing. Helena Gbelcová PhD and Petra Priščáková PhD for their invaluable help and advice. This study was supported by the project VEGA1/0168/18 “The study of pre-eclampsia etiology – the second most frequent cause of the pregnant women mortality”.

TABLE OF CONTENT

ABBREVIATIONS AND SYMBOLS

AFLP	acute fatty liver of pregnancy
ALT	alanine aminotransferase
APS	antiphospholipid syndrome
aPTT	activated partial thromboplastin time
ART	assisted reproductive technology
AST	aspartate aminotransferase
BMI	body mass index
CBC	complete blood count
CI	confidence interval
DIC	disseminated intravascular coagulation
ERVW-1	endogenous retrovirus group W member 1
FHR	fetal heart rate
FIGO	International Federation of Gynaecology and Obstetrics
FLT1	fms related tyrosine kinase 1
GPV	genetic polymorphism variable
HCTZ	hydrochlorthiazide
HELLP	hemolysis, elevated liver enzymes and low platelet
hs-CRP	high sensitivity C-reactive protein
HUS	haemolytic-uraemic syndrome
INR	international normalised ratio
ITP	immune thrombocytopenic purpura
IUGR	intrauterine growth restriction
LDH	lactate dehydrogenase
MAP	mean arterial pressure
OR	odds ratio
PAPP-A	pregnancy-associated plasma protein-A
PE	pre-eclampsia
PET	pre-eclamptic toxaemia
PI	pulsatility index
PIGF	placental growth factor
PRES	posterior reversible leukoencephalopathy syndrome
RBC	red blood cell
RIND	reversible neurological deficit < 48 h
ROC	receiver operator curve
RR	risk ratio
RUQ	right upper quadrant
sFLT1	soluable fms-like tyrosine kinase
SNP	single nucleotide polymorphism
TIA	transient ischemic attack
TTP	thrombotic thrombocytopenic purpura
VEGF	vascular endothelial growth factor
WHO	World Health Organisation

1 INTRODUCTION

Pre-eclampsia remains one of the top five causes of maternal and perinatal mortality worldwide. It is estimated that pre-eclampsia claims the lives of more than 70,000 women per year (Kassebaum et al., 2014) and more than 500,000 of fetuses and newborns; this is equivalent to the loss of 1600 lives per day (Firoz et al., 2011). For every woman who dies due to pre-eclampsia, it is estimated that other 20 suffer a life-altering morbidity (Pattinson and Hall, 2003).

More than 99 % of these losses occur in low- and middle-income countries, particularly those on the Indian subcontinent and sub-Saharan Africa (Khan et al., 2006). However, it is still relevant for Europe and Slovakia. Although the number of fatalities due to pre-eclampsia is significantly lower in Europe than in the developing countries it is still the most important cause of maternal death accounting for 17 - 24% of all maternal deaths (Widman and Bouvier-Colle, 2004). It is estimated that 9 billion EUR is spent every year on treatment of pre-eclampsia in Europe (Meads et al., 2008). It is why effective and efficient management of pre-eclampsia continues to be in the focus of medical research and practice in the field of obstetrics and gynaecology in Europe and worldwide.

Pre-eclampsia is a multisystem endothelial disorder. Its aetiology is not fully known and it can manifest with a variety of signs and symptoms that occur in different stages of pregnancy. The classification of pre-eclampsia in relation to other hypertension disorders varies in different medical systems. Also, the guidelines for diagnosis and management of pre-eclampsia differ in different countries and medical set ups (von Dadelszen et al., 2016). In most cases the treatment strategy includes antihypertensive therapy and other therapies addressing symptoms rather than the cause of the disease. In the case of severe pre-eclampsia the immediate delivery is the only therapeutic option (Pels et al., 2016). In patients with known risks of pre-eclampsia a preventive treatment can be applied to minimise the chance that the disease will manifest itself later (Han et al., 2016).

Preventive strategies and therapeutic interventions for pre-eclampsia are based on risk assessment. Risk stratification plays the key role in deciding about appropriate management of individual patients. Due to the systemic character of the disease, unknown aetiology, variability of signs and symptoms, and a high number of suspected risk factors it

is not easy to develop a reliable risk assessment method and an accurate predictive model. The challenge is further amplified by the low probability and high impact of the disease, including the death of the mother and/or the fetus. A highly sensitive and specific predictive method would make decisions of medical practitioners easier and lead to more effective and efficient care (Magee et al., 2016).

This work will contribute to the development of a quantitative predictive model for risk assessment and prediction of occurrence of pre-eclampsia in individual patients. The study reviews the existing knowledge of risk factors and predictive indicators and assesses whether any of the existing predictive method is reliable enough to be of practical use. The study further builds on the results of the newest research on genetic markers of pre-eclampsia at the Department of Medical Biology, Genetics and Clinical Genetics and hypothesises that ERVW-1 gene polymorphism could be a new predictive indicator, which, in combination with other predictive indicators, could improve the accuracy of multivariate predictive models.

Expecting that in the future more experimental data on ERVW-1 gene will be available this study focuses on the development of a methodology for a future experimental study that will explore the association between specific mutations and pre-eclampsia risk. The study also includes a simulation exercise to demonstrate how the method can be implemented and the result can be interpreted. It is hoped that this method will be picked up by other scholars who will then perform an experimental study and deliver a new predictive model that can be used in clinical practice.

2 PRE-ECLAMPSIA – AETIOLOGY, SIGNS AND SYMPTOMS, THERAPY AND PREVENTION

Pre-eclampsia is a serious pregnancy complication characterised by high blood pressure and damage to other organs, most frequently kidneys, liver and brain. The main symptom of pre-eclampsia is hypertension, often combined with proteinuria. Other symptoms may include impaired liver function, abdominal pain, nausea and vomiting, thrombocytopenia, headaches, changes or loss of vision, shortness of breath due to fluid in the lung and oedema. A more severe and life-threatening manifestation of the disease is eclampsia – clonic-tonic convulsions, as well as ischemic or haemorrhagic stroke, abruption of placenta, liver rupture, and acute renal failure (von Dadelszen et al., 2016).

Table 1 The adverse conditions that define pre-eclampsia and severe pre-eclampsia

Organ systems affected	***Adverse conditions that increase the risk of severe complications***	***Severe complications that warrant delivery***
CNS	Headache Visual symptoms	Eclampsia PRES Cortical blindness or retinal detachment Glasgow coma scale < 13 Stroke TIA RIND
Cardiorespiratory	Chest pain/dyspnoea Oxygen saturation < 97%	Uncontrolled severe hypertension (over a period of 12 hours despite use of three antihypertensive agents) Oxygen saturation < 90%, need for ≥ 50% oxygen for > 1 h, intubation, pulmonary oedema Positive inotopic support Myocardial ischaemia or infarction
Haematological	Elevated WBC count Elevated INR or aPTT Low platelet count	Platelet count < 50 x 10^9/L Transfusion of any blood product
Renal	Elevated serum creatinine Elevated serum uric acid	Acute kidney injury (creatinine > 150 μM with no prior renal disease) New indication for dialysis
Hepatic	Nausea or vomiting RUQ or epigastric pain Elevated serum AST, ALT, LDH, bilirubin Low plasma albumin	Hepatic dysfunction (INR > 2 in absence of DIC or warfarin) Hepatic haematoma or rupture
Feto-placental	Non-reassuring FHR IUGR Oligohydramnios Absent or reversed end-diastolic flow by Doppler velocimetry	Abruption with evidence of maternal or fetal compromise Reverse ductus venosus A wave Stillbirth

(Magee et al., 2014)

Because pre-eclampsia affects placental arteries it often results in foetal growth restriction, low birth weight and pre-term birth. In more severe forms of the disease the complications may lead to stillbirth. Severe forms of pre-eclampsia in most cases imply immediate delivery (Magee et al., 2014).

Pre-eclampsia does not present a risk for the mother and the child only in the gestational phase. Some studies demonstrated that the risk of cardiovascular, renal, hepatic and cerebral complications in mothers remained in the postpartum period (Filetti et al., 2012). The likelihood of recurrent pre-eclampsia in the next pregnancies is high and it increases with the severity of the current pre-eclampsia (Lin at al., 2015). People born from pre-eclampsia pregnancies tend to have life-long complications and higher incidence of chronic pulmonary diseases, cerebral paralyses, cognitive disorders, diabetes type 2, pulmonary hypertension and cardiovascular diseases (Tranquilli et al., 2012).

The diagnostic criteria for pre-eclampsia are summarised in table 2.

Table 2 Diagnostic criteria of pre-eclampsia and other conditions

Investigations for diagnosis	***Description in women with pre-eclampsia***	***Description in women with other conditions***
Maternal testing		
Urine testing		
Urinalysis (routine and microscopy with/without additional tests for proteinuria)	Proteinuria without RBCs or casts	Haemoglobinuria (dipstick 'haematuria' without RBCs): haemolytic anaemia RBCs alone: renal stones, renal cortical necrosis (also associated with back pain and oliguria/anuria) RBCs and/or casts are associated with other glomerular disease and scleroderma renal crisis and (about half of) TTP-HUS Bacteria: UTI or asymptomatic bacteriuria Proteinuria is usually absent in secondary causes of hypertension such as phenochromocytoma, hyperaldosteronism, thyrotoxicosis, coarctation of the aorta, and withdrawal syndromes
Oxygen saturation		
Pulse oximetry	SpO2 <97% associated with a heightened risk of severe complications (including non-respiratory)	May be decreased in any cardiorespiratory complication (e.g., pulmonary embolism)
CBC and blood film		
Haemoglobin	↑ due to intravascular volume depletion ↓ if microangiopathic haemolysis (with HELLP)	↑ due to volume depletion from any cause (e.g., vomiting) ↓ if microangiopathic haemolysis from other cause ↓ with any chronic anaemia (nutritional or

		myelodysplasia) ↓ with acute bleeding of any cause
WBC and differential	↔	↑ due to neutrophilia of normal pregnancy ↑ with inflammation/infection ↑ with corticosteroids
Platelet count	↓ associated with adverse maternal outcome	↓ with gestational, immune (ITP), or thrombotic thrombocytopoenia (TTP), APS, AFLP, myelodysplasia
Blood film	RBC fragmentation	Microangiopathy due to mechanical causes (e.g., cardiac valvopathy, cavernous haemangioma), DIC or other disorders of endothelial function (e.g., APS, TTP-HUS, vasculitis, malignant hypertension)
Tests of coagulation		
INR and aPTT	↑ with DIC which is usually associated with placental abruption ↑ is associated with adverse maternal outcome	May be ↑ in APS, DIC from other causes including sepsis, amniotic fluid embolism, stillbirth, massive haemorrhage, haemangiomas, shock ↑ is prominent in AFLP
Fibrinogen	↔	↓ with all causes of DIC including massive haemorrhage, genetic disorders ↓ more profound with AFLP than with HELLP Usually normal in TTP-HUS (ADAMTS13 vWF cleaving protein may be moderately decreased in HELLP106 but ADAMTS13 antibody should be absent
Serum chemistry		
Serum creatinine	↑ due to haemoconcentration and/ or renal failure ↑ associated with adverse maternal outcome	↑ with other acute or chronic kidney disease Renal failure prominent in malignant hypertension, TTP-HUS (along with thrombocytopoenia), AFLP (along with liver dysfunction)
Serum uric acid	↑ associated with adverse maternal and perinatal outcomes	↑ with dehydration, medication (e.g., HCTZ), genetic causes
Glucose	↔	↓ with AFLP, insulin therapy
AST or ALT	↑ associated with adverse maternal outcome	↑ with AFLP and other 'PET imitators' but to a lesser degree, and usually normal in TTP-HUS May be increased in other pregnancy-related conditions (e.g., intrahepatic cholestasis of pregnancy) or conditions not associated with pregnancy (e.g., viral hepatitis or cholecystitis)
LDH	↑ which may be prominent ↑ the is associated with adverse maternal outcome	↑ with AFLP, intravascular haemolysis ↑ LDH/AST ratio (>22) with TTP-HUS
Bilirubin	↑ unconjugated from haemolysis or conjugated from liver dysfunction	(early) ↑ in AFLP, ↑ with haemolytic anaemia, other liver disease with dysfunction, genetic diseases
Albumin	↓ associated with adverse maternal and perinatal outcomes	↓ as negative acute phase reactant with acute severe illness, malnutrition, nephrotic syndrome, crystalloid infusion
Fetal testing	*Abnormalities are not specific to the cause of poor placentation and/or placental dysfunction*	
Uterine artery Doppler velocimetry	Unilateral/bilateral notching, or elevated pulsatility index or resistance index may support a diagnosis of placental insufficiency including pre-eclampsia	

(Magee et al., 2014)

The therapeutic intervention depends on the severity of pre-eclampsia but often includes antihypertension therapy of severe and non-severe hypertension with oral and parenteral agents, fluid management, therapies for the HELLP syndrome including transfusion of blood products. Magnesium sulphate is administered for eclampsia prevention and treatment as well as fetal neuroprotection and corticosteroids for acceleration of fetal pulmonary development. In the case of severe pre-eclampsia that requires pre-term delivery, the key decision is the time of the childbirth. The decision should balance the maternal and fetal risks with the extension of the birth to a late pre-term gestation age with the view to reduction of neonatal morbidity (Magee et al., 2016).

Prevention strategy for pre-eclampsia high risk pregnancies includes low dose aspirin (75 - 100 mg/day) and calcium supplementation (1g/day) to be administered before the 16th week. Prophylactic dose of heparin can also be considered. Some national guidelines also recommend L-arginine, metformin, prostaglandin precursors, magnesium supplementation to prevent other pregnancy complications. Resting at home in the third trimester and reduction of workload or stress are the recommended lifestyle changes. In addition, periconceptional use of a folate-containing multivitamin, abstention from alcohol and smoking cessation are recommended for general beneficial effects in pregnancy (Han et al., 2016).

The aetiology of pre-eclampsia is not well understood. Several theories are proposed but none can fully explain the origin of this disease. It seems that there may be several origins and pathological pathways of pre-eclampsia. These may combine and result in variable signs and symptoms of the disease. The most supported theory assumes that there are two pathways – one is the placental disorder and the other is pre-existing maternal disorders such as renal disease, hypertension and diabetes. These processes can interact and result in a superposed placental pathology over the manifested or non-manifested maternal background disease leading to further organ damages and exacerbation of the symptoms of both (Staff et al., 2013).

The placental origin of pre-eclampsia is associated with an imbalance of angiogenic factors in the process of the development of placenta. An excess of anti-angiogenic factors like the soluble fms-like tyrosine kinase (sFLT1) and a reduction of pro-angiogenic factors like the vascular endothelial growth factor (VEGF) and the placental growth factor (PIGF) leading

to placental under-perfusion is observed in pre-eclampsia patients. The theory of placental origin of pre-eclampsia is confirmed by the recent research (InterPregGen, 2017) that associates the elevated sFLT1 with mutations of the fetal FLT1 gene thus confirming the placental FLT1 gene role in placental angiogenesis and the development of pre-eclampsia.

However, the angiogenic imbalance does not necessarily result in pre-eclampsia therefore the increased sFLT1 cannot play the role of the sole predictor. Maternal origin of pre-eclampsia plays another significant role the pre-eclampsia aetiology. A number of studies associate pre-eclampsia with the pre-existing maternal disease and propose a number of markers. For example, the SCOPE study (Kenny et al., 2014) links the late onset pre-eclampsia to factors that predict later cardiovascular disease through the metabolic syndrome. It is proposed that glycosylated fibronectin, the strong marker of gestational diabetes is a reliable marker for maternal pre-eclampsia. The InterPregGen project also identified the maternal gene variants which are associated with pre-eclampsia that have been previously linked to the essential (i.e. non-gestational) hypertension (InterPregGen, 2017).

The current albeit incomplete understanding of the aetiology suggests that pre-eclampsia is not a single disease but rather a set of overlapping and interacting conditions with different aetiology and pathogenic pathways. Further research in this subject will be needed to better understand this interaction and causality. It may result in reconsideration of the classification, prediction, diagnostics, therapy and prevention strategy.

Until the new research into the origin of pre-eclampsia discovers a causal predictive indicator the best approach to assessing and stratifying the risk and predicting the health outcome is to use stochastic methods, i.e. methods based on statistical prediction of the probability of occurrence of the disease. A statistical analysis of risk factors and various predictive indicators found to be associated with pre-eclampsia can provide a useful tool for practitioners to better manage the risk of pre-eclampsia and provide more effective and efficient care.

3 RISK FACTORS AND PREDICTIVE INDICATORS OF PRE-ECLAMPSIA

Extensive research has been done on risk factors and predictors of pre-eclampsia. This chapter summarises the findings and proposes suitable indicators as independent variable for a stochastic predictive model.

Risk factors are any factors that are found to increase the chance of a pregnant woman to develop pre-eclampsia. As pre-eclampsia is a complex disease of multiple origins it can be expected that there are numerous risk factors with different strength of association with the disease and consequently with different predictive values. They include a number of demographic, familial, life style, past medical history, gestational and paternal factors.

3.1 Demographic factors

Extremes of maternal age have been associated with the risk of pre-eclampsia and eclampsia (Redman et al., 2005). Maternal age of more than 40 years has been associated with an increased risk (Odds Ratio 1.49, 95 % Confidence Interval 1.22 – 1.82) (Khalil et al., 2013). The WHO Multi-country Survey of Maternal and Newborn Health reported that women older than 35 years were at high risk of pre-eclampsia, although not eclampsia. However, women younger than 19 years were at high risk for eclampsia, but not pre-eclampsia – this is probably related to under-diagnosis of pre-eclampsia in the population of women without full antenatal surveillance (Abalos et al., 2014).

Ethnicity of mother is also a risk factor. Women of South Asian or Afro-Caribbean ethnicity have higher risk when compared with Caucasians (Khalil et al.,2012, Wright et al.,2012). African-American women with severe pre-eclampsia demonstrate higher blood pressures and require more antihypertension treatment, while Caucasian women have a higher incidence of HELLP syndrome (Goodwin et al., 2005).

Socio-economic status seems to be a significant risk factor in developing countries. Rural inhabitants are twice more likely to suffer from pre-eclampsia compared to urban dwellers. (Endeshaw et al., 2015) Also the concurrent anemia, the diet poor in vitamins and nutrients

and low education level, which are often associated with a lower socioeconomic status, caused an increased risk of pre-eclampsia (Bilano et al., 2014)

3.2 Familial factors

A number of studies suggested that those with a family history of pre-eclampsia are at an increased risk for this disease (RR 2.90, 95 % CI 1.70 – 4.93) (English et al, 2015). A large population-based study reported a significantly higher risk of pre-eclampsia in sisters where the other sibling was diagnosed with pre-eclampsia (RR 2.6, 95 % CI 1.8 – 3.6) (Dawson et al., 2002). This risk increased further with the severity of disease (i.e., 2+ proteinuria) (RR 3.7, 95 % CI 2.5 – 5.5) (Dawson et al., 2002). Among family relatives of women who experienced pre-eclampsia, the rate of disease is higher in sisters (37 %), daughters (26 %) and grand-daughters (16 %) when compared with daughters-in-law (6 %) (Chesley and Cooper, 1986). Another study reported that the history of early- or intermediate-onset pre-eclampsia in mother or sister increased the risk of a similar form of pre-eclampsia by at least 150 % compared with an absence of such family histories (Boyd et al., 2013).

The maternal family history of pre-eclampsia is not the only familial factor. Esplin et al. (2001) suggested that women whose male partners were born from a pregnancy complicated by pre-eclampsia are more likely to suffer pre-eclampsia than partners of men who were born from normal pregnancies.

In addition, women with maternal or paternal familial history of hypertension or diabetes have an increased risk of developing pre-eclampsia (Qiu et al., 2003).

3.3 Life-style factors

Although smoking is considered to have adverse effect on all organ systems a number of studies show that smoking during pregnancy halves the risk of pre-eclampsia (England and Zhang, 2007). This protective effect is seen irrespective of parity and severity of disease.

Studies are not conclusive on the effect of physical activity. While some studies find that physical activity has a protective effect on the development of pre-eclampsia (OR 0.99, 95 % CI 0.93 - 1.05) others do not find any significant effect (Kasawara et al., 2012).

3.4 Mental health

Lifetime stress and stress during pregnancy is known to double the risk of pre-eclampsia (Yu et al., 2013). In addition, anxiety and depression in the first trimester of pregnancy was found to increase the risk of pre-eclampsia 2- to 3-fold (Kurki et al., 2000).

3.5 Past medical history

One of the most significant risks for pre-eclampsia is when the women already suffered from pre-eclampsia in the past. Women with a history of pre-eclampsia are at increased risk (15 – 65 %) of developing pre-eclampsia in subsequent pregnancies. The risk depends on the time of onset and the severity of pre-eclampsia in the previous pregnancy. Women who had HELLP syndrome have 25 % risk, those who had preterm delivery have 55 % risk and those who had early onset pre-eclampsia have 65 % risk of pre-eclampsia recurrence (Odegard et al., 2000).

Women with low birth weight (< 2500g) have been shown to have double the risk of experiencing pre-eclampsia (OR 2.3, 95 % CI 1.0 – 5.3) when compared with women with normal weight at birth. Further, the risk increased 4-fold for those women who weighed < 2500g at birth and were overweight as adults (Dempsey et al., 2003). A Danish cohort study reported that there was an increased frequency of pre-eclampsia in women who were born prematurely and were small for their gestational age (á Rogvi et al., 2012).

Women who are overweight or obese are known to be at increased risk for pre-eclampsia (Shamsi et al., 2013). A recent review of studies concluded that overweight or obesity as well as maternal adiposity is associated with an increased risk of pre-eclampsia (Wang et al., 2013). Increased BMI is an important risk factor for all levels of severity of pre-eclampsia with the risk of 64 % (Pare et al., 2014). This risk increases 2- to 3-fold as the BMI increases from 21kg/m^2 to 30 kg/m^2 (Bodnar et al., 2005). In addition to overweight the short stature of women (< 164 cm) was found to increase the risk of severe pre-eclampsia (Sohlber et al., 2011).

The pre-existing medical conditions may play an important role in the development of pre-eclampsia. Diabetes (type 1 and type 2) increases the risk of pre-eclampsia 2 to 4 times (Sibai et al., 2000). Pre-gestational diabetes increases also the risk of late-postpartum pre-eclampsia (Bigelow et al., 2013).

Chronic hypertension is another significant risk factor of pre-eclampsia. 23 % of women with chronic hypertension (MAP > 95 mmHg) are at risk of superimposed pre-eclampsia (Lecarpentier et al., 2013). A systematic review found that the risk of superimposed pre-eclampsia in women with chronic hypertension was nearly eight-fold higher than in the general pregnancy population (Bramham et al., 2014).

Other medical conditions found to be associated with an elevated risk of pre-eclampsia are chronic kidney disease, lupus nephropathy, as well as diabetic nephropathy (Hirose et al., 2014). For women with diabetes and proteinuria of 190 – 499 mg/day the risk of pre-eclampsia is significantly higher (Bramham et al., 2011).

Both inherited and acquired thrombophilia have been assessed for their association with pre-eclampsia. Studies found that factor V Leiden single nucleotide polymorphism (SNP) is associated with an increased risk of pre-eclampsia. No association was found between methylene tetrahydrofolate reductase SNP and prothrombin SNP and the risk of pre-eclampsia (Lin and August, 2005). Another review concluded that the risk of pre-eclampsia was two times higher in women who tested positive for antiphospholipid syndrome (Abou-Nassar et al., 2011).

3.6 Pregnancy related factors

Pre-eclampsia is recognised to more frequently affect the first pregnancy. An older large population-based study concluded that nulliparous women are at increased risk of pre-eclampsia compared with parous women (OR 3.6, 95 % CI 2.6 – 5.0) (Odegard et al., 2000). A recent large cohort study reported that nulliparity significantly increased the risk of late-onset pre-eclampsia when compared with early-onset disease (Lisonkova and Joseph, 2013).

Despite the fact that the risk of pre-eclampsia is generally less frequent in the second and consecutive pregnancies if conceived with the same partner, it seems that the risk rises with the increase of interval between pregnancies. Each year of delay between pregnancies increases the risk of pre-eclampsia with the odds ratio 1.12 (95 % CI 1.11 – 1.13) (Skjaerven et al., 2002). Consistently with this finding another study found that if the period between pregnancies is four years the risk of pre-eclampsia is increased with odd ratio 1.4 (95 % CI 1.2 – 1.6) (Mostello et al., 2008).

Multiple pregnancy also increases the risk of pre-eclampsia. For twin pregnancy the risk ratio is 2.62 (95 % CI 2.03 – 3.38) (Sibai et al., 2000). It is assumed that a large placenta mass results in increased circulating levels of sFLT-1 – the anti-angiogenic factor described above (Bdolah et al., 2008).

Assisted reproductive technology (ART) (especially in vitro fertilization) was found to be associated with higher risk of gestational hypertension and pre-eclampsia (Thomopoulos et al., 2013). A cohort study reported that hypertensive disorders occurred more frequently in ART pregnancies than in spontaneously conceived pregnancies (Opdahl et al., 2015).

Infections are also suspected to increase the risk of pre-eclampsia. A UK study reported that antibiotic prescriptions (as a proxy for acute infection) and urinary tract infection increased the risk of pre-eclampsia with odds ratio 1.28 (95 % CI 1.14 – 1.44) and 1.22 (95 % CI 1.03 – 1.45) respectively (Minassian et al., 2013). However, other studies concluded that only certain infections were associated with the risk of pre-eclampsia, e.g. urinary tract infection and periodontal infections but other maternal infections such as chlamydia, malaria, treated or untreated HIV were not (Conde-Agudelo et al., 2008).

Congenital malformation is associated with an increased risk of pre-eclampsia with risk ratio 1.26 (95 % CI 1.16 – 1.37) (Conde-Agudelo and Belizan, 2000). Fetal malformation is more strongly associated with early-onset pre-eclampsia rather than late-onset (Lisonkova and Joseph, 2013).

3.7 Paternal factors

Paternal age seems to play a role in pre-eclampsia, most likely as a result of genetic mutations that occur with aging or due to environmental factors such as exposure to radiation and heat. The risk of pre-eclampsia doubles if the father is older than 45 years (Dekker et al., 2011).

It is also suggested that the length of sexual relationship between partners or the length of sperm exposure can affect the risk of pre-eclampsia as repeated intercourse can lead to better accommodation of maternal organism to paternal antigens. However, studies only confirmed this association for gestational hypertension and not for pre-eclampsia (Olayemi et al., 2010).

Medical history of paternal of cardiovascular diseases is also suspected to be a risk factor for pre-eclampsia but the evidence is conflicting. One case control study reports that early chronic hypertension and myocardial infarction in the father increased the risk of pre-eclampsia 3-fold (Rigo et al., 2006). On the other hand a population-based study reports that there was no association between the hypertensive disorders of pregnancy and paternal cardiovascular risk factors such as BMI, blood pressure and lipid profile (Myklestad et al., 2011).

The risk factors above are usually used in the clinical practice to screen women with a higher risk of pre-eclampsia. Alone, these are not reliable predictors because only about 30 % of cases of pre-eclampsia are identified by their use (Park et al., 2013). This is why researchers looked for other clinical and laboratory markers that could be used for more reliable and precise prediction of pre-eclampsia.

3.8 Laboratory markers

Although the blood pressure could be considered as the key predictor of pre-eclampsia because it is a basis of its diagnosis, it seems that no index of blood pressure predicts pre-eclampsia well enough to be clinically useful (Ukah et al, 2016). A meta-analysis of 34 studies that applied blood pressure indicators in the second trimester for predicting pre-

eclampsia concluded that these indicators were only weakly associated with the actual development of pre-eclampsia (Cnossen et al., 2008).

Proteinuria is another diagnostic parameter of pre-eclampsia. The onset of proteinuria in early pregnancy, as a sign of an underlying renal disease, is associated with the risk of pre-eclampsia. The most precise predictor seems to be the albumin to creatinine ratio measured in early pregnancy (Baweja et al., 2011). This test is not used beyond experimental studies, most likely because of the difficult access to the sophisticated analytical technology (this indicator requires a high-performance liquid chromatograph) in a standard clinical practice (Ukah et al, 2016).

Glomerular epithelial cells – podocytes – are suggested as a possible predictor for pre-eclampsia. Studies analysing the predictive ability of podocytes that use podocyte proteins (podocin and nephrine) as markers reported no (Jim et al., 2014) to medium (Kelder et al., 2012) to high predictive performance (Craici et al., 2013).

As renal dysfunction results in increased creatinine and decreased calcium so the calcium to creatinine rate was tested for its predictive potential but it was found to be a poor predictor for pre-eclampsia (Vahdat et al., 2012).

The insulin mediator inositol phosphoglycan-P type is found to be high in urine in pre-eclampsia. It was tested for its predictive function and is now considered as a moderately good short-term predictor for pre-eclampsia (up to 2 weeks before diagnosis) (Dawonauth et al., 2014).

Placental proteins are considered to be potential markers for pre-eclampsia as they can indicate the endothelial dysfunction and damage. The high sensitivity C-reactive protein (hs-CRP) and fibronectin were studied with mixed results. Hs-CRP was found as having a poor (Kashanian et al., 2013) and fibronectin only moderate (Leeflang et al., 2007) predictive performance.

Angiogenic factors were suggested as pre-eclampsia predictors because they are suspected to be connected to the causal mechanism of pre-eclampsia. The placental growth factor (PIGF) that is a pro-angiogenic factor produced by the syncytiotrophoblast has lower

maternal circulation concentration in pre-eclampsia (McElrath et al., 2012). Its predictive performance was evaluated by several studies with contradictory results. While Gosh et al. reported that PIGF was a poor predictor of pre-eclampsia (Ghosh et al., 2013a; Ghosh et al., 2013b) Chappell et al. concluded that PIGF was strongly associated with the negative likelihood of pre-eclampsia with delivery within 14 days of diagnosis (Chappell et al., 2013). Part of this controversy can be attributed to the different methods of detection and quantification of PIGF in different studies (Ukah et al, 2016).

The soluble fms-like tyrosine kinase 1 (sFlt-1) in an anti-angiogenic factor produced by the placenta. It antagonises the pro-angiogenic factors like PIGF by binding to them. As sFlt-1 reduces the level of PIGF in the maternal circulation it was suggested that the sFlt-1 to PIGF ratio could be used as a pre-eclampsia predictor. A number of studies found the sFlt-1:PIGF ratio to be a good predictor of pre-eclampsia (Engels et al., 2013; Teixeira et al., 2013; Hanita et al., 2014; Forest et al., 2014) contradicted by a study that did not find any association of this indicator with pre-eclampsia (McElrath et al., 2012).

3.9 Ultrasound markers

As pre-eclampsia is associated with the abnormal development of arteries in uterus during pregnancy it is expected that Doppler ultrasonography could be the method to detect these abnormalities and predict pre-eclampsia reliably. Doppler ultrasound is a non-invasive, routinely used method that can study the utero-placental circulation and changes in blood flow resistance. The parameters of flow change that can be measured are the pulsatility index and the resistance index. The abnormal uterine Doppler velocimetry can be considered: bilateral notching or no notching with the resistance index > 0.7, mean resistance index > 0.55 or unilateral notching > 0.65 (Papageorghiou, 2002).

A few studies assessed abnormal uterine Doppler velocimetry parameters as predictors for pre-eclampsia with mixed and sometimes contradictory results. While one study found the uterine artery pulsatility index measured in the first and the second trimester to be a poor predictor of pre-eclampsia (Bolin, 2012) another study concluded that the pulsatility index is a good predictor and in particular the mean uterine artery difference (i.e. the difference between multiples of medians) between the first and the second trimester was a good predictor of early onset pre-eclampsia (Napolitano, 2012). Meta analyses of other studies

concluded that Doppler velocimetry is a moderately accurate predictive method particularly in the second trimester (Papageorghiou, 2002, Cnossen, 2008). However, these analyses reviewed heterogenous studies with different Doppler techniques and indices therefore the results cannot be considered conclusive.

