Liet. matem. rink., T. 42, spec. nr., 553–558

Statistical analysis of the influence of papillomavirus infection on human immune and antioxidative systems

Laura LAPKAUSKAITĖ (MII),
Janina DIDŽIAPETRIENĖ, Birutė KAZBARIENĖ (OI)

e-mail: laural@ktl.mii.lt

1. Introduction

The aim of this paper is to make a statistical analysis of the influence of Human Papillomavirus (HPV) infection on the state of immune and antioxidative systems, based on the data collected about patients of the Oncology Institute (OI) of Vinius University. So far the traditional statistical analysis (t-test, multivariate analysis of variance (MANOVA), regression and correlation analysis) was used to analyze indices of human immune and oxidative systems. Most of these methods rely on the assumption of normality of distribution which is violated rather frequently. Therefore it is worthwhile using nonparametric tests of homogeneity of the groups of interest. Logistic regression was used to describe the effect of immune and antioxidative systems on the presence and type of HPV. Logistic regression has a similar interpretation as MANOVA, however, it is less sensitive to model assumptions.

The research from the past several years has definitively shown HPV infection to play a significant role in cervical carcinogenesis. The International Agency for Cancer Research categorized HPV types 16 and 18 as human carcinogens. The role of viral and host interactions is not yet clarified. The immune and antioxidative systems respond to many tumors but the process for tumors is not understood so well as for viral infections. The behavior of tumors is altered in immunosuppressed individuals, and certain tumors have immunosuppressive activity [3]. In healthy individuals, the antioxidative system defends tissues against free radical attack. According to the present understanding, the antioxidative system plays an important positive role in carcinogenesis [4]. Therefore, disturbances of the functions of the immune and antioxidative systems reveal a deviation of the cell, organ, and the whole organism homeostasis from the normal state. The aim of this study was to analyze the changes of indices reflecting the state of the immune system or activity of the antioxidative system of cancer patients and healthy women.

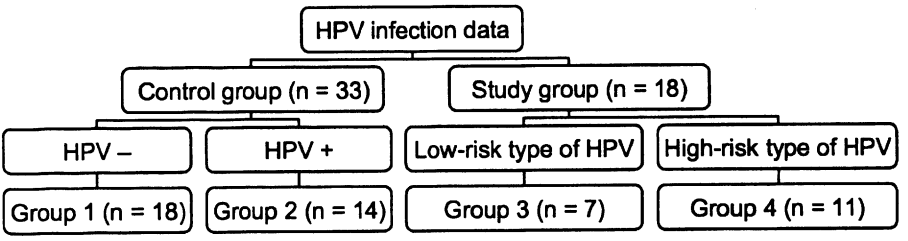

Fig. 1. Investigated groups.

2. Description of the data

The reported results were obtained when investigating 18 patients with cervical carcinoma (study group) hospitalized in the Clinic of Oncology Institute and 33 healthy volunteers (control group). The immunologic and antioxidative investigations were carried out in the Clinic of Cancer Research and Prevention of Oncology Institute of Vilnius University by researchers of this Clinic in 2001. Four groups of data were formed according to the presence of HPV infection and type of HPV in control and study groups (Fig. 1). Types 6 and 6/11 of HPV are categorized as low-risk types and 16 and 18 are categorized as high-risk types.

