Logistic Regression Model for determining risks factor for Hypertensive Disorders in Pregnancy

ANNA GABRIELA PÉREZ*, ELIZABETH TORRES RIVAS**, FRANCKLIN RIVAS ECHEVERRÍA***, CARLOS RIVAS ECHEVERRÍA^ 
*Escuela de Estadística
**Instituto de Estadística Aplicada y Computación (IEAC)
***Laboratorio de Sistemas Inteligentes
^Departamento de Toxicología y Farmacología
Universidad de Los Andes
Mérida, VENEZUELA

Abstract: - Preeclampsia (PE), one of the hypertensive disorders of pregnancy (HDP), is the onset of hypertension accompanied with proteinuria that occurs after 20th week of gestation. It is a syndrome of worldwide distribution, that complicates up to 10% of pregnancies and remain as the major cause of maternal and neonatal mortality and morbidity [1, 13]. Multiple conditions have been associated with an increased risk for PE like maternal age, nulliparity, previous preeclampsia, obesity, hypertension, among others. In this study, a logistic regression model was performed in order to estimate the main risk factors for PE and others HDP.

Key-Words: - preeclampsia, multiple correspondence analysis, logistic regression model

1 Introduction
High blood pressure is common during pregnancy; more than 10% of women will have their blood pressure recorded as above normal at some point before or within the following days after delivery. This wide range of disorders is generally grouped as HDS. In those women who develop hypertension but have no other complications, pregnancy outcome is similar, or even better, to that for women with normal blood pressure. But PE, hypertension associated with proteinuria, complicates up to 10% of all pregnancies, and can affect the mother’s and baby’s organs, leading to problems as: eclampsia, renal failure, HELLP syndrome, abruptio placentae, preterm delivery, small for gestational age, etc. These are just some reason why PE is a serious public health problem, according to the World Health Organization (WHO). It is the first cause of maternal and perinatal morbimortality in the world, particularly in the underdeveloped countries. PE has large incidence, generates the highest costs (medical and non – medical) in the obstetrical field. Antenatal care itself is not enough to prevent PE, the treatment could almost completely reverse the disease and prevent complications, as long as the diagnosis and the indication of the appropriate treatment are instituted as soon as needed. Preventive strategies could be applied among high risk women in order to reduce its incidence and complications (morbimortality). In Mérida, Venezuela, a Preclampsia Prevention Program has been applying preventive strategies to pregnant women for more than a decade. To evaluate which women are at risk and who are not is the key point of any decision to make. [1, 11, 12, 13]. Multiple conditions have been associated with an increased risk for PE like maternal age, nulliparity, previous preeclampsia, obesity, hypertension, diabetes, among Africans and descendents, familiar history of hypertension or preeclampsia, twin pregnancy, systemic erythematous lupus, poor nutrition, among others. [1, 4, 5].

The results of different clinical trials using several interventions to prevent HDP and PE have been contradictory. Interventions on women without risk not only increases medical costs but increases the risk of iatrogenic problems that could cause morbimortality in women who were going to have a satisfactory outcome. This is why estimation of the risk factors for PE, and other HDP, is a must.

We conducted a logistic regression model to evaluate these risks among pregnant women enrolled in a cohort study done by the Preeclampsia Prevention Program. 327 pregnant women (127 nulliparous and 200 multiparous) enrolled at the prenatal care clinics in Mérida, Venezuela, between January 1999 and September 2000; were prospectively followed up from their first visit (less
than 20 weeks of pregnancy) to the end of pregnancy at the University Hospital of Los Andes. Baseline parity status, as well as, their demographic characteristic and medical, obstetric, and family history were assessed at enrollment. Hypertensive disorders of pregnancy were classified according to recommendations of the National High Blood Pressure Education Program Working Group on High Blood Pressure in Pregnancy.

2 Statistical methodology developed to obtain the model for determining the main risk factors of hdp

The following statistical computational packages were used: S-Plus version 6.0 [8] and SAS version 8.02 [9] under UNIX for IBMRS6000 under AIX system, installed at the National Center for Scientific Calculations (CeCalCULA) [10]. Three statistical analyses were conducted:

1. Exploratory Data analysis: In this stage the organized description of the data was made considering both studied groups (primiparous and multiparous). Some central tendency, dispersion and form for the quantitative variables indices are presented and frequency and contingency tables for the qualitative variables; also some graphical representations were made. This stage is very important, since it reveals the variables that are candidates for being included in the later analyses.

2. Multiple Correspondence Analyses: With this multivariate technique, from the hypertable obtained with the selected variables in the exploratory analysis, we identify possible associated risk factors for PE.

3. Logistic Regression Model: When the possible associated risk factors for PE have been identified, a logistic regression model is developed, where the dependent variable is the PE (presence or absence) and the independent, or explanatory, variables are the selected risk factors according to the multiple correspondences analysis results.