4 MULTIVARIABLE PREDICTIVE METHODS OF PRE-ECLAMPSIA

The above review of risk factors and predictive markers shows that currently no single clinical or laboratory indicator has sufficient sensitivity and specificity to be an acceptable predictor of pre-eclampsia that could be used in clinical setting. This explains the number of attempts to develop multivariate predictive models that test different combinations of predictive indicators, both clinical and laboratory, and assess their accuracy for prediction of pre-eclampsia.

One of the recent studies is the SCOPE project that assessed 47 biomarkers that are either associated with pre-eclampsia, play biological role in placentation or play a role in cellular mechanism involved in the pathogenesis of pre-eclampsia. The result was a multivariable model using the angiogenin to PIGF ratio, MAP, miscarriage < 10 weeks, and the mean uterine artery resistance index. This model had the area under the receiver operator curve 0.73 for the model development cohort and 0.68 for the validation cohort, which was evaluated as a 'modest' predictive ability (Kenny, 2014).

Another large study developed a multivariate model based on the combination of MAP, uterine artery pulsatility index, pregnancy-associated plasma protein-A (PAPP-A) and PIGF. The model was able to identify 93 % of early onset pre-eclampsia, 36 % of late onset pre-eclampsia and 18 % of gestational hypertension (Poon, 2009).

It can be concluded that models based solely on laboratory markers like the two above do not provide sufficient accuracy. Most other models combine biomarkers with maternal characteristics or other clinical factors, despite the fact that these indicators alone are not considered useful predictors of pre-eclampsia (Ukah, 2016). The table 3 summarises multivariable models from different studies.

Table 3 A summary of studies and parameters entering multivariate models

Model	*Biomarkers*	*Blood pressure*	*Doppler ultrasound*	*Maternal characteristics*
Gallo et al., 2014		MAP		Gestation age Weight Height Racial origin Family history of PE Personal history of PE Smoking Chronic hypertension
Kleinrouweler et al., 2013		BPs, BPd, MAP	Lower PI, Higher PI Lower RI, Higher RI, Mean RI Bilateral notching, any notching	Age Height Weight BMI Obesity Smoking Alcohol consumption Ethnicity
Caradeux et al., 2013		BPs, BPd, MAP	PI	Age BMI Parity History of PE History of preterm labour History of hypertension Diabetes mellitus
Austdal et al., 2015	Urinary hippurate to creatinine ratio	MAP		Age
Elia et al., 2017	Urinary albumin to creatinine ratio	MAP		Gestational age Maternal age Essential hypertension Preexisting diabetes Gestational diabetes Smoking Nulliparity Social deprivation index BMI
Lai et al., 2013	PAPP-A PIGF ß-hCG			Weight Height Race Parity Chronic hypertension
Schneuer et al., 2012	PAPP-A ß-hCG PP-13			History of hypertension Parity Weight Maternal age
Di Lorenzo et al., 2012	PIGF ß-hCG			Chronic hypertension
Keikkala et al., 2013	PAPP-A hCG-h	MAP		Parity
Kuc et al., 2013 (99) and Kuc et al., 2014	Taurine PAPP-A PIGF ADAM12	MAP		Parity Weight
Goetzinger et al., 2014	PAPP-A ADAM12		PI	Chronic hypertension History of PE

				Pre-gestational diabetes Obesity
Park et al., 2013	PAPP-A	MAP	PI	History of PE
Myers et al., 2013	PIGF	MAP		Family history of PE History of infertility treatment
Rizos et al., 2013	Difference in PIGF between the 1st and 2nd trimester		PI	BMI
Perales et al., 2016	sFLT-1 to PIGF ratio	MAP		Parity History of PE

The models that combine blood pressure indicators, biomarkers, Doppler ultrasound indicators and maternal characteristics show improved performance compared to models based on biomarkers only, however, they still have only moderate predictive ability. Only a few models have the area under ROC curve larger than 0.7. Very few models were validated by sufficiently large validation studies. In those few models that were validated the accuracy of prediction was not fully confirmed.

In particular the sensitivity of these multivariate models is rather limited and most of them is capable to predict only some 50 – 60 % of pre-eclampsia cases. They seem to be better in predicting the early onset pre-eclampsia than the late onset pre-eclampsia. Adding more indicators, e.g. taurine, placental protein 13, free beta subunit of hCG, (i.e. pregnancy associated plasma proteins that are involved in implantation, trophoblast invasion and remodelling of spiral arteries) does not increase the accuracy of models significantly compared to models based on angiogenic markers (e.g. sFLT:PIGF ratio) or indicators of blood pressure, Doppler indicators and maternal characteristics.

Consequently, the multivariate models are rarely used in clinical practice. Few national guidelines give recommendation concerning the early screening of pre-eclampsia. The most recent recommendation of the International Federation of Gynaecology and Obstetrics (FIGO) considers as the best practice to screen pregnant women for clinical markers of pre-eclampsia in early pregnancy including history of pre-eclampsia, multiple pregnancy, antiphospholipid antibody syndrome, significant proteinuria and pre-existing conditions of hypertension, diabetes mellitus or renal disease (Ukah et al., 2016). Screening of other markers is not considered necessary because it has not shown any improvement in medical outcomes yet.

However, the FIGO calls for further research on predictive methods including multivariate models. There is a need for the following:

- Large prospective studies on the risk of pre-eclampsia;
- Standardisation of definitions and analytical methods to enable comparison of results and meta-analyses;
- Predictive models should aim to predict medical outcomes in terms of the severity and the time of disease onset for population groups with increased risk. Prediction of early onset pre-eclampsia (in the first trimester) is critical for sound decisions about prophylactic interventions.
- Predictive models need to be tested on larger samples to ensure their validity for a broader population.

5 SEARCHING FOR GENETIC PREDICTIVE INDICATORS

In the absence of reliable clinical and biological markers of pre-eclampsia researchers are looking for a new approach. One of the promising avenues is to look at the genetic make-up of pregnant women and their fetuses and identify genetic indicators that might be associated with genetic polymorphism that is hypothesised to have a causal role in pre-eclampsia.

One of such promising indicators is syncytin-1. Syncytin-1 is a protein encoded by the ERVW-1 gene. This gene belongs to the human endogenous retrovirus genes that form about 8 % of the human genome. These retroviral genes entered human genome in the early phase of its development millions years ago and most of them are transcriptionally inactive as a result of mutations and immune responses of the host organism. It is however assumed that the ERVW-1 gene that entered the human genome as the last one was silenced selectively and preserves some of its function that are beneficial for the host organism. The beneficial function of syncytin-1 is that it stimulates cell-cell fusion and the best-known role of this function is in the placenta development. Other beneficial roles may include the modulation of maternal immune tolerance to the fetus (Knerr et al., 2004).

Syncytin-1 mediates trophoblast fusion, which is a precondition for normal placental development. The placenta is composed by two cell layers – the syncytiotrophoblast and cytotrophoblast. The expression of syncytin-1 on the surface of both facilitates their fusion and the growth and maintenance of the syncytiotrophoblast layer. In addition, it is assumed that syncytin-1 also regulates proliferation of cytotrophoblasts and ensures continual replenishment of the cytotrophoblast pool (Huang et al., 2013).

In line with this understanding of syncytin-1's role, the expression of syncytin-1 is observed to rise in the first trimester, i.e. in the time of rapid expansion of placental size and function, in normal pregnancies (Holder et al., 2012). In pre-eclamptic pregnancies significant reduction of syncytin-1 expression is observed (Roland et al., 2016). This can lead to the conclusion that syncytin-1, measured as mRNA or protein in trophoblasts, could be used as a new biomarker for pre-eclampsia.

However, the predictive potential of syncytin-1 is difficult to assess due to the gaps in our understanding of details of regulatory mechanisms in the differentiation and fusion of trophoblast cells. For example, syncytin-1 cannot be the only placental growth regulatory agent and the exclusive mediator of the cell fusion process in human placentogenesis because it is an acquired mechanism. There must be another mechanism that exists in other mammals that do not have ERVW-1 gene (Knerr et al., 2004). Another uncertainty about the predictive potential of syncytin-1 stems from the observation that the increased syncytin-1 expression plays the role of a compensatory mechanism in IUGR pregnancies (Gao et al., 2012). Such a secondary increase can mask the primary underexpression of syncytin-1 as a root cause of pre-eclampsia.

A more practical drawback of the use of syncytin-1 as a predictor of pre-eclampsia is that so far there is no non-invasive method that would allow to measure the syncytin-1 level in the early stage of pregnancy. Syncytin-1 analysis requires a biopsy of the placental tissue and it can be too risky and inefficient to perform such an analysis as part of early pregnancy pre-eclampsia screening. Therefore other related indicators could be considered. One possibility is to analyse polymorphism of the ERVW-1 gene itself as a root cause of the decrease of expression of syncytin-1. No study has yet proved the link between ERVW-1 polymorphism and pre-eclampsia but these studies could not be considered conclusive due to their limiting design such as the small size of the sample (Priscakova et al., 2016). If the link between the ERVW-1 polymorphism and pre-eclampsia is proven a genetic test of pregnant women in early pregnancy could be considered as a feasible screening method. Then the disruption of ERVW-1 regulatory region that results in anomalies of syncytin-1 expression could be used as a predictor of pre-eclampsia either alone or in combination with other predictors in a multivariate predictive model.

Before ERVW-1 gene mutation is established as a predictor of pre-eclampsia (either as a single predictor or a co-predictor in a multivariate predictive model) the following steps need to be taken:

1) Proving the association between anomalies in syncytin-1 expression and specific mutations of ERVW-1 gene by performing an experimental study of appropriate design;
2) Testing a statistical correlation between the specific ERVW-1 gene mutations and pre-eclampsia by conducting a retrospective case control study;

3) Development of a multivariate predictive model based on clinical and laboratory indicators including an indicator of ERVW-1 polymorphism and testing the predictive value of the model and the added predictive value of the ERVW-1 mutation;
4) Validation of the predictive model by conducting a prospective study on a larger population.

The step 1 is the subject of the research conducted at the Institute of Medical Biology, Genetics and Clinical Genetics of the Faculty of Medicine of Comenius University Bratislava. The ongoing research collects data on:

- genetic variability of the ERVW-1 gene;
- expression of syncytin-1 (at the level of mRNA and proteins);
- immunolocalisation of syncytin in histological samples

in a sample of pregnant women who developed pre-eclampsia and a control group (Priscakova, 2017, Priscakova, 2018). This research will assess the strength of association between ERVW-1 mutations and the reduction of syncytin-1 expression in placenta.

This ongoing research focus on those ERVW-1 mutations that are considered to be associated with the rate of expression of syncytin-1. In the first phase of the research four DNA variants in the 2139 bp long region at positions -1340, -1046, -246, -30 were tested but no association of these variants with expression of syncytin-1 or frequency of pre-eclampsia was found (Priscakova et al., 2016). In the next phase of this research polymorphisms 128C>A, 134G>A, 157_158insA, 422C>T, 533A>T, 534G>T, 992T>C, 1222G>C, 1438C>T, 1878C>A and 1987C>T are being investigated (Priscakova, 2018). If this research shows an association between these mutations and the rate of expression of syncytin-1, these mutations will be considered as possible predictive indicators.

Before this research produces sufficient amount of data the methodology for next steps can be developed. The next sections of this study propose such a methodology and test it on simulation data. The basic assumption for this further work is that there is at least one mutation for which the association with pre-eclampsia will be found. For the development of a predictive model we create a generic dichotomous independent variable “GPV” representing the investigated ERVW-1 gene polymorphism that will be used in the design of further studies. In case more associations are found, additional variables (e.g. GPV1, GPV2, …, GPVn) can enter the further studies.

6 TESTING THE ASSOCIATION BETWEEN ERVW-1 GENE MUTATION AND PRE-ECLAMPSIA

Testing of the potential association between ERVW-1 gene mutation and pre-eclampsia can be done in the form a retrospective case control study.

For this study a sample of patients diagnosed with pre-eclampsia is compared to a control group of pregnant women without the disease. The frequency of gene mutation in the two groups will be compared by using an appropriate statistical test. The statistical hypothesis is that the two samples come from the same population, in other words, that there is no statistically significant difference between the frequency of gene mutation in the case group and the control group. If this null hypothesis is rejected, we can conclude the association between the ERVW-1 gene mutation and pre-eclampsia exists.

To ensure statistical validity, the sample and the control group must represent their respective populations, which is normally achieved by random selection of a sample of sufficient size.

The minimum size of the sample depends on the expected effect size of the investigated risk factor and the required level of statistical confidence and power of the study. The required confidence and power were set conventionally to 95 % and 80 % respectively. The 95 % confidence level means that if we find an association there is 95% certainty that this association really exists. The 80 % power means that if we do not find an association there is 80 % certainty that the association does not exist.

There are number of methods for calculation of the sample size but for this particular case the method based on comparison of proportions is most suitable. The challenge for this study is that even an approximative value of the effect size of ERVW-1 genetic mutation and the corresponding proportions in the pre-eclamptic and healthy populations are not known because studies of this kind have not been done before. The minimum sample size can be calculated on the basis of hypothetical scenarios exploring the whole range of possible values of the effect size.

For this study ten hypothetical scenarios were created where the effect size of the ERVW-1 mutation was expressed in terms of the relative risk and prevalence of the mutation. We assumed four levels of possible relative risk and three levels of prevalence. For each scenario the expected frequency of mutation in the case and the control group was calculated and from these values the minimal sample size to achieve the required level of confidence and power was derived. These scenarios are summarised in table 4. The mathematical basis for these calculations is described in Annex 1.

Table 4 Scenarios for calculations of sample size

Expected prevalence of ERVW-1 mutation in the population* High: P = 0.2 Moderate: P = 0.05 Low: P = 0.01	Expected risk ratio associated with the ERVW-1 mutation** Very high: RR = 8 High: RR = 4 Low: RR = 2 Very low: RR = 1.2	Calculated frequency of ERVW-1 mutation in the case group (proportion π_1)	Calculated frequency of ERVW-1 mutation in the control group (proportion π_2)	Calculated minimum sample size for confidence 95% and power 80%
0.20	1.2	0.231	0.199	> 1000
0.20	2.0	0.333	0.197	161
0.20	4.0	0.500	0.194	34
0.05	1.2	0.059	0.050	> 1000
0.05	2.0	0.095	0.049	489
0.05	4.0	0.174	0.047	93
0.01	1.2	0.012	0.010	> 1000
0.01	2.0	0.020	0.010	> 1000
0.01	4.0	0.039	0.009	422
0.01	8.0	0.075	0.009	140

*The prevalence of different ERVW-1 gene mutations typically ranges between 0.001 and 0.2 (1000 Genome Browser, 2018)
**RR = 4 is conventionally considered as high, RR = 2 as low (Garb, 1996)

It is clear from the table 4 that the scenarios with the expected low and very low relative risk or with low frequency of genetic mutation require unrealistically large samples that exceed the number of all cases of pre-eclampsia in Slovakia per year. We consider that the maximal achievable size of the case group for the study corresponds to the number of pre-eclamptic women in the region covered by the participating hospitals during one year. As there were 107 cases of pre-eclampsia in 2015 in the Bratislava region (Národné centrum zdravotníckych informácií, 2017) the size of the case group was set to n = 100.

This sample size should be sufficient if the observed mutation has high to moderate frequency and high risk. The sufficiency of sample size n = 100 is described in figure 1. This sample size is sufficient for those combination of expected frequency of genetic mutation and expected relative risk of this mutation that are positioned above the curve. If

the sample size is smaller than 100 the curve moves up and to the right (e.g. see the dotted line for n = 50).