Thirty seven variables (general information, the means of indices reflecting the immune state as well as the means of indices reflecting antioxidative system's activity) were analyzed. General information consists of the person's age, presence of HPV and type of HPV. The immune state is specified by the total leukocyte number *(Leuk)*, immunoregulation index ($CD4^+/CD8^+$), three indices of leukocyte migration inhibition reaction *(MI25, MI50, MI75, %)*, phagocytosis index *(PhI, units)*, phagocytosis number *(PhN, %)*, concentration of immunoglobulins G, A, and M *(IgG, IgA, IgM, g/l)*, percentage and absolute number (measured in $n \cdot 10^9/l$) of lymphocyte *(Lymph)*, monocyte *(Mon)*, neutrophile *(Neu)*, eosinophile *(Eos)*, basophile *(Bas)*, total T lymphocyte population *(CD5)*, T helpers ($CD4^+$), T suppressors ($CD8^+$), B lymphocyte ($CD72^+$) and NK cells ($CD16^+$). Activity of the antioxidative system is characterized by the concentration of lipid peroxidation product malondialdehyde *(MDA, nmol/ml)*, the activity of antioxidative enzymes superoxiddismutase (SOD, $u/ml \cdot 10^4$) and catalase *(CAT, nmol/l/min)*, and the concentration of antioxidants retinol ($VitA$, $\mu mol/l$) and alpha-tocopherol ($VitE$, $\mu mol/l$).

3. Statistical analysis

3.1. *Analysis of variance*

A model for the analysis of variance has been applied in the investigation of difference between means (medians) of the groups. A hypothesis on distribution's normality for the majority of variables was rejected (because of outlying observations, multimodality and skewness of distributions). Therefore nonparametric tests for homogeneity of the groups

in location (Wilcoxon, Kruskal–Wallis, median, Van der Waerden) and scale (Ansari–Bradley) were performed with respect to indices of immune and antioxidative systems.

A linear rank statistic has the following simple form $S = \sum_{j=1}^{n} c_j a(R_j)$, where R_j is a rank of the jth observation, $a(x)$ is a rank score function, c_j is an indicator variable denoting the class to which the jth observation belongs, and n is the total number of observations. For example, the Wilcoxon score function is $a(x) = x$, the median score function is $a(x) = \mathbf{1}\{x > (n+1)/2\}$, and the Van der Waerden score function is $a(x) = \Phi^{-1}(x/(n+1))$. The Van der Waerden score function is optimal for normal distribution. A comprehensive description of nonparametric methods one can find in [2].

For each score function mentioned above, asymptotic and exact tests for the null hypothesis of no difference among groups were performed. Since the logarithmic transformation enables us to reduce positive skewness and influence of large positive outliers, a t-test was performed for logarithms of variables as well. For comparison, the results of t-tests are also presented. The variables that significantly differ among groups are tabulated in Table 1.

A significant difference among groups 1 and 2 has been established for a unique variable VitE, the concentration of alpha-tocopherol (R T L). Also, a significant difference among groups 3 and 4 has been obtained for a unique variable, namely, leukocyte migration inhibition reaction MI25 (T L).

The table demonstrates that all the tests for location lead to the same conclusions in the majority of cases. The tests based on Ansari–Bradley scores for scale differences revealed no differences among the groups for these variables.

Table 1

Statistical significance is established by rank tests (R), t-test (T), t-test on the logarithmic scale (L)

Variable	Group 1–3	Group 1–4	Group 2–3	Group 2–4
Age	R T L	R T L	R T L	R T L
Leuk	R			
Lymph	R T L	R T L		
Neu	T L	T L		
NeuAb	R T L			
CD5		R T L	T	R T L
CD5Ab				L
CD16		R T L		
CD72	R T L	R T L	R T L	R T L
CD72Ab	L	R T L		R T L
IgG		R T L		
MI25	R T L	T L	R T L	R T L
PhN	T L		L	
Phl	R T L	R T L	R T L	R T L
MDA		R L		R T L

3.2. *Logistic regression*

Model. Instead of the traditional multivariate analysis of variance (MANOVA), a logistic regression can be used. Both of these methods determine indices that reflect differences between the groups, however, they differ in the underlying model. The logistic regression is reverse to MANOVA in some sense. The model for the analysis of variance describes how qualitative differences among groups reveal themselves through quantitative indices, while the model of logistic regression, on the contrary, describes how quantitative characteristics (explanatory variables) of individuals enable one to predict a group the person belongs to.