The studied variables were: nulliparity, maternal age, blood pressure, hypertension, transient hypertension, history of previous PE, family history of hypertension, or PE, diabetes, obesity, twin pregnancy, to be African or decedent, low socioeconomical income, preterm labour, caesarean section, maternal death, renal insufficiency, eclampsia, HELLP syndrome, abruptio placentae.

The studied variables from the new born were: apgar test at first and the fifth minute, weight, development of asphyxia, being small for gestacional age and perinatal death.

A total of 15 (15/327) cases developed PE, 8 (8/127) among nulliparous and 7 (7/200) among multiparous.

Blood pressure is one of the most important variables when studying PE, therefore, we compared the average blood pressure in both studied groups with a “t” test. Hypothesis tests and confidence intervals were made. For these tests a significance level of $\alpha = .10$ was considered.

The relevant hypotheses are:

- **Systolic blood pressure:**
  
  \[ H_0 : \mu_{psis(primiparous)} = \mu_{psis(multiparous)} \quad \text{Vs} \]

  \[ H_1 : \mu_{psis(primiparous)} \neq \mu_{psis(multiparous)} \]

- **Diastolic blood pressure:**
  
  \[ H_0 : \mu_{pd\,(primiparous)} = \mu_{pd\,(multiparous)} \quad \text{Vs} \]

  \[ H_1 : \mu_{pd\,(primiparous)} \neq \mu_{pd\,(multiparous)} \]

Hypothesis test for averages equality (table 1), reveals that there are no significant differences between the average systolic blood pressure among primiparous and multiparous patients, but there are differences between the average diastolic blood pressure among primiparous and multiparous patients, as observed on p value and on 90% confidence interval (table 2). 90% confidence interval shows that among primiparous diastolic blood pressure is slightly superior that among multiparous patients.

<table>
<thead>
<tr>
<th>Table 1. Hypothesis Test for average blood pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>t Statistic</strong></td>
</tr>
<tr>
<td>Systolic pressure</td>
</tr>
<tr>
<td>Diastolic pressure</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2. 90% Confidence interval for mean differences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Low Limit</strong></td>
</tr>
<tr>
<td>Systolic pressure</td>
</tr>
<tr>
<td>Diastolic pressure</td>
</tr>
</tbody>
</table>
Multiple correspondences analysis shows the relations between the different modalities of all included variables, and it is wanted to find those modalities associated to PE. Incidence of PE is the dependent variable, thus in the program is considered as a supplementary variable, so, it does not affect the axis formation.

While interpreting the first three principal axes (see Graph 1) the following information can be extracted:

**Axis 1:** Is highly associated to perinatal death. Often, information concerning newborns’ weight and sizes was not registered in the medical record. Some newborns died shortly after birth and presented a low apgar test.

**Axis 2:** Is highly associated to mild or severe PE and preterm labour. Newborns from these patients presented low or very low weight at birth, low apgar test at first minute, asphyxia and were small for gestational age.

**Axis 3:** Is highly associated to younger patients (less than 15 years old) and lacking information about history of hypertension or previous PE, socioeconomic income and race.

After exploratory analysis and the multiple correspondences analysis was done, we chose the independent or explanatory variables concerning the presence or absence of PE. The chosen independent variables were:

- Diastolic Pressure (mm Hg)
- Transient hypertension (0 = absence 1 = presence)

Nulliparity is also an independent variable, since on the hypothesis test for averages equality of diastolic blood pressure showed that average diastolic blood pressure is higher among nulliparous.

All candidates for independent variables are related to features of the mother, since the usefulness of this model is to be used to apply preventive strategies.

The logistic model obtained was:

\[
\text{Logit(Pre eeeclampsia)} = -24.40302 + 0.3597\text{diaspres} + \\
+ 10.98619\text{transhyper} - 0.0333\text{nulliparous}
\]

where:

diaspres: diastolic blood pressure
transhyper: Transient hypertension
Nulliparous: nulliparous or multiparous condition.

Reference group is composed by multiparous patients who do not present transient hypertension and whose diastolic blood pressure is a given value (mm Hg).

**Coefficients Interpretation:**

- **Constant:** this value indicates the logit of probability that a multiparous patient without transient hypertension and whose diastolic pressure is a given value (mm Hg) will develop PE. When comparing the significance of this parameter with the value ±1.96 of normal standard distribution we conclude that it is not significant, that is that the probability of developing or not developing PE is the same within the reference group.

- **Diastolic Blood Pressure:** The coefficient shows an increase in the logit of probability of developing PE according to the reference group, it was the only significant coefficient (value t>1.96) and this variable also shows differences between the primiparous and multiparous patients.