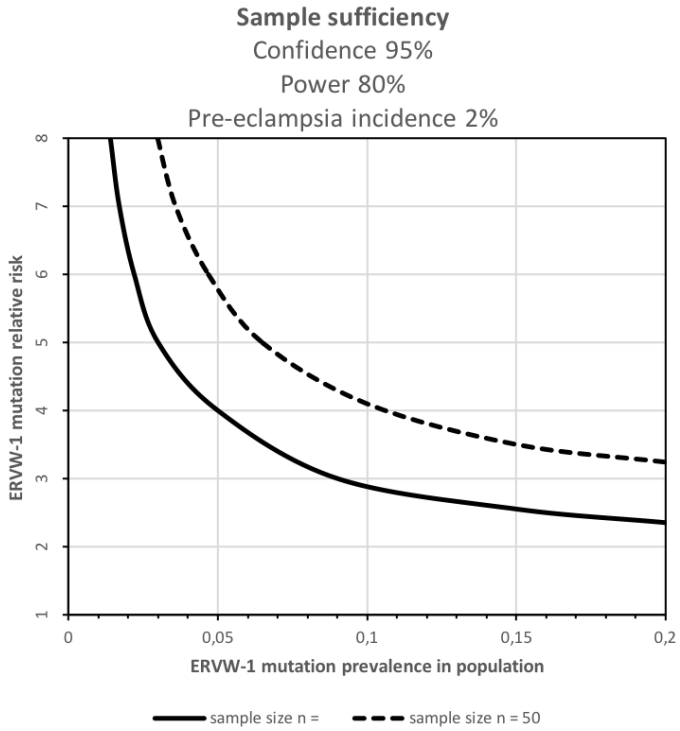


Figure 1 Sample sufficiency for sample size n = 100

The control group should be ideally as large as the case group in order to achieve the minimal pooled variance of the case and control group (Woodward, 1999). The control sample with n = 100 should be created by random sampling from the population of non-preeclamptic pregnant women giving birth in the participating hospitals.

The recruitment of participants to both the case and the control group should follow the research ethic principles (e.g. the informed consent of participants) and the design of the whole study should be approved by the ethical committee.

The correct selection of participants is important to avoid errors and bias. One important source of error can be the diagnosis of pre-eclampsia. As pre-eclampsia manifests in several different signs and symptoms there may be a number of false positive or false negative diagnoses. The study protocol should provide clear diagnostic criteria for the participating doctors to observe, e.g. the criteria in table 1. The correctness of diagnosis

and assignment to the case group could be verified according to diagnostic data that can be collected in the study.

Another source of error can be caused by biased selection of the control group, e.g. if the recruitment of controls is done in the hospital that deals with difficult pregnancies from the whole region so there is a higher concentration of patients with pathologies that may have some connection to pre-eclampsia. The representativeness of the control group can be checked by comparison of the control group characteristics to the general population of pregnant women. For example, some demographic characteristics of the control group could be compared to the general population, e.g. the average age in the control group could be compared to the average age of pregnant women in Slovakia by using an appropriate statistical test. If this check is negative the control group may be biased.

After the data are collected the association between ERVW-1 mutation and pre-eclampsia will be tested by a suitable statistical test considering the sample size and the frequency of occurrence of gene mutation. If we have a sample of appropriate size, we can use the chi-squared test that is widely used for testing association between two categorical variables. The result of chi-squared test gives the p value, i.e. the approximate probability that there is no difference between the compared frequencies. If the sample is small the Fisher exact test may provide better results (Harris and Taylor, 2008).

The statistical null hypothesis is that the frequency of gene mutation in the case group is the same as in the control group. The statistical test will confirm or reject this hypothesis at certain confidence level. If the $p \leqq 0.05$ the null hypothesis is rejected. It means that there is 95 % certainty that the frequency of mutation in pre-eclamptic sample differs from the frequency of this mutation in the control sample so we can conclude that there is an association between the gene mutation and pre-eclampsia. The GPV variable can be considered as a potential predictor and it can be used for building a predictive model.

If the results are of less than acceptable significance level, i.e. $p > 0.05$, and the null hypothesis cannot be rejected it could be analysed whether the reason is that the expected association does not exist or that the strength of this association, and therefore also the effect size, is so low that the study is not able to capture it. Several statistical programmes (including SPSS that is used in this study) offer a retrospective analysis of observed data to

determine the post hoc power but many statisticians consider it as irrelevant (Hoenig and Heisay, 2001). Instead they propose a complex method that includes analysis of confidence intervals, setting alternative hypothesis and equivalence testing. The complexity of such analysis goes beyond the scope of this study. It is also not necessary to achieve the objective of this study to find a good genetic predictor of pre-eclampsia. It can be argued that if our sample of n = 100 results in $p > 0.05$ the power is too low to reliable identify associations between pre-eclampsia and an ERVW-1 mutation with relatively low prevalence and low to very low relative risk. Such mutation would most likely not be a good pre-eclampsia predictor. Therefore, for this study the further analysis of insignificant results is not critical.

7 THE DEVELOPMENT OF MULTIVARIATE PREDICTIVE MODEL

The multivariate predictive model assumes that there are several independent factors that individually increase the risk of pre-eclampsia (risk factors) or decrease the risk (protective factors). They interplay in a complex way but if considered together they can provide a sufficiently sensitive and specific predictive tool for clinical practice that can assist physicians to assess the risk of individual patients and decide about the best prophylaxis or therapy.

The conceptual model of the predictive multivariate model is described in figure 2.

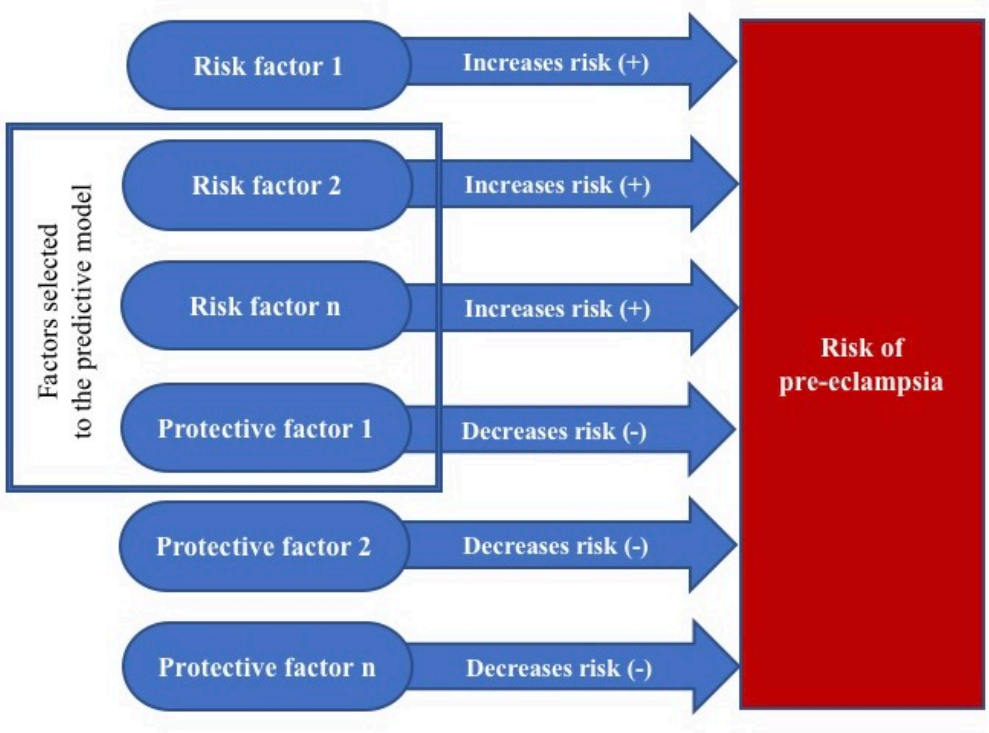


Figure 2 The conceptual model of a multivariate predictive model for pre-eclampsia

The first task in the design of the model is selection of risk and protective factors as predictors, i.e. the independent variables of the model. Numerous studies have identified a broad range of risk factors (and a few protective factors) as demonstrated in chapter 3 and 4. In theory all these factors could enter the multivariate model if we can consider them as independent from each other. It is however important that the model uses variables that

make sense for clinical practice, i.e. the variable can be measured in the first trimester by applying routinely used methods. For example, it is inutile if the predictive variable requires very sophisticated and costly analytical instruments that may be used in highly specialised hospitals but are not available for majority of gynaecological practices. Methods that are not standardised or are not sensitive or specific enough are not suitable for a clinical predictive tool.

The previous research can also inform the design of the model by indicating which risk factors have only weak association with pre-eclampsia and are likely to have marginal contribution to the model and can be omitted from the start. The additional complication in the pre-eclampsia model is that the causal relationship between the independent variables and the dependent variable is not known and some of the seemingly independent variables may not be independent. This is the case when the risk factors identified in previous research represent just different manifestations of the same causal process. In the multivariate predictive model these quasi-independent variables may mutually decrease their predictive values.

The usual approach to the development of a multivariate model is that the initial design starts with a larger number of independent variables and it is optimized by eliminating useless variables through iterative calculations. As the main objective of this study is to investigate the predictive role of ERVW-1 gene mutation we will simplify the initial multivariate model (the reference model) and include independent variables that are widely recognized as useful for risk assessment. For this purpose, we will limit the reference model to those variables recommended by FIGO. These variables are routinely measured and recorded during medical checks in early pregnancy. The list of reference model's independent variables therefore includes: history of pre-eclampsia, multiple pregnancy, proteinuria, pre-gestational hypertension and diabetes mellitus type 1 (the antiphospholipid syndrome recommended as a predictive variable by FIGO is excluded from the predictive model because antiphospholipid antibodies are not regularly screened in pregnant women in Slovakia). These variables will be complemented by the ERVW-1 gene polymorphism variable GPV that was proven to be associated with pre-eclampsia in the previous research stage. The list of independent variables entering the model is in table 5.

The continuous variables like blood pressure and proteinuria may be converted into binomial variable by using a threshold value. It should be tested if such conversion affects negatively the accuracy of the model.

Table 5 Independent variables entering the model

Independent variable	*Type*	*Unit*	*Threshold for conversion into binomial variable*[*]
Hypertension (MAP)	Continuous	mmHg	> 95
Urinary protein	Continuous	mg/24h	> 300
Multiple pregnancy	Binomial	yes/no	
Diabetes mellitus type 1	Binomial	yes/no	
History of pre-eclampsia	Binomial	yes/no	
ERVW-1 gene polymorphism	Binomial	yes/no	

[*]The threshold values are based on the FIGO guidelines. If the model with converted variables loses the predictive ability significantly compared to the model with continuous variables, the threshold values can be modulated to find the optimal way of converting continuous to binomial values.

The statistical method selected for this model is the logistic regression. It calculates the probability of expected outcome using the logistic function. Because the logistic function is considered as the best approximation of the probability of health outcomes the logistic regression is widely used in medicine for multivariate analysis where the dependent variable is the health outcome and independent variables are observed characteristics of patients. The mathematical background of the logistic regression is explained in Annex 2.

This study will be designed as a retrospective case control study where two representative samples of patients with pre-eclampsia (the case group) and without pre-eclampsia (the control group) will enter the study. Selection and randomisation rules apply in the same way as in the study in chapter 6 above.

The size of the case group can be decided by applying the "one in ten" rule. This rule stipulates that for each independent variable there should be at least 10 cases (Wittinghoff and McCulloch, 2007). The size of control group should be sufficiently big to reduce the variance of regression coefficient estimates. In practical terms this means that the control group should be at least 5 times larger than the case group (Woodward, 1999).

After the case and control samples are recruited to the study a DNA analysis is performed to find out the presence of investigated ERVW-1 gene polymorphism. Other independent variables will be based on medical records of patients entering the case and control groups.

Collected data will be subject of statistical analysis by a statistical software SPSS. The following sequence of analyses will be performed.

(1) Assessment of representativeness of the sample based on demographic and prevalence data (not performed on simulation data because these data are created to take into account the knowledge on prevalence).
(2) Analysis of correlation between independent variables. If a correlation is found between any two independent variables one of these variables can be excluded from the study. A variable that produces a more precise model should be maintained (not performed on simulation data as these data were generated randomly in order to be independent).
(3) Logistic regression analysis is performed to obtain estimates of the regression coefficients b_0, b_1, …, b_n. One or more tests will be carried out to assess the goodness of fit of the model. The Wald test will be performed to determine the significance level of each coefficient estimate. Depending on the value of estimates obtained calculations can be repeated for different combinations of independent variables, not excluding single variable models. As the particular interest of this study the contribution of the ERVW-1 polymorphism variable GPV will be assessed and discussed.
(4) The predictive model will be tested on the observed data. To determine the accuracy of the model the count pseudo-R^2 test, where $R^2 = \frac{\text{number of correct predictions}}{\text{sample size}}$, will be performed. To understand the trade-off between the sensitivity and specificity of the model the area under (AUC) receiver operation curve (ROC) will be calculated. Different variants of the model based on different combinations of independent variables can be compared.

If the model is found to have acceptable accuracy it can be further developed into a practical tool. The b_0 intercept should be corrected by using the formula (7) for the real frequency ratio so it can be used on the real population. As the model produces probabilities in the range from 0 to 1 the cut off values from which the probability will be considered as positively predicted case can be set for the optimal combination sensitivity and specificity for clinical practice. Alternatively, more than one cut off values can be defined to create several levels of risk, e.g. high risk for prophylactic treatment, moderate risk for frequent monitoring and low risk for standard management. A user-friendly calculator based on the model's formula can be developed for easy use by practitioners.

Before the model is used in medical practice it should be validated. The model parameters (estimates of the intercept b_0 and coefficients b_i, …, b_n) were calculated to best fit the observed data but it may not fit equally well other samples from the same population. As mentioned in chapter 4 some pre-eclampsia predictive models lost their predictive accuracy when they were used on data from a new validation cohort. Therefore, any new predictive model needs to be validated to understand if it is sufficiently accurate for the whole population.

One possible approach to validation is cross-validation. In this method the sample is divided into data subsets. The model is calculated again for one or more subsets and validated on other subsets. The accuracy of models based on subsets is compared to the accuracy of the model based on the complete dataset. Any relevant quantitative measure of fit can be used for cross-validation, for example the count pseudo-R^2.

However, in our study the cross-validation may not be applicable as the sample size is close to the minimum size and using data subsets may not produce statistically significant results. Therefore, it is proposed that the new predictive model of pre-eclampsia is validated by a new dataset from a prospective study on a cohort of pregnant women.

8 SIMULATION EXERCISE

In absence of observed data a simulation exercise can be carried out to demonstrate how the methodology proposed in this paper should be applied and how the results could be interpreted.

8.1 Analysis of association between genetic polymorphism variables and pre-eclampsia

8.1.1 Data generation

The simulation data were generated on the basis of the assumptions below. Some of these assumptions are purely arbitrary and they were selected only to demonstrate certain aspects of the methodology, others are approximation of the real situation, e.g. the prevalence of non-genetic risk factors in the population.

The simulation data on ERVW-1 gene polymorphism were generated for ten scenarios based on combination of four different levels of risk associated with the mutation and three levels of frequency of this mutation in the population. It is assumed that the most prevalent ERVW-1 mutations reach or exceed the level of 20 % of the population, while many mutations have very low prevalence of 1 % and less (1000 Genome Browser, 2018). This is why the three prevalence levels – high, moderate and low – were set to P = 0.2, 0.05 and 0.01 respectively. The four risk levels were set to RR = 1.2, 2, 4 and in one case 8. Conventionally, RR = 2 is considered as low risk and RR = 4 as high risks (Garb, 1996). We assume that RR values 1.2 and 8 can represent very low and very high risks.

Frequencies of presence and absence of mutation in the case group and the control group were calculated for each of ten scenarios as well as the odds ratio and the minimum sample size. Then data were produced as values of the independent variables GPV1, GPV2, …, GPV10. The data were generated by random assignment of value 1 (the risk factor present) or 0 (risk factor absent) so that the dataset met the required numbers of mutation presence or absence in a sample of the size n = 100, as indicated in table 6. Random generation of data ensures that the GPV variables are mutually independent. Data were produced only for six ‘realistic’ scenarios, i.e. where the required confidence level and power can be achieved with a sample size smaller than 1000. It should be noted that due to rounding to

integer of the numbers of exposed and unexposed cases and controls the odds ratios of generated data differ from the odds ratios calculated on the basis of frequencies. This means that the generated simulation data have slightly different risk characteristics from the original risk assumptions for data generation. The most affected is the GPV10 variable.

Table 6 Overview of scenarios for generation of data for the GPV variable

Scenario	*Variable*	*Frequency in the case group*	*Frequency in the control group*	*Minimum sample size for 95% confidence and 80% power*	*Data generated: Number in the case group of n = 100*	*Data generated: Number in the control group of n = 100*
High risk & high prevalence (RR = 4, P = 0.2)	GPV1 = 1	0.500	0.194	34	50	19
	GPV1 = 0	0.500	0.806		50	81
High risk & moderate prevalence (RR = 4, P = 0.05)	GPV2 = 1	0.174	0.047	93	17	5
	GPV2 = 0	0.826	0.953		83	95
High risk & low prevalence (RR = 4, P = 0.01)	GPV3 = 1	0.039	0.009	422	4	1
	GPV3 = 0	0.961	0.991		96	99
Low risk & high prevalence (RR = 2, P = 0.2)	GPV4 = 1	0.333	0.197	161	33	20
	GPV4 = 0	0.667	0.803		67	80
Low risk & moderate prevalence (RR = 2, P = 0.05)	GPV5 = 1	0.095	0.049	489	10	5
	GPV5 = 0	0.905	0.951		90	95
High risk & low prevalence (RR = 2, P = 0.01)	GPV6 = 1	0.020	0.010	2282	Data not generated	
	GPV6 = 0	0.980	0.990			
Very low risk & high prevalence (RR = 1.2, P = 0.2)	GPV7 = 1	0.231	0.199	2681	Data not generated	
	GPV7 = 0	0.769	0.8			
Very low risk & moderate prevalence (RR = 1.2, P = 0.05)	GPV8 = 1	0.059	0.050	8783	Data not generated	
	GPV8 = 0	0.941	0.950			

Very low risk & low prevalence (RR = 1.2, P = 0.01)	GPV9 = 1	0.012	0.010	41831	Data not generated	
	GPV9 = 0	0.988	0.990			
Very high risk & low prevalence (RR = 8, P = 0.01)	GPV10 = 1	0.075	0.009	140	7	1
	GPV10 = 0	0.925	0.991		93	99

8.1.2 Data analysis

In the next step the generated simulation data are considered as observed experimental data and processed as real observed data would be processed if they were available.

The data were analysed by applying the Pearson's chi-square test and Fischer's exact test. These tests assess if the difference in the frequency of ERVW-1 mutation in the case group and the control group are statistically significant. The critical parameter is the p-value. If p is lower than 0.05 we can conclude with at least 95% certainty that the frequencies are not the same, i.e. there is an association between the genetic mutation and pre-eclampsia. The results of the analysis are summarised in the table 7.

Table 7 Chi-square and Fischer test for GPV1, GPV2, GPV3, GPV4 and GPV5 variables

Variable	*Scenario*	*p (chi-square)*	*p (Fisher test)*
GPV1	High risk & high prevalence	<0.001	<0.001
GPV2	High risk & moderate prevalence	0.007	0.011
GPV3	High risk & low prevalence	0.174	0.369
GPV4	Low risk & high prevalence	0.037	0.054
GPV5	Low risk & moderate prevalence	0.179	0.283
GPV10	Very high risk & low prevalence	0.030	0.065

The table shows that the difference in frequency of occurrence of genetic mutation in the case and control groups is not statistically significant for GPV3 and GPV5 variables as the p-value is greater than the required level of significance 0.05. These variables represent scenarios with high risk and low prevalence, and low risk and moderate prevalence. For these two scenarios the sample of n = 100 is too small and our study is not able to detect association despite the fact the data were generated with pre-defined level of risk. This is an example of the statistical type II error. The solution would be to increase the sample

size, which may not be realistic for a study of pre-eclampsia. As a consequence of the undetected association the variables GPV3 and GPV5 were excluded from further studies.

Variables GPV4 and GPV10 are border cases where the chi-square test shows statistical significance but the Fisher test does not. We can further assess the significance by studying the confidence intervals for observed risk expressed as odds ratio (table 8).

Table 8 Risk of pre-eclampsia associated with the ERVW-1 polymorphism

Variables of ERVW-1 polymorphism	Odds ratio	95% confidence interval
GPV1	4.263	2.259 – 8.045
GPV2	3.892	1.376 – 11.007
GPV4	1.970	1.035 – 3.749
GPV10	7.452	0.900 – 61.729

Predictably, the analysis showed the levels of risk corresponding to the assumptions that were used in generating simulation data. It can noted that the odds ratio values are similar to the relative risk values albeit these two risk concepts have very different statistical meaning. This is the effect of the low frequency of pre-eclampsia. The important fact is that the confidence interval is rather wide for all GPV variables. We can interpret this as we have some level of certainty that the variables are risk factors (OR > 1) but we are quite uncertain about the level of risk. This is not valid for GPV10 variable where the confidence interval includes the value of 1, which means we do not have 95 % certainty that the variable is a risk factor, neutral or protective factor. This variance of risk parameter could be in theory reduced by a bigger sample size.

We can conclude that by analysing the case and control groups of size n = 100 we found a statistically significant association between variables GPV1, GPV2 and GPV4, and pre-eclampsia. These variables can be tested further for their predictive ability in a multivariable model. The variable GPV10 does not seem to be a significant predictor despite the fact that the data were generated with a very high risk ratio assumption. The sample of 100 does not have sufficient power to identify this variable as a risk factor. We will include this parameter in the multivariable model to see if this method can identify GPV10 as a significant predictor.

8.2 Developing and testing logistic regression models for prediction of pre-eclampsia

Variables GPV1, GPV2, GPV4 and GPV10 will be tested for their predictive ability in a logistic regression model.

8.2.1 Data generation

To build a model an additional dataset for variables representing risk factors other than ERVW-1 polymorphism is needed. Based on the FIGO recommendations these variables will be included in the model:

Pre-gestational hypertension: Simulation data for variable PGHT will be generated on the assumption that the prevalence of hypertension in the population of pregnant women equals to the prevalence of general population corresponding to approximately 35 % (WHO, 2018). Previous studies found the hypertension to be an important risk factor of pre-eclampsia with the relative risk RR = 8 (see the overview of pre-eclampsia risk factors in section 3).

Pre-gestational proteinuria: The variable PGPU will be built on known data about the prevalence of proteinuria. It is assumed that the population of pregnant women has the same frequency of proteinuria as is the prevalence in the general population, i.e. 8.5 % (Bezinque et al., 2017). Earlier studies report that the relative risk of pre-eclampsia associated with pre-gestational proteinuria is RR = 4.

Diabetes mellitus type 1: The simulation data for DMT1 variable take into consideration that in Slovakia the prevalence of diabetes type 1 in women in the age group 25 - 29 is 512 per 100,000 (Národné centrum zdravotníckych informácií, 2018). The relative risk was reported by previous studies to be RR = 3.

Multiple pregnancy: About 2.5 % of all pregnancies in Slovakia are multiple pregnancies (Národné centrum zdravotníckych informácií, 2017). The relative risk of pre-eclampsia associated with multiple pregnancies was reported to be RR = 2.6. This information will assist to build the dataset for variable MPRE.

History of pre-eclampsia: The data for variable PEHI will be based on a simple estimation of how many pregnant women can have a history of pre-eclampsia. There is about 35,000 pregnant women in Slovakia annually who have already given birth (Národné centrum zdravotníckych informácií, 2017). 2 % of them, i.e. 700, might have experienced pre-eclampsia. These women represent about 1.3 % of 55,000 pregnant women annually. The history of pre-eclampsia is reported to be a strong risk factor with RR = 8.

The frequencies and risk characteristics for these variables are summarised in the table 9.

Table 9 Overview of non-genetic risk factors for generation of simulation data

		Frequency in the case group	*Frequency in the control group*	*Odds ratio based on frequencies*
Pre-gestational hypertension (RR = 8, P = 0.35)	Risk factor present PGHT = 1	0.812	0.341	8.3
	Risk factor absent PGHT = 0	0.188	0.659	
Pre-gestational proteinuria (RR = 4, P = 0.085)	Risk factor present PGPU = 1	0.271	0.081	4.2
	Risk factor absent PGPU = 0	0.729	0.919	
Diabetes mellitus type 1 (RR = 3, P = 0.005)	Risk factor present DMT1 = 1	0.015	0.005	3.1
	Risk factor absent DMT1 = 0	0.985	0.995	
Multiple pregnancy (RR = 2.6, P = 0.025)	Risk factor present MPRE = 1	0.063	0.024	2.7
	Risk factor absent MPRE = 0	0.937	0.976	
History of pre-eclampsia (RR = 8, P = 0.013)	Risk factor present PEHI= 1	0.095	0.011	9.2
	Risk factor absent PEHI = 0	0.905	0.989	

The multivariate model will include the four genetic mutation GPV variables and the five non-genetic variables. By rule of thumb referred to above the multivariate model with 9 independent variables should have at least 90 cases and 5 times more in the control group. For the sake of consistency with the previous exercise the size of the case group was set to 100. The contingency tables for generated data for samples n = 100 for cases and n = 500 for controls for all independent variables are summarised in the table 10.

Table 10 Contingency tables for all independent variables

Variable		*Number in the case group n = 100*	*Number in the control group n = 500*	*Odds Ratio*
GPV1	GPV1 = 1	50	97	4.2
	GPV1 = 0	50	403	
GPV2	GPV2 = 1	17	24	4.1
	GPV2 = 0	83	476	
GPV4	GPV3 = 1	33	99	2.0
	GPV3 = 0	67	401	
GPV10	GPV4 = 1	7	4	9.3
	GPV4 = 0	93	496	
PGHT	PGHT = 1	81	170	8.3
	PGHT = 0	19	330	
PGPU	PGPU = 1	27	41	4.1
	PGPU = 0	73	459	
DMT1	DMT1 = 1	1	2	2.5
	DMT1 = 0	99	498	
MPRE	MPRE = 1	6	12	2.6
	MPRE = 0	94	488	
PEHI	PEHI = 1	10	6	9.2
	PEHI = 0	90	494	

Due to the rounding to integer the simulation data have slightly different risk characteristics than the original assumptions. The most affected variable is the DMT1 variable.

After the data were generated by random assignment (to ensure independence of variables) they could enter into the logistic regression model.

8.2.2 Data analysis and model development

All the calculated models were tested for their statistical significance and goodness of fit by several tests provided by the SPSS program including the likelihood ratio test, Cox and Snell R^2, Nagelkerke R^2, and Hosmer and Lemeshow test. The tests confirmed statistical significance for all models below. Models differed in their goodness of fit but all models had a better fir than the null model, i.e. the model where parameters b_1, …, b_n are equal to zero.

8.2.2.1 Single variable predictive models

First, the logistic regression was calculated for each variable separately. The results are summarized in the table 11.

Table 11 Logistic regression single variable models for individual variables

Variable	*Parameter b estimate*	*Significance (Wald)*	*Exp(b)*	*95% confidence interval of exp(b)*	*Correctly predicted cases in %*	*Correctly predicted controls in %*	*Cut off value*
GPV1	1.424	< 0.001	4.155	2.648 – 6.518	50	81	0.3
GPV2	1.402	< 0.001	4.062	2.092 – 7.888	17	95	0.3
GPV4	0.691	0.004	1.995	1.245 – 3.196	33	80	0.2
GPV10	2.234	< 0.001	9.333	2.679 – 32.520	7	99	0.5
PGHT	2.113	< 0.001	8.276	4.858 – 14.099	81	66	0.3
PGPU	1.421	< 0.001	4.141	2.401 – 7.141	27	92	0.3
DMT1	0.922	0.453	2.515	0.226 – 28.007	-	-	-
MPRE	0.954	0.063	2.596	0.951 – 7.088	-	-	-
PEHI	2.214	< 0.001	9.148	3.244 – 25.798	10	99	0.5

The main output from the calculation is the estimate of parameter b. This is the coefficient that expresses the change in log-odds if the value of the independent variable changes. The statistical significance of this estimate is assessed by the Wald test. The interpretation of these results can be demonstrated on the GPV1 variable. The b estimate for GVP1 means that the log-odds increases by 1.424 if the patient has the GVP1 mutation. This corresponds to increase of odds of pre-eclampsia by e^b, in this case by 4.155. This is expected as the data were generated to produce odds ratio OR = 4.

The single variate models produce significant estimates of b for all variables except for diabetes type 1 DMT1 and multiple pregnancy MPRE. The confidence interval include value 1 therefore there is uncertainty whether the variable is a risk factor at all. It should be noted that the confidence interval is wide also for the statistically significant estimates, especially for variables with low frequency. As mentioned above the low frequency results in high variance of the estimate. This variance could be reduced by more observations if feasible.

It is interesting to observe that the variable GPV10 that did not show a significant association with pre-eclampsia in the analysis before is significant predictor in the logistic regression and has the highest value of parameter estimate b. We can assume that the increased size of the control group from 100 to 500 reduced the variance and the parameter became significant. The confidence interval is still very wide.

The estimate b serves to construct the formula for calculation of probabilities of an event, i.e. of the occurrence of pre-eclampsia. For all variables with the significant b estimate the

predicted probabilities were calculated and compared to observed values. This provides information on the accuracy of the model and can be described in terms of sensitivity, i.e. the ability to detect patients with the disease, and specificity, i.e. the ability to identify healthy subjects. The predictive ability of the single variate logistics models is described by the percentage of correctly predicted values. The table 11 shows that most single variate models have poor ability to identify high risk patients as they correctly predicted only 8-50 % cases. The exception is pre-gestational hypertension variable PGHT that is able to predict 81 % of cases.

It can be concluded that the models we created on the basis of single variables using the simulation data are poor predictive tools, except the model based on pre-gestational hypertension.

8.2.2.2 The reference multivariable model

In the next step several variants of multi variable model were developed and tested. The reference multivariable model to which other models will be compared is the model that includes only non-genetic variables. The results for this reference model are described in table 12.

Table 12 Parameter estimates for a reference multivariable logistic regression model without genetic polymorphism variables

Variable	*b*	*P (Wald)*	*Exp(b)*	*95% Confidence Interval*	
PGHT	2.056	0.000	7.814	4.504	13.555
PGPU	1.421	0.000	4.141	2.228	7.695
DMT1	1.045	0.537	2.845	0.103	78.447
MPRE	1.243	0.033	3.466	1.104	10.884
PEHI	2.075	0.001	7.962	2.452	25.854
Intercept	-3.193	0.000			

The meaning of the variable b is the same as in the single variable models described above. The logistic regression analysis calculated b for estimates for all variables but it turned out that in this model the estimates for diabetes type 1 variable DMT1 is not statistically significant. As apparent from table 12 the dominant risk determinant in this model is again the pre-gestational hypertension but also the history of pre-eclampsia PEHI variable.

To assess the accuracy of the model the Receiver Operator Curve (ROC) was constructed for this model in figure 3. The ROC produced the area under the curve (AUC) of 0.792, which is conventionally considered as fair (models with AUC > 0.8 are considered as good and < 0.7 as poor).

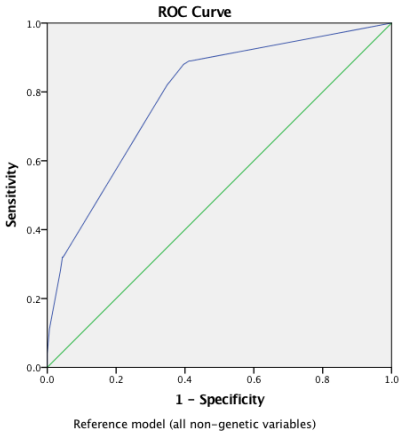


Figure 3 ROC for the multi variable logistic regression model without genetic polymorphism variables

8.2.2.3 The complete multivariable model with multiple genetic polymorphism variables

Next, the logistic regression was calculated for the model with all genetic polymorphism variables GPV and all non-genetic variables. It was then tested how the precision of the model changed compared to the reference model. The estimates of parameter b are summarised in the table 13 for this model.