For binary response models, the response Y can assume one of the two possible values $Y = 0$ and $Y = 1$. Suppose $x \in R^m$ is a vector of explanatory variables. In the logistic regression model, it is assumed that the conditional probability of the event $Y = 1$ under the condition that explanatory variable is equal to x is of the form

$$\Pr\left(\mathrm{Y} = 1|x\right) = \frac{\exp(\beta^T x)}{1 + \exp(\beta^T x)} \quad \text{(logistic function)}.$$

For estimating the parameter $\beta \in R^m$, either the maximum likelihood or the weighted least squares method is used.

Goodness-of-fit characteristics of the logistic model are the following: the p-value of Hosmer–Lemeshow goodness-of-fit statistic and rank correlation c.

The Hosmer–Lemeshow statistic is

$$\chi^2_{HL} = \sum_{i=1}^{g} \frac{(O_i - N_i\bar{\pi}_i)^2}{N_i\bar{\pi}_i(1 - \bar{\pi}_i)},$$

where g is the number of groups, N_i is the total frequency of subjects in the ith group, O_i is the total frequency of the event outcomes in the ith group and $\bar{\pi}_i$ is the average estimated probability of the event outcome for the ith group. Usually $g = 10$ and groups of approximately equal size are obtained by dividing all the observations sorted in increasing order of their estimated event probability. The Hosmer–Lemeshow statistic is then compared to a chi-square distribution with $(g - n)$ degrees of freedom (the default is $n = 2$).

The rank correlation $c = 0.5(D+1) = (n_c + 0.5(t - n_c - n_d))/t$, where D is Somer's rank correlation coefficient, t is the total number of pairs with different responses, n_c of which are concordant, n_d of which are discordant, and $t - n_c - n_d$ of which are tied. Recall that two pairs $(Y_i, \widehat{p}(x_i))$, $(Y_j, \widehat{p}(x_j))$ are called concordant if $z = (\widehat{p}(x_j) - \widehat{p}(x_i))(Y_j - Y_i) > 0$, discordant if $z < 0$, and tied if $z = 0$.

Results. Two models of the logistic regression were investigated.

For women in the control group not infected with HPV (group 1, in this case, the response variable $Y = 0$) and infected with a low-risk type of HPV (group 2, response $Y = 1$), a significant explanatory variable is the concentration of antioxidant alpha-tocopherol VitE ($p = 0.0147$, $\widehat{\beta} = 0.2088$). $HL\ p = 0.3588, c = 0.797$. Here and in

what follows p stands for the p-value, $\widehat{\beta}$ for the maximum likelihood estimate of β, and $HL\ p$ for the Hosmer–Lemeshow p-value.

For women in the study group infected with a low-risk (group 3, response $Y = 0$) and high-risk (group 4, response $Y = 1$) type of HPV, significant explanatory variables are leukocyte migration inhibition reaction MI25 ($p = 0.0766, \widehat{\beta} = 0.1627$) and percentage of natural killers CD16 in lymphocytes ($p = 0.0547, \widehat{\beta} = 0.7562$). $HL\ p = 0.5348$, $c = 0.843$. MANOVA revealed significant differences among the groups with respect to the pair of variables MI25 and CD16 (the hypothesis of no-overall group effect was rejected at the significance level 0.05, $p = 0.0260$).

4. Conclusions

32 indices reflecting the state of the immune system, 5 indices reflecting activity of the antioxidative system and information about the presence and type of HPV, and age of a person have been studied. Some studies have shown that the functional state of the immune system is closely associated with the development of cervical dysplasia and cancer in women with HPV infection [5], [6]. Some investigations indicate that the impact of HPV infection is greater on immunosuppressed individuals [7]. We have found that the immune system of patients reacts to HPV infection. In both study groups the percentage index of B lymphocyte was essentially lower and the phagocytosis index was essentially higher than in both control groups. The logistic regression made it possible to determine a significant difference among HPV-negative and HPV-positive healthy women groups with respect to the pair of variables MI25 and CD16 (these indices were not found to be significant by univariate methods). MANOVA leads to the same result.

