Binary Logit Regression Analysis

Summary of binary Logit regression analysis

name	option	frequency	percentage
divide into groups	0	134	57.26496%
	1	100	42.73504%
	amount to	234	100.0%
gather	valid	234	100.00000%
	hiatus	0	0.00000%
	amount to	234	100.0%

The intensity difference(ID), intensity ratio(IR),absolute enhancement intensity value (EIAV), lesion diameter, AFP≥7ng/mL,1, and <7ng/mL,0 were taken as independent variables, while the group was taken as the dependent variable for binary Logit regression analysis. As can be seen from the above table, a total of 234 samples participated in the analysis, and there was no missing data.

Results of likelihood ratio test of binary Logit regression model

model	-2 times the logarithmic likelihood value	Chi-square value	*df*	*p*	AIC price	BIC price
Just intercepts	319.43518					
Final model	210.27660	109.15858	5	0.00000	222.27660	243.00853

First, we conducted a comprehensive evaluation of the model's overall validity. As shown in the table above: The original null hypothesis tested in this analysis was that the model quality remained consistent when either the ID, IR, EIAV, lesion diameter, or AFP≥7,1, <7,0 variables were included. The p-value was less than 0.05, which indicates rejection of the null hypothesis. This result confirms that the selected independent variables demonstrated statistical significance, validating the model's validity and confirming the model construction's practical relevance.

Summary of binary Logit regression analysis results

Item	regression coefficient	standard error	z price	Wald χ^2	p price	OR price	OR 95% CI
ID	0.04699	0.09924	0.47356	0.22426	0.63581	1.04812	0.86286 ~ 1.27315
IR	11.61627	3.47760	3.34031	11.15767	0.00084	110887.30849	121.54611 ~ 101163216.94603
EIAV	-0.05658	0.04250	-1.33125	1.77223	0.18311	0.94499	0.86947 ~ 1.02708
Diameter of lesion	0.05401	0.06379	0.84664	0.71681	0.39719	1.05549	0.93145 ~ 1.19606
AFP≥7，1、<7，0	1.02916	0.37797	2.72289	7.41411	0.00647	2.79872	1.33424 ~ 5.87064
nodal increment	-12.64119	4.53374	-2.78825	7.77432	0.00530	0.00000	0.00000 ~ 0.02340

Note: The dependent variable = group

As shown in the table above, when conducting binary Logit regression analysis with ID, IR, EIAV, lesion diameter, AFP≥7,1, and <7,0 as independent variables, and grouping as the dependent variable, the results indicate that these factors collectively explain 0.34 percentage point variation in group distribution. The model equation derived from the data is: ln(p/1-p) = -12.641 + 0.047* ID+ 11.616* IR-0.057* EIAV+ 0.054* lesion diameter + 1.029* ≥7,1, <7,0 (where p represents the probability of group 1, and 1-p denotes the probability of group 0).

According to the summary and analysis, it is known that IR, AFP≥7,1, and <7,0 will have a significant positive influence on the grouping. However, the ID, EIAV, and the diameter of lesions will not have an influence on the grouping.

Summary of binary Logit regression prediction accuracy

		predicted value		forecast accuracy	Prediction error rate
		0	1		
ground truth	0	114	20	85.07463%	14.92537%
	1	24	76	76.00000%	24.00000%
gather				81.19658%	18.80342%

The model fitting quality was determined by the prediction accuracy of the model. As shown in the table above, the overall prediction accuracy of the research model was 81.20%, indicating acceptable model fitting. When the true value was 0, the prediction accuracy was 85.07%; when the true value was 1, the prediction accuracy was 76.00%

Hosmer-Lemeshow fit test

χ^2	*free degree df*	p price
8.54968	8	0.38169

Hosmer-Lemeshow fit test is used to analyze the model fit. According to the table above,

The original hypothesis tested by the model here is that the degree of conformity between the model fitting value and the observed value is consistent; the p-value here is greater than 0.05 (χ^2 = 8.550, p = 0.382> 0.05), so it indicates that the original hypothesis is accepted, that is, the model passes the HL test and the model fitting adequacy is good.

Hosmer and Lemeshow test process table

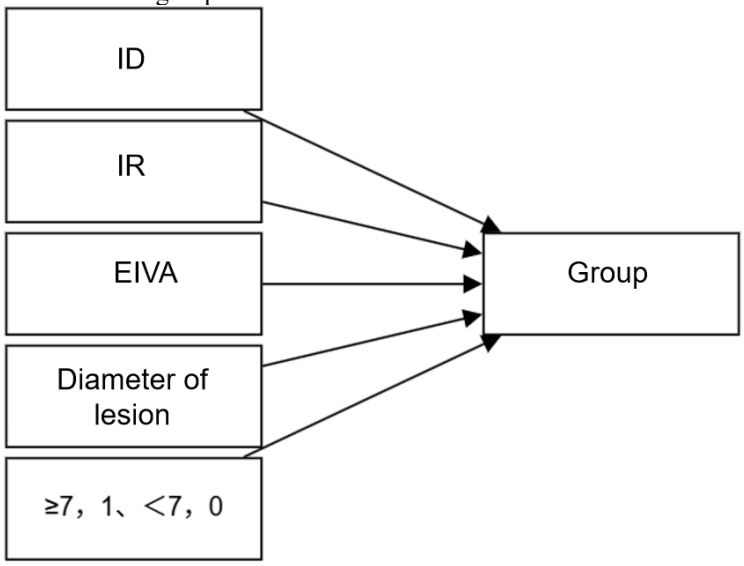
Predict the probability of the percentile group	Group =0		Group =1		gather
	observed value	expected value	observed value	expected value	
[0,10]	22	22.05919	1	0.94081	23
(10,20]	23	22.15888	1	1.84112	24
(20,30]	21	19.97131	2	3.02869	23
(30,40]	21	19.51773	3	4.48227	24
(40,50]	15	15.68749	8	7.31251	23
(50,60]	14	12.99910	10	11.00090	24
(60,70]	6	9.94239	17	13.05761	23
(70,80]	6	7.15842	18	16.84158	24
(80,90]	3	3.45229	20	19.54771	23
(90,100]	3	1.05322	20	21.94678	23

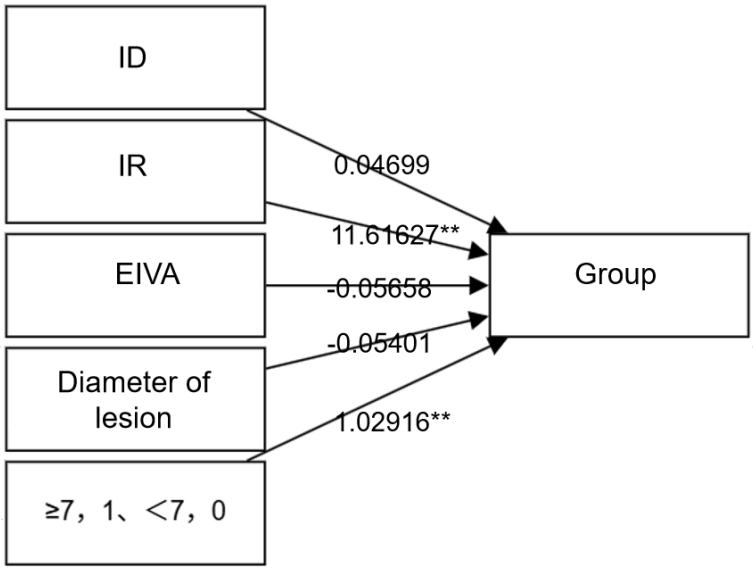
Results of binary Logit regression analysis-simplified format

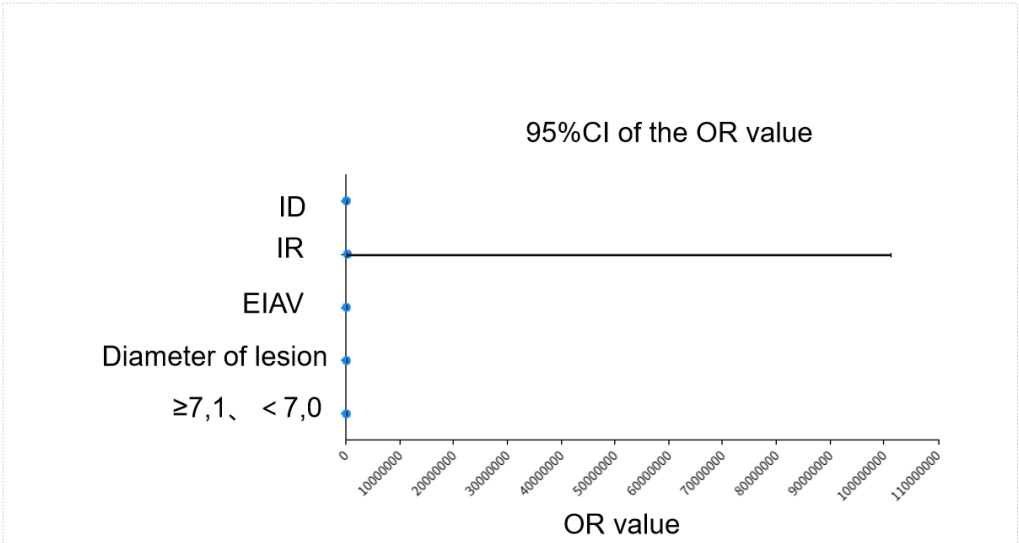
Item	regression coefficient
ID	0.04699
	(0.47356)

IR	11.61627** (3.34031)
EIAV	-0.05658 (-1.33125)
Diameter of lesion	0.05401 (0.84664)
≥7，1、＜7，0	1.02916** (2.72289)
nodal increment	-12.64119** (-2.78825)
likelihood ratio test	χ2 (5)=109.15858,p=0.00000
Hosmer-Lemeshow test	χ2 (8)=8.54968,p=0.38169

Note: The dependent variable = group







Marginal effects result

Item	Marginal effect dy/dx	standard error	z price	p price	95%CI
intensity difference	0.00686	0.01445	0.47425	0.63532	-0.02148 ~ 0.03519
strength ratio IR	1.69451	0.46340	3.65668	0.00026	0.78626 ~ 2.60276
Absolute value of enhanced intensity of lesion	-0.00825	0.00612	-1.34812	0.17762	-0.02025 ~ 0.00375
Diameter of lesion	0.00788	0.00925	0.85194	0.39424	-0.01025 ~ 0.02600
≥7，1、＜7，0	0.15707	0.05734	2.73941	0.00615	0.04469 ~ 0.26945

Summary of sample loss

Item	sample number	proportion
Valid samples	234	100.00000%
Eliminate invalid samples	0	0.00000%
amount to	234	100%