Table 13 Parameter estimates for a logistic regression model with all predictor variables

Variable	*b*	*P (Wald)*	*Exp(b)*	*Confidence Interval*	
GPV1	1.358	0.000	3.889	2.267	6.671
GPV2	1.196	0.009	3.307	1.355	8.073
GPV4	0.629	0.033	1.875	1.053	3.341
GPV10	1.792	0.024	6.004	1.263	28.542
PGHT	1.983	0.000	7.268	4.083	12.936
PGPU	1.529	0.000	4.612	2.367	8.985
DMT1	0.048	0.984	1.049	0.011	104.042
MPRE	1.420	0.019	4.137	1.265	13.534
PEHI	1.694	0.010	5.444	1.500	19.760
Intercept	-3.929	0.000			

All independent variables produced statistically significant b estimates with $p < 0.05$ except the diabetes mellitus type 1 DMT1 variable. This variable could be excluded from the model without any negative effect. This model shows better goodness of fit compared to the reference model, as measured by Cox and Snell, and Negelkerke pseudo R^2 tests.

To assess the accuracy of this model the ROC curve was calculated, see figure 4. The area under the curve is 0.846 which implies that the model can be considered as good and is superior to the reference model.

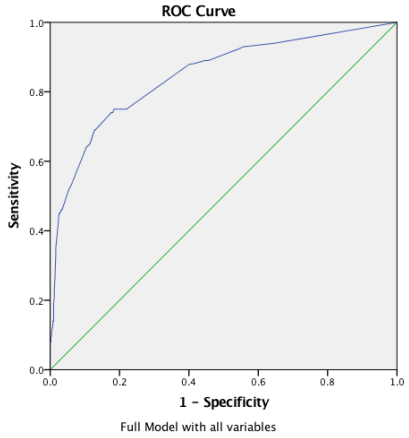


Figure 4 ROC for the multi variable logistic regression model based on all studied predictor variables

8.2.2.4 Reduced models with one genetic polymorphism variable

The genetic variables added more accuracy to the model, which could probably be expected for any variable that has a positive value of b parameter. However, it is not likely that there will be a number of ERVW-1 mutation that will be able to play a predictive role. More likely, it will be just one mutation. It can be analysed whether a single genetic variable could improve the reference model. The figure 5 shows the ROC curves for the four combinations of individual GPVs and all non-genetic variables. The areas under the ROC curve are AUC = 0.836, 0.805, 0.798 and 0.802 for models with variables GPV1, GPV2, GPV4 and GPV10 respectively. We can see that all GPV variable increase the area under the curve compared to the reference model, but the GPV1 variable has the biggest effect despite the fact that this is not the variable with the highest b estimate in single variable nor in multivariable models.

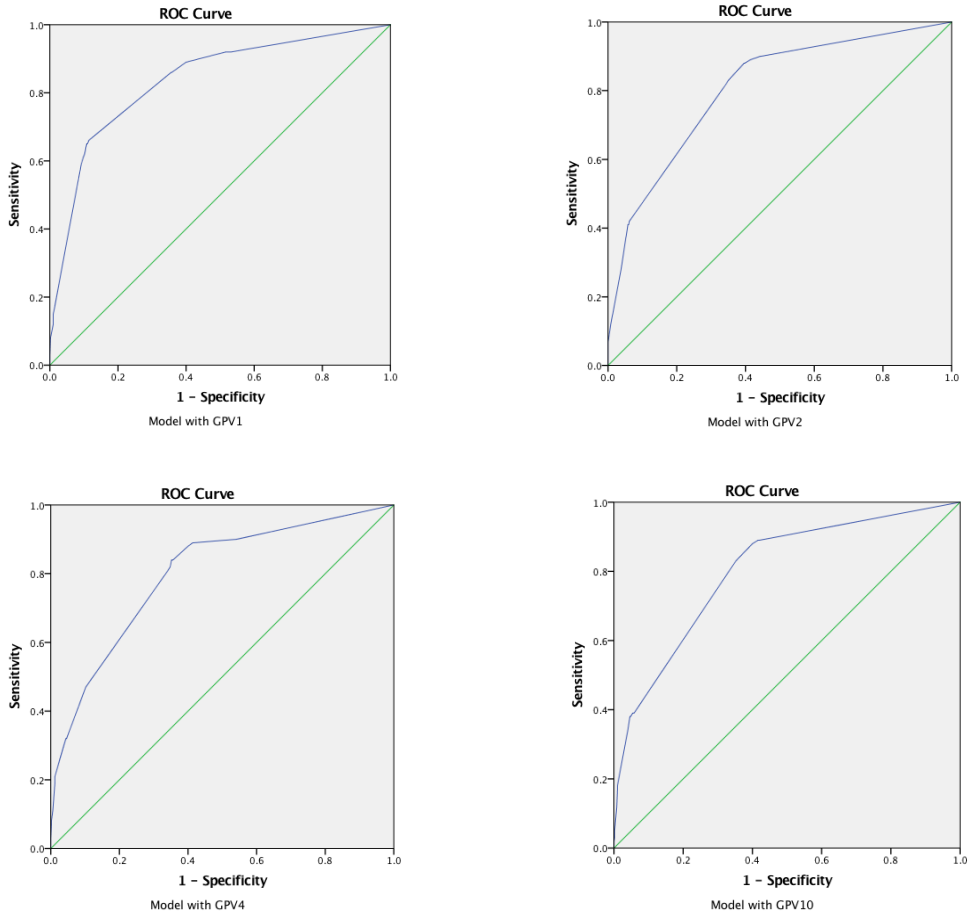


Figure 5 ROC curves for models with GPV1, GPV2, GPV4 and GPV10 variables

9 DISCUSSION

This study provides a comprehensive overview of recent scientific studies that analyse the risk factors of pre-eclampsia and provide quantitative description of the risk. It can be concluded from the review that until now no reliable predictor of pre-eclampsia has been found. The most reliable predictive models are multivariable models but even the best of them do not seem to be accurate enough. The best models so far have the area under the ROC curve at the level 0.8, which indicates that there may be too much trade-off between sensitivity and specificity. This can be illustrated on the model developed in the simulation exercise for non-genetic variables. This model has the AUC = 0.792. The model parameters can be modulated so that the model is more sensitive at the expense of specificity. For example, it can correctly predict 32 % of pre-eclampsia cases and 96 % of healthy cases, which means this model has very weak sensitivity but high specificity. If we improve sensitivity up to maximum and correctly predict 89 % of cases we significantly lose specificity and correctly predict only 58 % of healthy cases. To understand what this means in the clinical practice it can be described in the following way: if the model was applied to predict risk of all pregnant women in Slovakia each percentage point of increase of sensitivity means that about 10 more patients[1] with the high risk of pre-eclampsia (who may receive a prophylactic treatment) are correctly identified. One percentage point of decrease in specificity means that about 500 more patients[2] are incorrectly identified as risky (for who a prophylactic treatment is useless or harmful).

The simulation showed that the multivariable model can be moderately improved by adding more independent variables even if these variables are poor predictors individually. The GPV10 variable is a good example. Non-significant in the comparison of frequencies of this mutation in the case control study, it is significant and one of the strongest predictors in the full multivariable model. Adding this variable improved the accuracy of the reference model substantially. The question whether these improved models are more suitable for clinical practice should be answered by practitioners who are in the best position to decide what the right balance between the sensitivity and specificity for the disease is and whether an imperfect but simple model meets their needs.

[1] Based on an assumption of pre-eclampsia incidence of 1000 cases a year.
[2] Based on an assumption of 50,000 pregnancies a year.

Because the challenges above the quest for better predictive variables to enter multivariable models continues. This study demonstrates that there is a feasible methodology to investigate new predictors that could improve the predictive ability multivariable models. As the least explored predictors are genetic predictors this study focuses on the ERVW-1 gene polymorphism but the described methodology can be applied to any other predictor that can be considered as an independent variable in the model.

The proposed methodology is feasible but its application for ERVW-1 mutations as predictors of pre-eclampsia has some limitations and shortcomings that were discovered during the simulation. The most important limit is the sample size. Because pre-eclampsia is rare it may be impossible to get the required sample size. For a case sample with n = 100 almost all cases of pre-eclampsia in one year in the Bratislava region would have to enter the study as there were 107 pre-eclampsia cases in the whole Bratislava region in 2015. The simulation shows that a sample size of 100 is already too small to identify risk factors with low prevalence (< 0.05) and low risk.

Because of the low frequency of cases and the low prevalence of genetic risk factors in the population the level of risk is rather uncertain (i.e. the confidence interval is wide). This has effect on the added value these predictors bring to the predictive model. The simulation showed that the best improvement of accuracy was achieved by the predictor with high risk (OR = 4) rate and high prevalence (P = 0.2). On the other hand the predictor with very high relative risk (OR = 8) but low prevalence (P = 0.01) had much more limited ability to improve the accuracy. We can conclude that not only the risk but also the prevalence of the risk factor must be considered when designing a new predictive model.

This has also implications for expectations from further investigation of ERVW-1 polymorphism in relation to pre-eclampsia. One implication is that small sample studies may not identify association of polymorphism with pre-eclampsia if the risk is small and the prevalence of the mutation is low. Unfortunately, the prevalence of most ERVW-1 mutations is low to extremely low therefore small sample for many of these mutations may not produce positive results. From the perspective of prediction this is not a big loss because, as the simulation showed, low prevalence predictors included in the model do not improve the predictive ability substantially.

10 CONCLUSIONS

This study provided a comprehensive methodology and guidance for a future study of risk factors of pre-eclampsia and for the development of a predictive model. The study builds on the research in ERVW-1 gene polymorphism and hypothesises that this polymorphism could be a reliable predictor of pre-eclampsia that will improve the accuracy of predictive model, so the model can be used in clinical practice for pre-eclampsia risk assessment and stratification. The methodology can be used as it is presented with the two-fold objective to prove the predictive value of ERVW-1 polymorphism and construct an accurate predictive model. The presented methodology can also be adapted to other research objectives and hypotheses concerning pre-eclampsia as the theoretical basis and statistical tools are generic and can be applied, for example, for different sets of predictors. Thus, the methodology can be used for testing the predictive value of any new predictor of pre-eclampsia.

The hypothesis that ERVW-1 gene polymorphism could be a predictor of pre-eclampsia was not selected by chance. There is a plausible theory about biological processes that lead to pre-eclampsia. This theory assumes that ERVW-1 mutation leads to underexpression of syncytin-1, which leads to the abnormal development of placenta and to clinical manifestations of pre-eclampsia. While the details of this biological process are not fully known the confirmation of ERVW-1 gene polymorphism as a strong predictor would give the theory more credibility.

This indicates another dimension of this study and reinforces the call for continuation of research in this area. A future experimental study based on the methodology and the insight provided by this study can have an added value in broadening the knowledge about pre-eclampsia. As the aetiology of pre-eclampsia is not fully understood, the stochastic methods can provide a hint for the direction of research in this area. If a new genetic predictor is found to be strong it will attract the attention of researchers who may want to identify and explain the underlying causal process. This will improve the knowledge of pre-eclampsia, which in turn will lead to better diagnostics, prediction, prophylaxis and treatment. The study showed that even relatively small scale studies can identify such strong predictors.

It is hoped that this study will have some impact, either a direct one through the development of a new predictive tool potentially used in clinical practice or by contributing to the knowledge about pre-eclampsia. There is a hope that the ultimate impact of this study will be more mothers and their children enjoying a healthier life.

ANNEX 1 MATHEMATICAL BASIS OF THE MINIMAL SAMPLE SIZE CALCULATION

By applying the method in Wang and Chow (2007) the minimum sample size can be calculated for different confidence levels and power, and proportions of the risk factor in the case and the control group by using the formula:

$$n = (Z_{\alpha/2} + Z_{\beta})^2 * (\pi_1(1 - \pi_1) + \pi_2(1 - \pi_2)) / (\pi_1 - \pi_2)^2 \quad (1)$$

where $Z_{\alpha/2}$ is the critical value of the normal distribution at α/2 (e.g. for a confidence level of 95%, α is 0.05 and the critical value is 1.96), Z_{β} is the critical value of the normal distribution at β (e.g. for a power of 80%, β is 0.2 and the critical value is 0.84), and π_1 and π_2 are the expected proportions of the two groups.

The expected proportions of the risk factor in the case and the control group are not known. They can be expressed in terms of parameters that may be known or can serve for the development of plausible scenarios. In our case these parameters are:
(1) the prevalence of the ERVW-1 mutation, and
(2) the relative risk of ERVW-1.

Once the specific ERVW-1 mutation is selected for analysis the approximate global prevalence of this mutation can be found in the 1000 Genome Browser (2018). The relative risk can be estimated by comparing the expected risk of the genetic mutation to the known relative risk of other risk factors identified in previous scientific studies such as the risk factors described in chapter 3.

The proportions and π_1 and π_2 can be then expressed as:

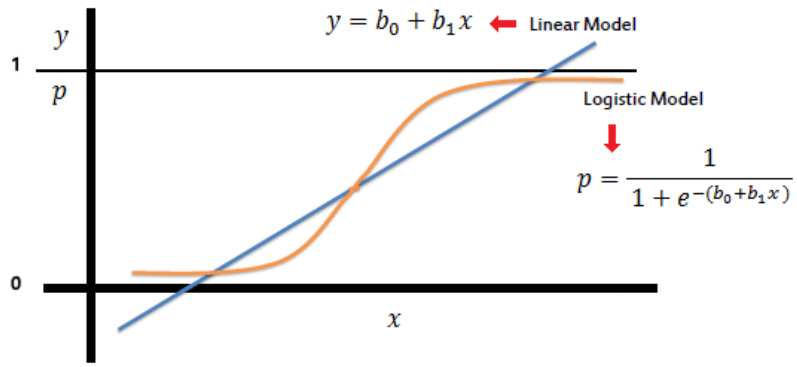
$$\pi_1 = (RR * P / (1 - P)) / (RR * P / (1 - P) + 1) \quad (2)$$

$$\pi_2 = P / (1 - F) - (F / (1 - F)) / ((1 - P) / (RR * P) + 1) \quad (3)$$

where RR is the relative risk, P is the prevalence of the ERVW-1 mutation in the population and F is the frequency of pre-eclampsia in the population of pregnant women (F = 0.02).

ANNEX 2 MATHEMATICAL BASIS OF A PREDICTIVE MODEL BASED ON THE LOGISTIC REGRESSION

Together with the linear discriminant analysis and the probit regression, the logistic regression is one of the most frequently used methods for multivariate analysis. These alternative methods use different functions for the probability of expected outcome. The linear method uses a linear function and the probit method uses the cumulative normal distribution function, compared to the logistic method that uses the logistic function.



The intercept b_0 shifts curves to the left and the coefficient b_1 determines the steepness of the curve (Sayad, 2018)

Figure 6. Graphical comparison of single variable linear and logistic regression

Mathematically, the multivariate logistic function can be expressed as:

$$p = \frac{1}{1 + e^{b_0 + b_1 x_1 + \cdots + b_n x_n}} \quad (4)$$

where p is the probability that the dependant variable equals a case (in this study it means that pre-eclampsia occurs in a patient) and x_i, i=1,…,n, are independent variables (predictors). The equation above can be converted into the logit form:

$$ln\left(\frac{p}{1-p}\right) = b_0 + b_1 x_1 + \cdots + b_n x_n \quad (5)$$

where $\frac{p}{1-p}$ is odds of the case and ln(odds) is called logit or log-odds (Stanford University, 2018). This formula shows the similarity of the logistic regression to the linear regression. Essentially, the logistic regression is a special form of linear regression where the y-value is the natural logarithm of odds of the case. This helps to interpret the results as a unit

change of the coefficient x_i multiplies the odds of the case by e^{b_i}. There is a direct relationship between coefficients b_i and the odds ratio (OR):

$$OR = e^{b_1 + \cdots + b_n} \quad (6)$$

The similarity of the logistic regression to the linear regression implies that similar tests can be used to check the goodness of fit and significance of the coefficients. Similar to the R^2 test for linear regression, several pseudo-R^2 test were developed for the logistic regression, including McFadden, Efron, Cox and Snell, Nagelkerke and the count tests, each of them expressing a measure of 'goodness of fit' of the regression model. The Wald test is the equivalent of the t-test in linear regression and is often used to assess the statistical significance of each individual predictor (Cohen et al., 2003).

The advantage of the logistic regression is that it can deal with unbalanced data, i.e. data that have the ratio between the number of cases and the number of controls in the sample higher than in the normal population. This is important for the situation when cases are rare, such as in pre-eclampsia. It means that we can take more cases of pre-eclampsia to the sample, while the number of non-pre-eclamptic women will be disproportionally low. The calculated coefficients b_i, i = 1, …, n, expressing the effect of predictive variables will be the same but the intercept b_0 will have to be corrected using the formula:

$$b_0' = b_0 + ln\left(\frac{\pi'}{1-\pi'}\right) - ln\left(\frac{\pi}{1-\pi}\right) \quad (7)$$

where π is the frequency in the sample and π' is the frequency in the population (Stanford University, 2018).

REFERENCES

1000 Genome Browser (2018). http://internationalgenome.org, 24/9/2018.

a Rogvi R, Forman J, Damm P, Greisen G (2012). Women born preterm or with inappropriate weight for gestational age are at risk of subsequent gestational diabetes and pre-eclampsia. PLOS ONE, 7(3):e34001.

Abalos E, Cuesta C, Carroli G, Qureshi Z, Widmer M, Vogel JP (2014). Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: a secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. BJOG, 121 Suppl 1:14–24.

Abou-Nassar K, Carrier M, Ramsay T, Rodger MA (2011). The association between antiphospholipid antibodies and placenta mediated complications: a systematic review and meta-analysis. Thromb Res, 128(1):77–85.

Austdal M, Tangerås LH, Skråstad RB, Salvesen KÅ, Austgulen R, Iversen AC, Bathen TF (2015). First Trimester Urine and Serum Metabolomics for Prediction of Preeclampsia and Gestational Hypertension: A Prospective Screening Study. International Journal of Molecular Sciences, 16(9):21520–21538.

Armitage P and Berry G (1993). Statistical Methods in Medical Research. Oxford: Blackwell Scientific Publications.

Baweja S, Kent A, Masterson R, Roberts S, McMahon L (2011). Prediction of pre-eclampsia in early pregnancy by estimating the spot urinary albumin: creatinine ratio using high-performance liquid chromatography. BJOG, 118(9):1126–1132.

Bezinque A, Noyes SL, Kirmiz S, Parker J, Dey S, Kahnoski RJ, Lane BR (2017). Prevalence of Proteinuria and Other Abnormalities in Urinalysis Performed in the Urology Clinic. Urology, 103:34-38.

Bdolah Y, Lam C, Rajakumar A, Shivalingappa V,Mutter W, Sachs BP (2008). Twin pregnancy and the risk of preeclampsia: bigger placenta or relative ischemia? Am J Obstet Gynecol, 198(4):428e1-6.

Bigelow CA, Pereira GA, Warmsley A, Cohen J, Getrajdman C, Moshier E (2013). Risk factors for new-onset late postpartum preeclampsia in women without a history of preeclampsia. Am J Obstet Gynecol, 210(4):338e1-8.

Bilano V, Ota E, Ganchimeg T, Mori R, Souza J (2014). Risk factors of pre-eclampsia /eclampsia and its adverse outcomes in low- and middle-income countries: a WHO secondary analysis. PLOS ONE 9(3):e91198.

Bodnar LM, Ness RB, Markovic N, Roberts JM (2005). The risk of preeclampsia rises with increasing prepregnancy body mass index. Ann Epidemiol, 15(7):475–482.

Bolin M, Wikstrom A, Wiberg-Itzel E, Olsson A, Ringvall M, Sundstrom-Poromaa I (2012). Prediction of Preeclampsia by Combining Serum Histidine-Rich Glycoprotein and Uterine Artery Doppler. Am J Hypertens, 25(12):1305–1310.

Boyd HA, Tahir H, Wohlfahrt J, Melbye M (2013). Associations of personal and family preeclampsia history with the risk of early-, intermediate- and late-onset preeclampsia. Am J Epidemiol, 178(11):1611–1619.

Bramham K, Briley AL, Seed PT, Poston L, Shennan AH, Chappell LC (2011). Pregnancy outcome in women with chronic kidney disease: a prospective cohort study. Reprod Sci, 18(7):623–630.

Bramham K, Parnell B, Nelson-Piercy C, Seed PT, Poston L, Chappell LC (2014). Chronic hypertension and pregnancy outcomes: systematic review and meta-analysis. BMJ, 348:g2301.

Caradeux J, Serra R, Nien J, Perez-Sepulveda A, Schepeler M, Guerra F (2013). First trimester prediction of early onset preeclampsia using demographic, clinical, and sonographic data: a cohort study. Prenat Diagn, 33(8):732–736.

Chappell LC, Duckworth S, Seed PT, Griffin M, Myers J, Mackillop L (2013). Diagnostic accuracy of placental growth factor in women with suspected preeclampsia: a prospective multicenter study. Circulation, 128(19):2121–2131.

Chesley LC, Cooper DW (1986). Genetics of hypertension in pregnancy: possible single gene control of pre-eclampsia and eclampsia in the descendants of eclamptic women. Br J Obstet Gynaecol, 93(9):898–908.

Cnossen JS, Morris RK, ter Riet G, Mol BWJ, van der Post JAM, Joris AM, Coomarasamy A (2008). Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. Can Med Assoc J, 178(6):701–711.

Cohen, J, Cohen, P, West, SG, Aiken, LS (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.). Mahwah: Luarence Erlbaum Associates Publishers.

Conde-Agudelo A, Belizan JM (2000). Risk factors for pre-eclampsia in a large cohort of Latin American and Caribbean women. BJOG, 107(1):75–83.

Conde-Agudelo A, Villar J, Lindheimer M (2008). Maternal infection and risk of preeclampsia: systematic review and meta-analysis. Am J Obstet Gynecol, 198(1):7–22.

Craici IM, Wagner SJ, Bailey KR, Fitz-Gibbon PD, Wood-Wentz CM, Turner ST (2013). Podocyturia predates proteinuria and clinical features of preeclampsia: longitudinal prospective study. Hypertension, 61(6):1289–1296.

Dawonauth L, Rademacher L, L'Omelette A, Jankee S, Yan M, Jeeawoody R (2014). Urinary inositol phosphoglycan-P type: Near patient test to detect preeclampsia prior to

clinical onset of the disease. A study on 416 pregnant Mauritian women. J Reprod Immunol, 101:148–152.

Dawson LM, Parfrey PS, Hefferton D, Dicks EL, Cooper JM, Young D (2002). Familial Risk of Preeclampsia in Newfoundland: A Population-Based Study. J Am Soc Nephrol, 13(7):1901–1906.

Dekker G, Robillard P, Roberts C (2011). The etiology of preeclampsia: the role of the father. J Reprod Immunol, 89(2):126–132.

Dempsey JC, Williams MA, Luthy DA, Emanuel I, Shy K (2003). Weight at birth and subsequent risk of preeclampsia as an adult. Am J Obstet Gynecol, 189(2):494–500.

Di Lorenzo G, Ceccarello M, Cecotti V, Ronfani L, Monasta L, Brumatti LV (2012). First trimester maternal serum PIGF, free [beta]-hCG, PAPP-A, PP-13, uterine artery Doppler and maternal history for the prediction of preeclampsia. Placenta, 33(6):495.

Elia EG, Robb AO, Hemming K, Price MJ, Riley RD, French-Constant A, Stock SJ (2017). Is the first urinary albumin/creatinine ratio (ACR) in women with suspected preeclampsia a prognostic factor for maternal and neonatal adverse outcome? A retrospective cohort study. Acta Obstetricia et Gynecologica Scandinavica, 96(5):580–588.

Endeshaw M, Abebe F, Bedimo M, Asart A (2015). Diet and Pre-eclampsia: A Prospective Multicentre Case-Control Study in Ethiopia. Midwifery 31(6):617–624.

Engels T, Pape J, Schoofs K, Henrich W, Verlohren S (2013). Automated measurement of sFlt1, PlGF and sFlt1/PlGF ratio in differential diagnosis of hypertensive pregnancy disorders. Hypertens Pregnancy, 32(4):459–473.

England L, Zhang J (2007). Smoking and risk of preeclampsia: a systematic review. Front Biosci, 12:2471–2483.

English FA, Kenny LC, McCarthy FP (2015). Risk factors and effective management of preeclampsia. Integr Blood Press Contro,l 8:7–12.

Esplin M, Fausett M, Fraser A, Kerber R, Mineau G, Carrillo J (2001). Paternal and maternal components of the predisposition to preeclampsia. N Engl J Med, 344(12):867–872.

Filetti LC, Imudia AN, Al-Safi Z, Hobson DT, Awonuga AO, Bahado-Singh RO (2012). New onset delayed postpartum preeclampsia: different disorders? J Matern Fetal Neonatal Med, 25(7):957–60.

Firoz T, Sanghvi H, Merialdi M, von Dadelszen P (2011). Pre-eclampsia in low and middle income countries. Best Pract Res Clin Obstet Gynaecol, 25(4):537–48.

Forest J, Theriault S, Masse J, Bujold E, Giguere Y (2014). Soluble Fms-like tyrosine kinase-1 to placental growth factor ratio in mid-pregnancy as a predictor of preterm preeclampsia in asymptomatic pregnant women. Clin Chem Lab Med, 52(8):1169–1178.

Gallo D, Poon L, Fernandez M, Wright D, Nicolaides K (2014). Prediction of Preeclampsia by Mean Arterial Pressure at 11–13 and 20–24 Weeks' Gestation. Fetal Diagn Ther, 36(1):28–37.

Gao Y, He Z, Wang Z, Luo Y, Sun H, Zhou Y, Huang L, Li M, Fang Q, Jiang S (2012). Increased expression and altered methylation of HERVWE1 in the human placentas of smaller fetuses from monozygotic, dichorionic, discordant twins. PLOS One, 7(3):e33503.

Garb, JL (1996). Understanding medical research: a practitioner's guide. Boston: Little, Brown and Company.

Ghosh S, Raheja S, Tuli A, Raghunandan C, Agarwal S (2013b). Serum placental growth factor as a predictor of early onset preeclampsia in overweight/obese pregnant women. J Am Soc Hypertens, 7(2):137–148.

Ghosh SK, Raheja S, Tuli A, Raghunandan C, Agarwal S (2013a). Is serum placental growth factor more effective as a biomarker in predicting early onset preeclampsia in early second trimester than in first trimester of pregnancy? Arch Gynecol Obstet, 287(5):865–873.

Goetzinger K, Tuuli M, Cahill A, Macones G, Odibo A (2014). Development and Validation of a Risk Factor Scoring System for First-Trimester Prediction of Preeclampsia. Am J Perinatol, 31(12):1049–1055.

Goodwin AA, Mercer BM (2005). Does maternal race or ethnicity affect the expression of severe preeclampsia? Am J Obstet Gynecol, 193(3 Pt 2):973–978.

Han, A, Bujold E, Belizan M, Jaime J, Sharma S, Magee LA (2016) Preventing pre-eclampsia and its complications. In: Magee, LA, von Dadelszen, P, Stones, W, Mathai, M eds. The FIGO Textbook of Pregnancy Hypertension. London: The Global Library of Women's Medicine, p. 101-122.

Hanita O, Alia N, Zaleha A, Azlin M (2014). Serum soluble FMS-like tyrosine kinase 1 and placental growth factor concentration as predictors of preeclampsia in high risk pregnant women. Malays J Pathol, 36(1):19–26.

Harris, M., Taylor G (2008). Medical statistics made easy. Bloxham: Scion Publishing.

Hirose N, Ohkuchi A, Rie US (2014). Risk of Preeclampsia in Women with CKD, Dialysis or Kidney Transplantation. Med J Obstet Gynecol, 2(2):1028.

Hoenig JM, Heisay DM (2001). The abuse of power: the pervasive fallacy of power calculations for data analysis. The American Statistician, 55(1):19-24.

Holder BS, Tower CL, Abrahams VM, Aplin JD (2012). Syncytin-1 in the human placenta. Placenta, 33(6):460-6.

Huang Q, Li J, Wang F, Oliver MT, Tipton T, Gao Y, Jiang SW (2013). Syncytin-1 modulates placental trophoblast cell proliferation by promoting G1/S transition. Cellular Signalling, 25(4):1027–1035.

INTERPREGGEN (Genetic studies of pre-eclampsia in Central Asian and European populations), Final Report Summary (2017). (http://cordis.europa.eu/result/rcn/194531_en.html), 29/8/2017.

Jim B, Mehta S, Qipo A, Kim K, Cohen HW, Moore RM (2014). A comparison of podocyturia, albuminuria and nephrinuria in predicting the development of preeclampsia: A prospective study: PLOS One, 9(7):e101445.

Kasawara K, do Nascimento S, Costa M, Surita F, e Silva J (2012). Exercise and physical activity in the prevention of pre-eclampsia: systematic review. Acta Obstet Gynecol Scand, 91(10):1147–1157.

Kashanian M, Aghbali F, Mahali N (2013). Evaluation of the diagnostic value of the first-trimester maternal serum high-sensitivity C-reactive protein level for prediction of pre-eclampsia. J Obstet Gynaecol Res, 39(12):1549–1554.

Kassebaum NJ, Bertozzi-Villa A, Coggeshall MS, Shackelford KA, Steiner C, Heuton KR (2014). Global, regional, and national levels and causes of maternal mortality during 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. Lancet, 384(9947):980–1004.

Keikkala E, Vuorela P, Laivuori H, Romppanen J, Heinonen S, Stenman U (2013). First trimester hyperglycosylated human chorionic gonadotrophin in serum – A marker of early-onset preeclampsia. Placenta, 34(11):1059–1065.

Kelder TP, Penning ME, Uh H, Cohen D, Bloemenkamp KWM, Bruijn JA (2012). Quantitative polymerase chain reaction–based analysis of podocyturia is a feasible diagnostic tool in preeclampsia. Hypertension, 60(6):1538–1544.

Kenny LC, Black MA, Poston L, Taylor R, Myers JE, Baker PN (2014). Early pregnancy prediction of preeclampsia in nulliparous women, combining clinical risk and biomarkers: the Screening for Pregnancy Endpoints (SCOPE) international cohort study. Hypertension, 64(3):644–52.

Kenny LC, Black MA, Poston L, Taylor R, Myers JE, Baker PN, McCowan LM, Simpson NA, Dekker GA, Roberts CT, Rodems K, Noland B, Raymundo M, Walker JJ, North RA (2014). Early pregnancy prediction of preeclampsia in nulliparous women, combining clinical risk and biomarkers: the Screening for Pregnancy Endpoints (SCOPE) international cohort study. Hypertension, 64(3):644-52.

Khalil A, Syngelaki A, Maiz N, Zinevich Y, Nicolaides K (2013). Maternal age and adverse pregnancy outcome: a cohort study. Ultrasound Obstet Gynecol, 42(6):634–643.

Khan KS, Wojdyla D, Say L, Gulmezoglu AM, Van Look PF (2006). WHO analysis of causes of maternal death: a systematic review. Lancet, 367(9516):1066–74.

Kleinrouweler CE, Bossuyt PMM, Thilaganathan B, Vollebregt KC, Arenas Ramirez J, Ohkuchi A (2013). Value of adding second-trimester uterine artery Doppler to patient

characteristics in identification of nulliparous women at increased risk for pre-eclampsia: an individual patient data meta-analysis. Ultrasound Obstet Gynecol, 42(3):257–267.