- **Transient hypertension:** The value of this coefficient shows an increase of 10.98619 in the logit of probability of developing PE according to the reference group, maintaining constants the other variables. The coefficient was nonsignificant.

- **Nulliparity:** This coefficient shows a decrease of 0.0033 in the logit of probability of developing PE according to the reference group.
group. The coefficient was nonsignificant when comparing it with the value 1.96 of the normal standard distribution.

Table 3. Significance of the Logistic Model Coefficients

<table>
<thead>
<tr>
<th>Coefficients</th>
<th>Value</th>
<th>Standard Error</th>
<th>T Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>-24.40302</td>
<td>19.7662</td>
<td>-1.2345</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>0.35970</td>
<td>0.11253</td>
<td>3.19635</td>
</tr>
<tr>
<td>Transient hypertension</td>
<td>10.98619</td>
<td>17.9055</td>
<td>0.61356</td>
</tr>
<tr>
<td>Nulliparity</td>
<td>0.00333</td>
<td>0.8021284</td>
<td>0.00415</td>
</tr>
<tr>
<td>Residual Deviance</td>
<td>14.48174</td>
<td>195 degree of freedom</td>
<td></td>
</tr>
</tbody>
</table>

In order to test the model accuracy, it is necessary to compare the residual deviance (14.48174) with the Chi Square distribution value \( \chi^2_{(0.95;295)} = 336.0577 \). Considering that 14.48174 < 336.0577 we can conclude that the binomial model accuracy is appropriate.

The probabilities of developing PE due to a factor if the remaining factors are constants are:
- The probability that a patient develops PE if she is nulliparous is 0.05637 and 0.0562 if she is multiparous. In this case, a diastolic pressure of 60 mm Hg was assumed.
- The probability that a patient develops PE if she has transient hypertension is 0.9997 and 0.0563 if she did have it. In this case, a diastolic pressure of 60 mm Hg was assumed.

Since diastolic pressure is a quantitative variable, it is analyzed when an increase or decrease is observed. Thus, an increase of 10 mm Hg is a measurement that quantifies the change in the risk for PE if the remaining variables keep as constants. Therefore the risk for developing preeclampsia when an increase of 10 mm Hg in the diastolic blood pressure occurs, for instances a change from 80 to 90 mm Hg, increases 36.488 times the risk for PE in the patients included in the study, independently if she is nulliparous or multiparous.

3 Conclusion
1. The obtained logistic model turned out to be appropriate for determining the presence or not of preeclampsia in the patients included in the study. This model is useful, since it allows quantifying the probability of having preeclampsia in primiparous and multiparous and also calculates the risk of developing the disease when it is increased the blood pressure.
2. The multiple correspondences analysis is an appropriate technique for investigations in which it is desired to visualize graphically the existing relations between the modalities or categories of the studied variables.
3. Hypothesis test for averages equality in the diastolic blood pressure, reveals that there is a statistical difference between the average diastolic blood pressure among primiparous and multiparous patients. The 90% confidence interval is positive but the inferior limit is close to zero, which indicates that the diastolic pressure among primiparous patients is slightly superior to that among multiparous patients.
4. In the multiple correspondences analysis who were slight or severe preeclamptics, pathologies like high pressure or pregnancy induced hypertension were associated with childbirths before time. New born of these patients were born by means of vaginal childbirth, they have presented very low or low weight when being born, low apgar test at the first minute, asphyxia and were small for the gestacional age.
5. In the multiple correspondences analysis, the primiparity does not appear associated with preeclampsia. This finding is very interesting, because the primiparity has been the main reported risk factor associated with preeclampsia for many years.
6. The other risk factors that have been proposed, such as patient age, diabetes, obesity, black race, multiple pregnancy and low socioeconomic level do not appear associated to preeclampsia, although there exist some relations that have to be taken into account:
   • Black race patients associated to multiple pregnancy, obese, diabetic and with transitory hypertension.
   • In patients with multiple pregnancy, diabetic, obese, with transitory hypertension, that do not present low socioeconomic level, the births take place by means of Caesarean and new born present asphyxia.
   • Patient with ages between 15 and 20 years, primiparous, with hypertensive and preeclampsia antecedents, childbirths before time whose new born have presented low apgar1 and low weight or very low weight.
7. One of the most important results obtained from the logistic regression model reflects that the probability that a patient develops preeclampsia is very similar in the primiparous and multiparous patients included in the study.
8. There exists high probability that a patient develops preeclampsia if she presents pregnancy induced hypertension.

9. A 10 mm Hg increase in diastolic pressure in the patients included in the study (for example changes from 80 to 90 mm Hg) increases the risk of developing preeclampsia in 36,488 times, independently if she is primiparous or multiparous.

References:

