

Prediction Using Logistic Regression Analysis of Peripheral Vascular Disease

Li Yanan^{1, 2}, Guo Xiaona³, Yan Chunsheng^{3, *}

¹ Henan Province Hospital of TCM, 450000, Henan, P. R. China

² The Second Affiliated Hospital of Henan University of Traditional Chinese Medicine, 450000, Henan, P. R. China

³ Huanghe Science and Technology College, Zhengzhou, 450003, Henan, P. R. China

*email: 331318321@qq.com

Abstract. Logistic regression model is to study the response variable is an important analytical method for non-continuous variables. Linear regression models and quantitative analysis is one of the most commonly used data mining methods of statistical analysis, linear regression analysis but generally require the response is a continuous variable, the data distribution is normal conditions. This study used logistic regression analysis to predict the study of peripheral vascular disease in the carotid atherosclerosis disease prediction model was established to provide scientific basis for the clinical treatment of peripheral vascular disease.

Keywords: Logistic; peripheral vascular disease; regression model; clinical evidence.

1 Introduction

Etiology of peripheral vascular disease can be divided into two categories with the external factors. External factors include Six Exogenous, special toxin (Opium), trauma; internal factors include diet, emotional injuries, organs and meridians dysfunction, injured workers wasting like. In the disease process, whether it is caused by internal or external trigger, can appear blood stasis. After blood stasis, damage the body's normal blood circulation causing various pathological changes. Its pathogenesis is key evil, Deficiency and Stasis. Which can be either external evil, justice may be the pathological product (such as bleeding, phlegm, wet) after blood stasis; not only by the evil imaginary bar even, it may be the result of injury is blood stasis; stasis is often the result of evil caused, but also because of some imaginary made^[1]. Although the lesions of peripheral vascular disease in the blood vessels of a local majority, but there is a close relationship with the blood and viscera. Because the loss of function of organs deep, unable to transport blood, unifying right, catharsis disorders, the blood does not run properly and stasis occurs; on the contrary, after blood stasis also causes the organs lose Ruyang wasting. In addition, peripheral vascular disease etiology and pathogenesis are still endowments intolerance, genetic factors, Chong and Ren disorders, temporary card also can not be ignored.

Peripheral vascular disease in clinical common symptoms are pain, abnormal skin temperature, skin color abnormalities, limb atrophy (dystrophy syndrome), limb thickening (swelling), ulcers and gangrene.

(1) Pain is the main arterial ischemic disease symptoms, it can be divided into intermittent pain, persistent pain (rest pain, nerve pain neutral).

① Intermittent pain: chronic limb insufficiency, arterial insufficiency decompensated, common in the early gangrene. The main reason there is physical activity, body position, temperature changes.

② Persistent pain (rest pain): pain was continuous member, particularly at night staggering. Because when a stationary state heart rate and blood pressure were lower, distal artery worse, cause severe ischemic extremities persistent severe pain, patients often can not sleep due to pain, tuck sit. When the body sagging, it may slightly reduce pain, swelling of the extremities can cause long drooping.

③ Nerve neutral pain: acute or chronic arterial occlusion, limb ischemia, can cause ischemic neuritis showed persistent dull with intermittent smarting, may radiate to the distal limb, numbness wells, hair.

cool, cold, or burning, ant line, acupuncture and other sensory abnormalities. Found anthrax found in fine or diabetes peripheral neuropathy and other patients.

(2) Skin temperature anomalies

Cold limbs, cold, caused by arterial occlusion stenosis; fever, burning, usually caused by excessive expansion of arteries by convergence, seen in Erythromelalgia. Limb venous stasis, it can lower limb stuffy feel, seen in varicose veins and other diseases.

(3) Abnormal skin color

Pale skin color, flushing, bruising to clogged arteries, insufficiency of performance; skin pigmentation, compared with venous stasis disease; skin redness arterial dilation; if the skin was a pale purple three-color change was a flushing artery spasm cause, such as Raynaud's syndrome.

(4) Limb atrophy (dystrophy syndrome)

Mostly caused by ischemia, visible dry skin, thin light, scaling, chapped, decreased sweating; hair thinning, loss, stop growing; toenails grow slowly, thick dry, brittle, dull, ingrown toenails like growth; muscle atrophy, especially in the feet and legs, severe atrophy is prompted severe ischemia.

(5) Limb thickening (swelling)

Or more for venous reflux disease and other disorders. Crude swollen limbs, according to the depression, pain, venous thrombosis; slightly swollen ankle and foot, after standing is even worse, raising subsided, reflux or venous valvular insufficiency, caused by venous stasis^[2]. Late Treads visible area of skin and subcutaneous tissue nutritional disorders, stasis dermatitis, pigmentation. Lymphedema, visible signs of cellulite, a wide range of hard and tough skin thickening, even as elephantiasis legs. Other systemic edema, mostly for heart, liver, kidney and other organ diseases.

(6) Ulcers and gangrene

If arterial ischemic disease, severe tissue ischemia, hypoxia, is the formation of ulcers and gangrene. Often starting at the toe, after the extension and each toe. That spread to the dorsum of the foot, ankle. Mostly dry gangrene, infection were mostly wet. It should be noted ulcer boundaries, location, extent, depth, granulation, pus; if venous stasis ulcers, occurs in the foot boots, an area may be larger, edge thickening and hardening, black, not of granulation fresh, foul smell, prolonged unhealed, recurrent.

2 Logistic Regression

Logistic model was proposed in 1838 by the Belgian biologist Verhulst, was first to study population issues and make improvements to the Malthus equation, at first he did not pay attention, and later American demographers Pear in the study of the US population to raise this issue again equation before it became popular, it is often called modern literature Verhulst-pearl hinder equation^[3]. This model is called Logistic model, because it has some meaning reasoning.

General Logistic model form as follows:

$$P(Y=1|x_1, x_2, \dots, x_n) = \frac{\exp(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)}{1 + \exp(\beta_0 + \beta_1 x_1 + \dots + \beta_n x_n)} \quad (1)$$

Do logit transformation of the above formula, Logistic regression model can be written as:

$$\text{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_n x_n \quad (2)$$

Because Logistic regression model involves more complex mathematical theory, the statistical analysis of data tend to be more complex and usually can not be calculated by hand, can only be achieved on the computer. Statistical software, commonly used with SAS, SPSS, S-PLUS, but these are mostly commercial software, you need to pay expensive purchase price^[4]. The R statistical software is a free software, it provides a flexible, interactive environment to analyze, display data, and provides a number of statistical packages as well as some statistical calculation tools and integrated function enables users to perform flexible data analysis simplifies data analysis. It can complete statistical calculations for most models, and helps the user to analyze the results.

3 Fitting Logistic Model

In the actual analysis, the non-continuous response variables commonly encountered, namely a categorical response variables, such as economic research involved whether to buy a product, whether under certain conditions the prevalence epidemiology studies, etc. In the study the relationship between dichotomous variables with a number of independent variables, it is usually chosen Logistic regression model. Its multiple regression model was mainly seeking an estimate parameters of the model parameters Logistic model commonly used estimate the maximum likelihood method (maximum likelihood, ML). ML basic idea of natural law to establish the likelihood function and the log-likelihood function, then by making the maximum log-likelihood function to solve the corresponding parameter values, the estimated value of the obtained parameter called Maximum Likelihood estimated value^[5]. Maximum Likelihood Estimation good statistical properties of consistency, efficiency and normality, etc., if the sample data, the more accurate its estimated value.

4 Case Analysis

In the prediction of carotid atherosclerosis as an example, Logistic regression analysis to predict the static atherosclerosis. 247 patients, data analysis, male 116 cases, 131 cases of female, age range 29-84 years, the overall mean age was 52.6 (± 9.4 years, 25.1% suffering from carotid atherosclerosis, diabetes 13.8%, 13.4% suffer from hypertension. Basic carotid atherosclerosis carotid atherosclerosis group and non-group are shown in Table 1.

Table 1. Basic clinical information of all the subjects

	Non-carotid arteriosclerosis group	Carotid arteriosclerosis group	p
Number	185	62	
Age	50.2 \pm 8.4	59.5 \pm 8.9	<0.001
Sex	42.7%	59.7%	0.020
BMI	24.15 \pm 2.95	24.41 \pm 3.66	0.638
WC	81.2 \pm 8.5	84.0 \pm 11.1	0.103
PP	48.2 \pm 14.2	57.1 \pm 19.1	0.001
Smoke	25%	%%	0.042
Hypertension	11.4%	94	0.109
DM	8.1%	40	<0.001

According to the research object Log model value calculated according to carotid atherosclerosis carotid atherosclerosis group and non-group grouped as a scatter plot. as shown in picture 2.

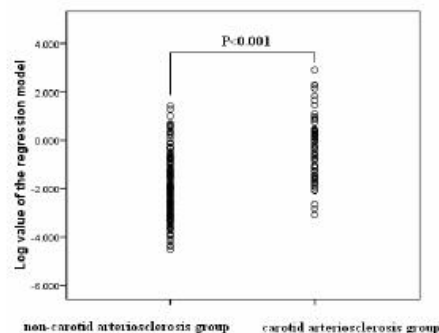


Figure 1. Scatter plot of Log value for the regression model

Make further ROC curve to evaluate the model's predictive value of carotid atherosclerosis, shown in Figure 2.

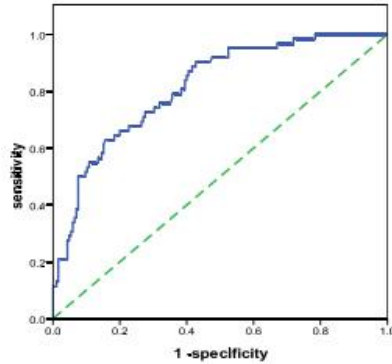


Figure 2. ROC curve of the regression model in detecting carotid arteriosclerosis

As can be seen from the figure, the area under the ROC curve (area under the curve, AUC) was 0.814, $P < 0.001$, when $PI = 0.157$, the maximum Youden index, the sensitivity of 0.903 and specificity of 0.573.

References

- [1] Gong Lansheng, Xu Yun Yu, Weizhong. China vascular disease early detection Technical Application Guide (the first report) [J]. Chinese People's Health, 2006,18 (9): 323-331.
- [2] Shizhao Jian, Zhang Mingming. Logistic regression analysis [J]. Computer Aided Engineering, 2005 (3).
- [3] Jin Shuigao. The correct interpretation and the correct application of the results of Logistic regression method [J]. Journal of Preventive Medicine, 2003 (3).
- [4] Antoniadis C, Antonopoulos AS, Tousoulis D, et al. Homocysteine and coronary atherosclerosis: from folate fortification to the recent clinical trials [J]. Eur Heart J, 2009,30(1):6-15.
- [5] Zheng Wei, Gao Ge, Wei Yibi. Improved Principal Component Logistic Regression Research and Application [J]. Suzhou University Journal of Medical Science, 2008, 28(4): 517-520.