Knerr I, Huppertz B, Weigel C, Dötsch J, Wich C, Schild RL, Beckmann MW, Rascher W (2004). Endogenous retroviral syncytin: compilation of experimental research n syncytin and its possible role in normal and disturbed human placentogenesis. Mol Hum Reprod, 10(8):581-8.

Kuc S, Koster MP, Franx A, Schielen PC, Visser GH (2013). Maternal characteristics, mean arterial pressure and serum markers in early prediction of preeclampsia. PLOS One, 8(5):e63546.

Kuc S, Koster MP, Pennings JL, Hankemeier T, Berger R, Harms AC, Dane AD, Schielen PC, Visser GH, Vreeken RJ (2014). Metabolomics profiling for identification of novel potential markers in early prediction of preeclampsia. PLOS One, 9(5):e98540.

Kurki T, Hiilesmaa V, Raitasalo R, Mattila H, Ylikorkala O (2000). Depression and anxiety in early pregnancy and risk for preeclampsia. Obstet Gynecol, 95(4):487–490.

Lai J, Pinas A, Poon L, Agathokleous M, Nicolaides, K (2013). Maternal Serum Placental Growth Factor, Pregnancy-Associated Plasma Protein-A and Free beta-Human Chorionic Gonadotrophin at 30–33 Weeks in the Prediction of Pre-eclampsia. Fetal Diagn Ther, 33(3):164–172.

Langbein M, Strick R, Strissel PL, Vogt N, Parsch H, Beckmann MW, Schild RL (2008). Impaired cytotrophoblast cell-cell fusion is associated with reduced syncytin and increased apoptosis in patients with placental dysfunction. Mol Reprod Dev., 75(1):175-183.

Lecarpentier E, Tsatsaris V, Goffinet F, Cabrol D, Sibai B, Haddad B (2013). Risk factors of superimposed preeclampsia in women with essential chronic hypertension treated before pregnancy. PLOS One, 8(5):e62140.

Leeflang MMG, Cnossen JS, van der Post JAM, Mol BWJ, Khan KS, ter Riet G. Accuracy of fibronectin tests for the prediction of pre-eclampsia: a systematic review. Eur J Obstet Gynecol, 133(1):12–19.

Lin J, August P (2005). Genetic thrombophilias and preeclampsia: a meta-analysis. Obstet Gynecol, 105(1):182–192.

Lin S, Leonard D, Co MA, Mukhopadhyay D, Giri B, Perger L, Beeram MR, Kuehl TJ, Uddin MN (2015). Pre-eclampsia has an adverse impact on maternal and fetal health. Tranls Res., 165(4):449-463.

Lisonkova S, Joseph KS (2013). Incidence of preeclampsia: risk factors and outcomes associated with early- versus late-onset disease. Am J Obstet Gynecol, 209(6):544.e1-544.e12.

Lisonkova S, Sabr Y, Mayer C, Young C, Skoll A, Joseph KS (2014). Maternal morbidity associated with early-onset and late-onset preeclampsia. Obstet Gynecol., 124(4):771-81.a

Magee LA, Pels A, Helewa M, Rey E, von Dadelszen P (2014). Diagnosis, evaluation, and management of the hypertensive disorders of pregnancy. Pregnancy Hypertension, 4(2):105–45.

Magee, LA, von Dadelszen, P, Stones, W, Mathai, M (2016). Introduction. In: Magee, LA, von Dadelszen, P, Stones, W, Mathai, M eds. The FIGO Textbook of Pregnancy Hypertension. London: The Global Library of Women's Medicine, p. 33-61 p. xiv-xviii.

McElrath T, Lim K, Pare E, Rich-Edwards J, Pucci D, Troisi R (2012). Longitudinal evaluation of predictive value for preeclampsia of circulating angiogenic factors through pregnancy. Am J Obstet Gynecol, 207(5): 407.e1-e7.

McElrath T, Lim K, Pare E, Rich-Edwards J, Pucci D, Troisi R (2012). Longitudinal evaluation of predictive value for preeclampsia of circulating angiogenic factors through pregnancy. Am J Obstet Gynecol, 207(5): 407.e1-e7.

Meads C, Cnossen J, Meher S, Juarez-Garcia A, ter Riet G, Duley L (2008). Methods of prediction and prevention of pre-eclampsia: systematic reviews of accuracy and effectiveness literature with economic modelling. Health Technol Assess., 12 (6): 1-270.

Minassian C, Thomas SL, Williams DJ, Campbell O, Smeeth L (2013). Acute maternal infection and risk of pre-eclampsia: a population-based case-control study. PLOS One, 8(9):e73047.

Mostello D, Kallogieri D, Tungsiripat R, Leet T (2008). Recurrence of preeclampsia: effects of gestational age at delivery of the first pregnancy, body mass index, paternity, and interval between births. Am J Obstet Gynecol, 199(1):55.e1-55.e7.

Myers J, Kenny L, McCowan L, Chan E, Dekker G, Poston L (2013). Angiogenic factors combined with clinical risk factors to predict preterm pre-eclampsia in nulliparous women: a predictive test accuracy study. BJOG, 120(10):1215–1223.

Myklestad K, Vatten L, Salvesen K, Smith G, Romundstad P (2011). Hypertensive disorders in pregnancy and paternal cardiovascular risk: a population-based study. Ann Epidemiol, 21(6):407–412.

Napolitano R, Melchiorre K, Arcangeli T, Dias T, Bhide A, Thilaganathan B (2012). Screening for pre-eclampsia by using changes in uterine artery Doppler indices with advancing gestation. Prenat Diagn, 32(2):180–184.

Národné centrum zdravotníckych informacií (2017). Starostlivosť o rodičku a novorodenca v SR 2015. http://www.nczisk.sk/Documents/publikacie/2015/zs1751.pdf

Národné centrum zdravotníckych informacií (2018). Health Statistics Yearbook of the Slovak Republic 2016.
http://www.nczisk.sk/Documents/rocenky/2016/Zdravotny_stav_obyvatelstva.pdf

Odegard RA, Vatten LJ, Nilsen ST, Salvesen KA, Austgulen R (2000). Risk factors and clinical manifestations of pre-eclampsia. BJOG, 107(11):1410–1416.

Olayemi O, Strobino D, Aimakhu C, Adedapo K, Kehinde A, Odukogbe A (2010). Influence of duration of sexual cohabitation on the risk of hypertension in nulliparous parturients in Ibadan: A cohort study. Aust N Z J Obstet Gynaecol, 50(1):40–44.

Opdahl S, Henningsen A, Tiitinen A, Bergh C, Pinborg A, Romundstad P (2015). Risk of hypertensive disorders in pregnancies following assisted reproductive technology: a cohort study from the CoNARTaS group. Hum Reprod, 30(7):1724–1731.

Papageorghiou AT, Yu CKH, Cicero S, Bower S, Nicolaides KH (2002). Second-trimester uterine artery Doppler screening in unselected populations: a review J Matern Fetal Neonatal Med, 12(2):78–88.

Papageorghiou AT, Yu CKH, Cicero S, Bower S, Nicolaides KH (2002). Second-trimester uterine artery Doppler screening in unselected populations: a review J Matern Fetal Neonatal Med, 12(2):78–88.

Pare E, Parry S, McElrath TF, Pucci D, Newton A, Lim K (2014). Clinical risk factors for preeclampsia in the 21st century. Obstet Gynecol, 124(4):763–770.

Park FJ, Leung CHY, Poon LCY, Williams PF, Rothwell SJ, Hyett JA (2013). Clinical evaluation of a first trimester algorithm predicting the risk of hypertensive disease of pregnancy. Aust N Z J Obstet Gynaecol, 53(6):532–539.

Park FJ, Leung CHY, Poon LCY, Williams PF, Rothwell SJ, Hyett JA (2013). Clinical evaluation of a first trimester algorithm predicting the risk of hypertensive disease of pregnancy. Aust N Z J Obstet Gynaecol, 53(6):532–539.

Pattinson RC, Hall M (2003). Near misses: a useful adjunct to maternal death enquiries. Br Med Bull, 67:231–43.

Pels A, van Dedelszen P, Engelbrecht S, Ryan H, Bellad M, Lalonde A, Magee LA (2016). Timing and mode of delivery. In: Magee, LA, von Dadelszen, P, Stones, W, Mathai, M eds. The FIGO Textbook of Pregnancy Hypertension. London: The Global Library of Women's Medicine, p. 167-183.

Perales A, Delgado JL, De La Calle M, García-Hernández JA, Escudero AI, Campillos JM, Sarabia MD, Laíz B, Duque M, Navarro M, Calmarza P, Hund M, Álvarez FV (2016). sFlt1/PlGF for early-onset pre-eclampsia prediction: STEPS (Study of Early Pre-eclampsia in Spain). Ultrasound Obstet Gynecol, 2016 Nov 24.

Priscakova P, Konkolova J, Petrovic R, Lipov J, Minarik G, Bohmer D, Repiska V, Gbelecova H (2016). ERVW-1 gene polymorhysm related to preeclampsia. Bratisl Med J, 117(6):340-344.

Priscakova P. (2017). Charakterizacia genetickej variability a genovej expresie ERVW-1 v placentach preeklamptickych zien. Bratislava: Comenius University. Unpublished PhD thesis.

Priscakova P. (2018). Personal communication.

Poon LCY, Kametas NA, Maiz N, Akolekar R, Nicolaides KH (2009). First-Trimester Prediction of Hypertensive Disorders in Pregnancy. Hypertension, 53(5): 812–818.

Qiu C, Williams MA, Leisenring WM, Sorensen TK, Frederick IO, Dempsey JC (2003). Family history of hypertension and type 2 diabetes in relation to preeclampsia risk. Hypertension, 41(3):408–413.

Redman CW, Sargent IL (2005). Latest Advances in Understanding Preeclampsia. Science, 308(5728):1592–1594.

Rigo J, Boze T, Derzsy Z, Derzbach L, Treszl A, Lazar L (2006). Family history of early-onset cardiovascular disorders is associated with a higher risk of severe preeclampsia. Eur J Obstet Gynecol Reprod Biol, 128(1–2):148–151.

Rizos D, Eleftheriades M, Karampas G, Rizou M, Haliassos A, Hassiakos D (2013). Placental growth factor and soluble fms-like tyrosine kinase-1 are useful markers for the prediction of preeclampsia but not for small for gestational age neonates: a longitudinal study. Eur J Obstet Gynecol Reprod Biol, 171(2):225–230.

Roland CS, Hu J, Ren CE, Chen H, Li J, Varvoutis MS, Leaphart LW, Byck DB, Zhu X, Jiang SW (2016). Morphological changes of placental syncytium and their implications for the pathogenesis of preeclampsia. Cell Mol Life Sci., 73(2):365-76.

Sayad, S (2018). Logistic regression. http://www.saedsayad.com/logistic_regression.htm, 24/9/2018.

Schneuer F, Nassar N, Khambalia A, Tasevski V, Ashton A, Morris J (2012). First trimester screening of maternal placental protein 13 for predicting preeclampsia and small for gestational age: In-house study and systematic review. Placenta, 33(9):735–740.

Shamsi U, Saleem S, Nishter N (2013). Epidemiology and risk factors of preeclampsia; an overview of observational studies. Epidemiology, 6(4):368–374.

Sibai B, Caritis S, Hauth J, Lindheimer M, VanDorsten J, MacPherson C (2000). Risks of preeclampsia and adverse neonatal outcomes among women with pregestational diabetes mellitus. National Institute of Child Health and Human Development Network of Maternal-Fetal Medicine Units. Am J Obstet Gynecol 2000;182(2):364–369.

Sibai B, Hauth J, Caritis S, Lindheimer M, MacPherson C, Klebanoff M (2000). Hypertensive disorders in twin versus singleton gestations. National Institute of Child

Health and Human Development Network of Maternal-Fetal Medicine Units. Am J Obstet Gynecol, 182(4):938–942.

Skjaerven R, Wilcox AJ, Lie RT (2002). The interval between pregnancies and the risk of preeclampsia. N Engl J Med, 346(1):33–38.

Sohlberg S, Stephansson O, Cnattingius S, Wikstrom A (2011). Maternal body mass index, height, and risks of preeclampsia. Am J Hypertens, 25(1):120–125.

Staff AC, Benton SJ, von Dadelszen P, Roberts JM,Taylor RN, Powers RW (2013). Redefining preeclampsia using placenta-derived biomarkers. Hypertension, 61(5):932–42.

Stanford University (2018). Lagunita online course in applied statistics. https://lagunita.stanford.edu/c4x/HumanitiesScience/StatLearning/asset/classification.pdf, 24/9/2018

Teixeira PG, Reis ZSN, Andrade SP, Rezende CA, Lage EM, Velloso EP (2013). Presymptomatic prediction of preeclampsia with angiogenic factors, in high risk pregnant women. Hypertens Pregnancy, 32(3):312–320.

Thomopoulos C, Tsioufis C, Michalopoulou H, Makris T, Papademetriou V, Stefanadis C (2013). Assisted reproductive technology and pregnancy-related hypertensive complications: a systematic review. J Hum Hypertens, 27(3):148–157.

Tranquilli A, Landi B, Giannubilo SR, Sibai BM (2012). Preeclampsia - No longer solely a pregnancy disease. Pregnancy Hypertens., 2(4):350-357.

Ukah UV, Payne B, Cote AM, Hoodbhoy Z, von Dadelszen P (2016): Risk factors and predictors of pre-eclampsia. In: Magee, LA, von Dadelszen, P, Stones, W, Mathai, M eds. The FIGO Textbook of Pregnancy Hypertension. London: The Global Library of Women's Medicine, p. 75-100.

Vahdat M, Kashanian M, Sariri E, Mehdinia M (2012). Evaluation of the value of calcium to creatinine ratio for predicting of pre-eclampsia. J Matern Fetal Neonatal Med, 25(12):2793–2794.

von Dadelszen, P, Ayres de Campos, D, Barivalala, W (2016). Classification of the hypertensive disorders of pregnancy. In: Magee, LA, von Dadelszen, P, Stones, W, Mathai, M eds. The FIGO Textbook of Pregnancy Hypertension. London: The Global Library of Women's Medicine, p. 33-61.

Wang H, Chow SC (2007). Sample Size Calculation for Comparing Proportions. Wiley Encyclopedia of Clinical Trials. https://onlinelibrary.wiley.com/doi/10.1002/9780471462422.eoct005, 24/9/2018.

Wang Z, Wang P, Liu H, He X, Zhang J, Yan H (2013). Maternal adiposity as an independent risk factor for pre-eclampsia: a meta-analysis of prospective cohort studies. Obes Rev, 14(6):508–521.

WHO (2011). WHO recommendations for prevention and treatment of pre-eclampsia and eclampsia. Geneva: WHO.

WHO (2018). Raised blood pressure: situation and trends. WHO Global observatory. http://www.who.int/gho/ncd/risk_factors/blood_pressure_prevalence_text/en/, 24/9/2018

Wildman K, Bouvier-Colle M (2004). Maternal mortality as an indicator of obstetric care in Europe. Br J Obstet Gynaecol., 111 (2): 164-169.

Wittinghoff E, McCulloch CE (2007). Relaxing the rule of ten events per variable in

logistic and Cox regression. Am J Epidemiol. 165(6):710-8.

Woodward M (1999). Epidemiology: Study design and data analysis. Boca Raton: Chapman & Hall.

World Diabetes Foundation (2018). https://www.worlddiabetesfoundation.org/sites/default/files/Europe.pdf, 24/9/2018

Wright D, Akolekar R, Syngelaki A, Poon LC, Nicolaides KH (2012). A competing risks model in early screening for preeclampsia. Fetal Diagn Ther, 32(3):171–178.

Yu Y, Zhang S, Wang G, Hong X, Mallow EB, Walker SO (2013). The combined association of psychosocial stress and chronic hypertension with preeclampsia. Am J Obstet Gynecol, 209(5): 438.e1-e12.