Evaluation of the antioxidant state has revealed that the concentration of lipid peroxidation product MDA is similar in both control groups and is higher in patients with cervical carcinoma infected by a high-risk type of HPV. The changes in the concentration of MDA imply that an alteration in pro-oxidant/antioxidant balance in favor of the pro-oxidant site in cancer patients is associated with tumors progression [8]. Nutritional factors are associated with the risk of cervical cancer. A statistically significantly lower level of retinol [9] or alpha-tocopherol [10] was observed in the blood serum of HPV-positive patients with cervical intraepithelial neoplasia. In our study, the level of alpha-tocopherol was higher in the blood serum of HPV-positive healthy women (logistic regression confirmed this result)and was similar in both study groups. Multivariate methods have not revealed any new effects. The level of retinol was similar in all groups. Our preliminary results indicate that the changes in the activity of the antioxidative system is more expressed in the women with cervical carcinoma.

Acknowledgement. The authors would like to thank Ž. Gudlevičienė, A. Krištaponienė, G. Prasmickienė, D. Šukelienė, and R. Žemaitytė for the data presented.

References

[1] A. Agresti, *Analysis of Ordinal Categorical Data*, John Wiley & Sons, Inc., New York (1984).
[2] M. Hollander, D.A. Wolfe, *Nonparametric Statistical Methods*, John Wiley & Sons, Inc., New York (1973).
[3] G. Touloumi, A. Hatzakis, I. Potouridou, I. Milona, J. Strarigos, A. Katsambas, G. Giraldo, E. Beth–Giraldo, R.J. Biggar, N. Mueller, D. Trichopoulos, The role of immunosuppression and immune-activation in classic Kaposi's sarcoma, *Int J. Cancer*, **82**(6), 817–821 (1999).
[4] D. Dreher, A.F. Junod, Role of oxygen free radicals in cancer development, *Eur. J. Cancer*, **32A**(1), 30–38 (1996).
[5] I.H. Frazer, The role of immune system in anogenital human papillomavirus, *Australs J. Dermatol.*, **39**, Suppl 1, 5–7 (1998).
[6] J. Konya, J. Diller, Immunity to oncogenic human papillomavirus, *Adv. Cancer*, **82**, 205–238 (2001).
[7] P.K. Nicholls, M.A. Staviley, The immunology of animal papillomavirus, *Vet. Immunol Immunopathol*, **73**(2), 101–127 (2002).
[8] M. Gerber, C. Astre, C. Segala, M. Saintot, J. Simony–Lafontaine, H. Pujol, Oxidant-antioxidant status alterations in cancer patients: relationship to tumor progression, *J. Nutr.*, **126**, Suppl 4, 1201S–1207S (1996).
[9] M. Lehtinen, T. Luostarinen, L.D. Youngman, T. Anttila, J. Dillner, T. Hakulinen, P. Koskela, P. Lenner, G. Hallmans, Low levels of serum vitamins A and E in blood and subsequent risk for cervical cancer: interaction with HPV seropositivity, *Nutr. Cancer*, **34**(2), 229–234 (1999).
[10] A. Kwasniewska, A. Tukendorf, M. Semczuk, Content of alpha-tocopherol in blood serum of human Papillomavirus – infected women with cervical dysplasias, *Nutr. Cancer*, **28**(3), 248–251 (1997).

Papilomos viruso įtakos žmogaus imuninei ir antioksidacinei sistemoms statistinė analizė

L. Lapkauskaitė, J. Didžiapetrienė, B. Kazbarienė

Šis taikomasis darbas skirtas papilomos viruso įtakos žmogaus imuninei ir antioksidacinei sistemoms statistinei analizei. Statistiškai reikšmingų rodiklių nustatymui buvo taikomi ranginiai kriterijai ir logistinė regresija.