

Construction and validation of a risk nomogram model for colorectal sessile serrated lesions

Journal of International Medical Research

2025, Vol. 53(5) 1–16

© The Author(s) 2025

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/03000605251337577

journals.sagepub.com/home/imr



Sihui Huang* , **Shiyu Liu***, **Fang Tan**, **Hu Chen**
and **Guangxia Chen** 

Abstract

Objective: This study aimed to explore the risk factors for colorectal sessile serrated lesions and construct a risk nomogram model.

Methods: Patients were enrolled retrospectively from the Affiliated Xuzhou Municipal Hospital of Xuzhou Medical University from January 2019 to September 2023 and randomized to the training and validation sets at a ratio of 7:3. The predictors for constructing the nomogram model were screened via univariate analysis and multivariate logistic regression analysis. Subsequently, the performance of the model was evaluated.

Results: Multivariate logistic regression analysis revealed that age, history of smoking, history of alcohol consumption, and triglyceride–glucose index were independent risk factors for colorectal sessile serrated lesions ($p < 0.05$). The area under the curve values of the nomogram model in the training and validation sets were 0.715 (95% confidence interval: 0.676–0.753) and 0.742 (95% confidence interval: 0.669–0.815), respectively. The calibration curves showed good homogeneity between the predicted and actual values. Decision curve analysis showed that this nomogram model can achieve positive clinical benefits.

Conclusions: Age, history of smoking, history of alcohol consumption, and triglyceride–glucose index are independent predictors of colorectal sessile serrated lesions. This nomogram model may predict the risk of colorectal sessile serrated lesions.

*These authors contributed equally to this work.

Corresponding author:

Guangxia Chen, Department of Gastroenterology, The Affiliated Xuzhou Municipal Hospital of Xuzhou Medical University. No. 269, University Road, Tongshan District, Xuzhou, Jiangsu, China.

Email: gx_chen2010@163.com

The Affiliated Xuzhou Municipal Hospital of Xuzhou Medical University, China



Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (<https://creativecommons.org/licenses/by-nc/4.0/>) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Keywords

Colorectal cancer, sessile serrated lesions, prediction model, nomogram, prevention

Date received: 29 December 2024; accepted: 9 April 2025

Introduction

Colorectal cancer (CRC) is a malignant tumor originating from the colorectal mucosal epithelium. According to the Global Cancer Statistics 2022, CRC ranked third in the number of new cases and second in the total number of deaths worldwide. In Asia, the proportion of malignant tumors with low survival rates, such as digestive tract tumors, is higher than that in other parts of the world. In China, approximately 517,000 new cases of CRC and 240,000 associated deaths were reported in 2022, ranking second and fourth among all malignant tumors, respectively. Therefore, CRC currently poses a substantial disease burden.¹ In addition to the traditional adenomatous polypcarcinoma pathway, the serrated pathway is an important carcinogenesis pathway. It has been reported that approximately 20%–30% of sporadic CRC cases occur through the serrated pathway.^{2–4} Serrated polyps refer to the serrated structures in crypts observed under a microscope. According to the World Health Organization guidelines (2019), serrated polyps can be classified into three categories: hyperplastic polyps, sessile serrated lesions (SSLs), and traditional serrated adenomas.⁵

Colorectal SSLs account for approximately 5%–15% of the serrated polyps. These lesions are flat and irregular with unclear boundaries under endoscopy. The surface of the lesion is often covered with a mucus cap or exhibits a cloud-like appearance. Approximately 75%–90% of colorectal SSLs are located in the proximal colon, making it difficult to detect via colonoscopy, with a higher rate of missed diagnosis

and incomplete resection under endoscopy.⁶ Therefore, colorectal SSLs are considered the main precancerous lesion of the serrated carcinogenesis pathway and one of the important lesions in the occurrence of inter-phase CRC. Screening of high-risk groups for SSLs may reduce the incidence of CRC, which could mitigate the adverse effects of colorectal malignancies. However, the current research data on the clinical risk factors for colorectal SSLs remain limited. This study aimed to explore the risk factors for colorectal SSLs by analyzing the clinical data of patients with colorectal SSLs and healthy controls as well as construct and verify a nomogram model to predict the risk of colorectal SSLs.

Materials and methods

General information

Sample size estimation was performed using the R software “4-step method” loaded into the ‘pmsampsize’ package. At least 759 cases were estimated for inclusion in the sample. This study included 310 patients who were diagnosed with colorectal SSLs by a pathologist between January 2019 and September 2023. The inclusion criteria were as follows: (a) aged ≥ 18 years; (b) good bowel preparation (Boston Bowel Preparation Scale ≥ 6 points) for the completion of colonoscopy and pathological biopsy; and (c) complete clinical data. The exclusion criteria were as follows: (a) incomplete colonoscopy or failure to reach the ileocecal region; (b) poor bowel preparation that affected observation; (c) patients with CRC or other intestinal

diseases, such as inflammatory bowel disease, colorectal ulcer, and familial adenomatous polyp syndrome, who underwent colonoscopy; (d) history of colorectal surgery; (e) history of malignant tumors; and (f) combined with infection or serious diseases of the heart, brain, liver, kidneys, and other organs or other diseases that affected the study outcome. A total of 626 healthy adults who underwent routine endoscopies were selected from the endoscopy center during the same period. Colonoscopy revealed no obvious abnormalities.

Methods

The following clinical data of the participants were collected retrospectively: (a) general clinical data: sex, age, height, weight, body mass index (BMI), history of chronic diseases (hypertension, diabetes), and lifestyle habits (smoking, alcohol consumption); (b) laboratory test results: white blood cell count (WBC), red blood cell count (RBC), lymphocyte count (LYM), neutrophil count (NEU), eosinophil count (EOS), basophil count (BAS), monocyte count (MON), hemoglobin (Hb) level, blood glucose level, total cholesterol (TC) level, triglyceride (TG) level, low-density lipoprotein (LDL) level, high-density lipoprotein (HDL) level, TC/HDL, TG/HDL, and LDL/HDL. The neutrophil/lymphocyte ratio (NLR) and triglyceride–glucose (TyG) index were calculated. The reporting of this study conforms to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines.⁷ This study was conducted following the Helsinki Declaration of 1975 as revised in 2013. This study was approved by the Affiliated Xuzhou Municipal Hospital of Xuzhou Medical University Research Ethics Committee (approval no. 2023-KY-059) in August 2023. Written informed consent was obtained from the participants before starting colonoscopy. Patient data were not

shared with third parties. All patient details have been de-identified and may not be identified in any way.

BMI refers to the index obtained by dividing the body weight (kg) by the square of body height (m^2). History of alcohol consumption is defined as still regular drinking, with an average of ≥ 1 time per week for >6 months. History of smoking is defined as continuous or cumulative smoking for >6 months. According to the Chinese Guidelines for the Prevention and Treatment of Hypertension (2018 Revision), history of hypertension is defined as a systolic blood pressure level of ≥ 140 mmHg and/or diastolic blood pressure level of ≥ 90 mmHg, measured three times in the clinic on different days in patients who did not take antihypertensive drugs in the past or were currently receiving antihypertensive medication. History of diabetes is defined as follows: (a) random intravenous blood glucose level of ≥ 11.1 mmol/L; (b) fasting blood glucose level of ≥ 7.0 mmol/L; (c) postprandial blood glucose level of ≥ 11.1 mmol/L or blood glucose level of ≥ 11.1 mmol/L at 2 h after oral glucose tolerance test of 75 g anhydrous glucose; (d) glycosylated Hb level of $\geq 6.5\%$; and (e) hyperglycemic patients with a history of diabetes who received medical treatment. NLR refers to the ratio of neutrophils to lymphocytes, which suggests a relationship between the systemic inflammatory response state and the progression of colorectal tumors. Neutrophils and lymphocytes are two types of inflammatory cells in the body. Inflammatory reactions lead to abnormal levels of blood inflammatory indicators, such as neutrophils and lymphocytes. The increased levels of neutrophils produce various cytokines and chemokines to promote tumor progression. Lymphocytes are involved in the body's immune response. Inflammation leads to decreased lymphocyte count and weakens the body's antitumor ability, which results in the proliferation of tumor cells. There are many studies on NLR in

neoplastic diseases. Some studies have indicated that NLR has a certain diagnostic value for colorectal tumors.^{8–10} The TyG index is calculated using the following formula: $\text{TyG index} = \ln [\text{TG (mmol/L)} \times 88.545 \times \text{fasting blood glucose (mmol/L)} \times 18/2]$. The TyG index is widely used to assess insulin resistance and is closely related to an increased risk of cardiovascular disease.¹¹ Moreover, studies have indicated that TyG index is associated with an increased risk of colorectal tumors.^{12–14}

Statistical methods

Data processing and statistical analysis were performed using R4.3.1 (R Software Programs (<https://cran.r-project.org/web/packages/>)) and SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, N.Y., USA). Normally distributed quantitative data were expressed as mean \pm SD and analyzed via independent sample *t*-test. Non-normally distributed quantitative data were expressed as median (quartile) and analyzed via the Mann–Whitney U test. Qualitative data were expressed as percentages and analyzed via Pearson's chi-square test or Fisher's exact probability test. In univariate analysis, statistically significant variables were further analyzed via multivariate logistic regression to explore the independent risk factors for colorectal SSLs. A nomogram model was constructed. A receiver operating characteristic curve was plotted to analyze the predicted values of the nomogram model. A calibration curve was constructed to analyze the calibration degree between the predicted results and the actual observation results. Decision curve analysis (DCA) was performed to evaluate whether the constructed nomogram model has clinical benefit for use in clinical applications. *p*-values of <0.05 were considered to indicate statistical significance.

Results

Clinical baseline characteristics

In this study, 936 participants were randomized to the training set ($n=680$) and validation set ($n=256$) at a sampling ratio of 7:3. Table 1 shows the comparison of participants' clinical baseline data between the training and validation sets. Among the included variables, no statistically significant differences were observed between the training and validation groups ($p > 0.05$), indicating that the grouping was reasonable.

Comparison of clinical data between patients with colorectal SSLs and healthy controls in the training set

Table 2 illustrates statistically significant differences between patients with colorectal SSLs and healthy controls in the training set with regard to age, sex, BMI, history of smoking, history of alcohol consumption, WBC, NEU, NLR, MON, RBC, Hb level, blood glucose level, TG level, TC/HDL, TG/HDL, and TyG index ($p < 0.05$). Patients diagnosed with colorectal SSLs were older, mostly male, had a history of smoking and alcohol consumption, and had higher levels of BMI, WBC, NEU, NLR, MON, RBC, Hb, blood glucose, TG, TC/HDL, TG/HDL, and TyG index. In contrast, there were no statistically significant differences between the two groups with regard to other measures such as LYM, EOS, BAS, TC, LDL/HDL, and history of diabetes and hypertension ($p > 0.05$).

Screening of independent risk factors for the occurrence of colorectal SSLs

The statistically different variables obtained from the above univariate analysis were included in the multivariate logistic regression analysis. The Hosmer–Lemeshow test

Table 1. Clinical baseline characteristics in the training and validation sets.

	Training set (n = 656)	Validation set (n = 280)	Statistics	p value
Age, years	51.000 (42.000–58.000)	51.000 (42.250–58.000)	Z = −0.329	0.742
BMI, kg/m ²	24.385 (22.150–26.830)	24.445 (22.490–26.780)	Z = −0.021	0.983
WBC, 10 ⁹ /L	5.580 (4.760–6.690)	5.440 (4.565–6.720)	Z = −1.090	0.276
NEU, 10 ⁹ /L	3.240 (2.543–3.998)	3.080 (2.500–4.030)	Z = −1.119	0.263
LYM, 10 ⁹ /L	1.820 (1.533–2.260)	1.820 (1.430–2.270)	Z = −1.096	0.273
NLR	1.715 (1.363–2.240)	1.685 (1.353–2.198)	Z = −0.369	0.712
MON, 10 ⁹ /L	0.340 (0.270–0.420)	0.325 (0.260–0.410)	Z = −1.320	0.187
EOS, 10 ⁹ /L	0.110 (0.070–0.190)	0.110 (0.060–0.180)	Z = −1.505	0.132
BAS, 10 ⁹ /L	0.020 (0.010–0.030)	0.020 (0.010–0.030)	Z = −0.352	0.725
RBC, 10 ¹² /L	4.545 (4.200–4.890)	4.520 (4.200–4.905)	Z = −0.186	0.852
Hb, g/L	140.000 (128.000–152.000)	140.000 (126.250–151.000)	Z = −0.501	0.616
Blood glucose, mmol/L	5.100 (4.640–5.658)	5.100 (4.700–5.600)	Z = −0.232	0.816
TC, mmol/L	4.820 (4.173–5.520)	4.870 (4.180–5.440)	Z = −0.040	0.968
TG, mmol/L	1.380 (0.910–2.108)	1.295 (0.880–1.965)	Z = −1.492	0.136
HDL, mmol/L	1.200 (1.020–1.448)	1.220 (1.050–1.470)	Z = −0.717	0.473
LDL, mmol/L	3.040 (2.570–3.550)	3.060 (2.580–3.510)	Z = −0.235	0.815
TC/HDL	3.950 (3.273–4.745)	3.965 (3.170–4.648)	Z = −0.682	0.496
TG/HDL	1.170 (0.650–1.960)	1.105 (0.660–1.715)	Z = −1.345	0.179
LDL/HDL	2.535 (2.030–3.088)	2.515 (1.990–3.055)	Z = −0.701	0.483
TyG index	8.660 (8.190–9.140)	8.560 (8.150–9.060)	Z = −1.376	0.169
Sex				
Male	337 (51.372%)	145 (51.786%)	$\chi^2 = 0.013$	0.908
Female	319 (48.628%)	135 (48.214%)		
Smoking				
Yes	134 (20.427%)	66 (23.571%)	$\chi^2 = 1.155$	0.283
No	522 (79.573%)	214 (76.429%)		
Alcohol				
Yes	94 (14.329%)	33 (11.786%)	$\chi^2 = 1.083$	0.298
No	562 (85.671%)	247 (88.214%)		
Hypertension				
Yes	63 (9.604%)	17 (6.071%)	$\chi^2 = 3.132$	0.077
No	593 (90.396%)	263 (93.929%)		
Diabetes				
Yes	162 (24.695%)	57 (20.357%)	$\chi^2 = 2.060$	0.151
No	494 (75.305%)	223 (79.643%)		

BAS: basophil count; BMI: body mass index; EOS: eosinophil count; Hb: hemoglobin; HDL: high-density lipoprotein; LDL: low-density lipoprotein; LYM: lymphocyte count; MON: monocyte count; NEU: neutrophil count; NLR: neutrophil/lymphocyte ratio; RBC: red blood cell count; TC: total cholesterol; TG: triglyceride; TyG index: triglyceride–glucose index; WBC: white blood cell count.

showed a good fitting degree ($\chi^2 = 4.308$, $p = 0.828$). As shown in Table 3, age, history of smoking, history of alcohol consumption, and TyG index were independent risk factors for the occurrence of colorectal SSLs ($p < 0.05$). The risk of colorectal

SSLs increased (odds ratio (OR): 1.048, 95% confidence interval (CI): 1.029–1.067) with aging ($p < 0.05$). Compared with healthy controls, patients with a history of smoking and alcohol consumption had a nearly doubled risk of SSLs (OR: 1.754,

Table 2. Comparison of baseline characteristics between patients with SSLs and healthy controls in the training set.

	SSL group (n = 218)	Control group (n = 438)	Statistics	p value
Age, years	54.000 (46.000–60.000)	50.000 (42.000–56.000)	Z = −4.470	0.000
BMI, kg/m ²	24.910 (23.025–27.430)	24.190 (22.015–26.670)	Z = −2.892	0.004
WBC, 10 ⁹ /L	5.935 (4.975–6.713)	5.470 (4.645–6.648)	Z = −2.732	0.006
NEU, 10 ⁹ /L	3.345 (2.740–4.108)	3.150 (2.410–3.943)	Z = −2.904	0.004
LYM, 10 ⁹ /L	1.805 (1.490–2.323)	1.830 (1.548–2.240)	Z = −0.003	0.998
NLR	1.815 (1.418–2.335)	1.680 (1.298–2.200)	Z = −2.280	0.023
MON, 10 ⁹ /L	0.360 (0.290–0.440)	0.320 (0.260–0.410)	Z = −2.733	0.006
EOS, 10 ⁹ /L	0.120 (0.070–0.210)	0.110 (0.070–0.170)	Z = −1.438	0.151
BAS, 10 ⁹ /L	0.020 (0.010–0.030)	0.020 (0.010–0.030)	Z = −0.145	0.885
RBC, 10 ¹² /L	4.653 ± 0.486	4.511 ± 0.496	t = 3.469	0.001
Hb, g/L	143.769 ± 18.232	137.132 ± 16.944	t = 4.606	0.000
Blood glucose, mmol/L	5.195 (4.700–5.820)	5.020 (4.600–5.600)	Z = −2.261	0.024
TC, mmol/L	4.895 (4.225–5.598)	4.790 (4.158–5.493)	Z = −1.007	0.314
TG, mmol/L	1.605 (1.068–2.343)	1.270 (0.850–1.950)	Z = −4.606	0.000
HDL, mmol/L	1.170 (1.008–1.443)	1.238 (1.028–1.458)	Z = −1.380	0.168
LDL, mmol/L	3.050 (2.588–3.550)	3.040 (2.540–3.553)	Z = −0.095	0.924
TC/HDL	4.020 (3.400–4.845)	3.900 (3.173–4.720)	Z = −2.009	0.045
TG/HDL	1.375 (0.790–2.263)	1.060 (0.600–1.790)	Z = −3.828	0.000
LDL/HDL	2.540 (2.120–3.098)	2.530 (2.005–3.083)	Z = −0.719	0.472
TyG index	8.820 (8.393–9.233)	8.550 (8.125–9.025)	Z = −4.719	0.000
Sex				
Male	144 (66.055%)	193 (44.064%)	$\chi^2 = 28.178$	0.000
Female	74 (33.945%)	245 (55.936%)		
Smoking				
Yes	72 (33.028%)	62 (14.155%)	$\chi^2 = 31.894$	0.000
No	146 (66.972%)	376 (85.845%)		
Alcohol				
Yes	56 (25.688%)	38 (8.676%)	$\chi^2 = 34.316$	0.000
No	162 (74.312%)	400 (91.324%)		
Hypertension				
Yes	25 (11.468%)	38 (8.676%)	$\chi^2 = 1.307$	0.253
No	193 (88.532%)	400 (91.324%)		
Diabetes				
Yes	62 (28.440%)	100 (22.831%)	$\chi^2 = 2.463$	0.117
No	156 (71.560%)	338 (77.169%)		

BAS: basophil count; BMI: body mass index; EOS: eosinophil count; Hb: hemoglobin; HDL: high-density lipoprotein; LDL: low-density lipoprotein; LYM: lymphocyte count; MON: monocyte count; NEU: neutrophil count; NLR: neutrophil/lymphocyte ratio; RBC: red blood cell count; TC: total cholesterol; TG: triglyceride; TyG index, triglyceride–glucose index; WBC, white blood cell count.

95% CI: 1.055–2.916 and OR: 1.996, 95% CI: 1.144–3.483, respectively; $p < 0.05$). In addition, the TyG index was an independent risk factor for the occurrence of colorectal SSLs (Figure 1).

Construction and verification of the nomogram model

Statistically significant variables were screened using multivariate logistic regression analysis to construct a nomogram

Table 3. Multivariate logistic regression analysis of SSLs in the training set.

	Regression coefficient	Standard error (SE)	Wald	p value	OR	95% CI	
						Lower bound	Upper bound
Age	0.047	0.009	25.734	0.000	1.048	1.029	1.067
Sex (male)	0.336	0.264	1.626	0.202	1.400	0.835	2.347
BMI	0.008	0.030	0.072	0.788	1.008	0.950	1.069
Smoking (yes)	0.562	0.259	4.698	0.030	1.754	1.055	2.916
Alcohol (yes)	0.691	0.284	5.919	0.015	1.996	1.144	3.483
WBC	0.228	0.173	1.728	0.189	1.256	0.894	1.765
NEU	−0.282	0.270	1.095	0.295	0.754	0.444	1.279
NLR	0.261	0.252	1.073	0.300	1.299	0.792	2.130
MON	−0.433	0.925	0.219	0.640	0.649	0.106	3.975
RBC	−0.016	0.360	0.002	0.964	0.984	0.486	1.990
Hb	0.011	0.010	1.110	0.292	1.011	0.991	1.032
Blood glucose	−0.126	0.080	2.497	0.114	0.881	0.754	1.031
TG	0.002	0.138	0.000	0.988	1.002	0.765	1.312
TC/HDL	−0.050	0.104	0.232	0.630	0.951	0.775	1.166
TG/HDL	−0.224	0.188	1.415	0.234	0.800	0.553	1.156
TyG index	0.798	0.362	4.859	0.027	2.221	1.093	4.517

BMI: body mass index; CI: confidence interval; Hb: hemoglobin; MON: monocyte count; NEU: neutrophil count; NLR: neutrophil/lymphocyte ratio; OR: odds ratio; RBC: red blood cell count; TG: triglyceride; TyG index: triglyceride–glucose index.

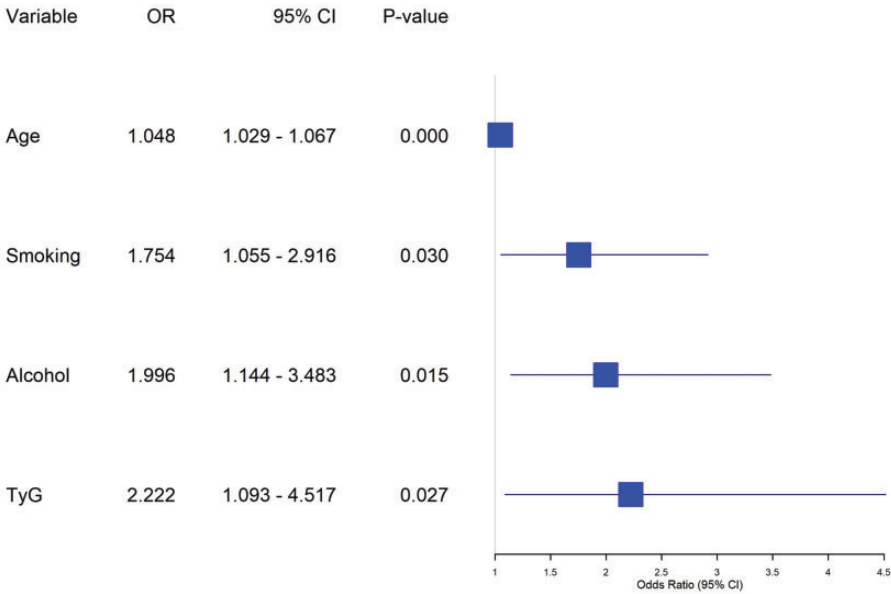


Figure 1. Forest plot. Multivariate logistic regression of colorectal sessile serrated lesions in the training set. Four statistically significant variables screened via multivariate logistic regression analysis were age, history of smoking, history of alcohol consumption, and TyG index. CI: confidence interval; OR: odds ratio; TyG: triglyceride–glucose.

model to predict the risk of colorectal SSLs (Figure 2). The scores of each variable were added to determine the risk of colorectal SSLs corresponding to the total score.

The area under the curve (AUC) values predicted by the nomogram model were 0.715 (95% CI: 0.676–0.753) in the training set and 0.742 (95% CI: 0.669–0.815) in the validation set (Figure 3).

Calibration curve analysis further showed that the predicted results of the nomogram model were in good agreement with the actual observation results. The Hosmer–Lemeshow test at $p > 0.05$ suggested a good fitting model (Figure 4).

DCA showed that the decision curve was above the None line and All line. The model provided a positive net benefit within a certain threshold probability range. This affirmed the model's potential ability for risk assessment of colorectal SSLs in clinical practice (Figure 5).

Discussion

Serrated colorectal polyps were first reported by Morson in 1962 and included metaplastic or hyperplastic polyps, which were considered non-neoplastic lesions lacking malignant potential at that time.¹⁵ Researchers gradually discovered other similar lesions such as SSLs, which are neoplastic lesions with cancerous potential and the main precancerous lesions of serrated CRC. They are also a significant contributing factor to the occurrence of postcolonoscopy CRC.¹⁶ The natural progression of SSLs is a long-term process that usually takes 10–15 years to develop into dysplasia and eventually into malignancy.^{17–19} However, the serrated pathway tends to be cancerous at a faster rate compared with CRC developed via the classical adenoma–carcinoma pathway. Once SSLs appear in areas with dysplasia, the lesion may develop into CRC within a short time, with a poor overall prognosis.^{20,21}

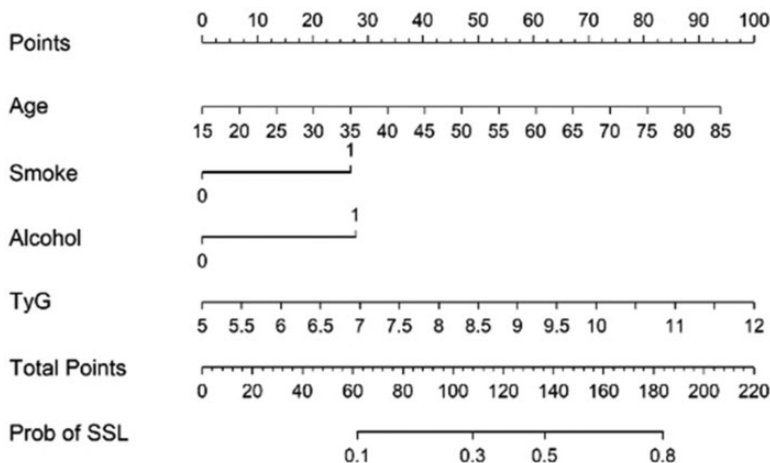


Figure 2. Nomogram for the prediction of colorectal SSLs. For each variable, a vertical line was drawn from its respective axis upward to intersect the corresponding score line to determine the score associated with the variable. The individual scores of all variables were then added to obtain the total score. A vertical line was subsequently drawn from the total score axis downward to intersect the probability axis to determine the predicted probability of SSLs. TyG: triglyceride–glucose; SSL: sessile serrated lesion.

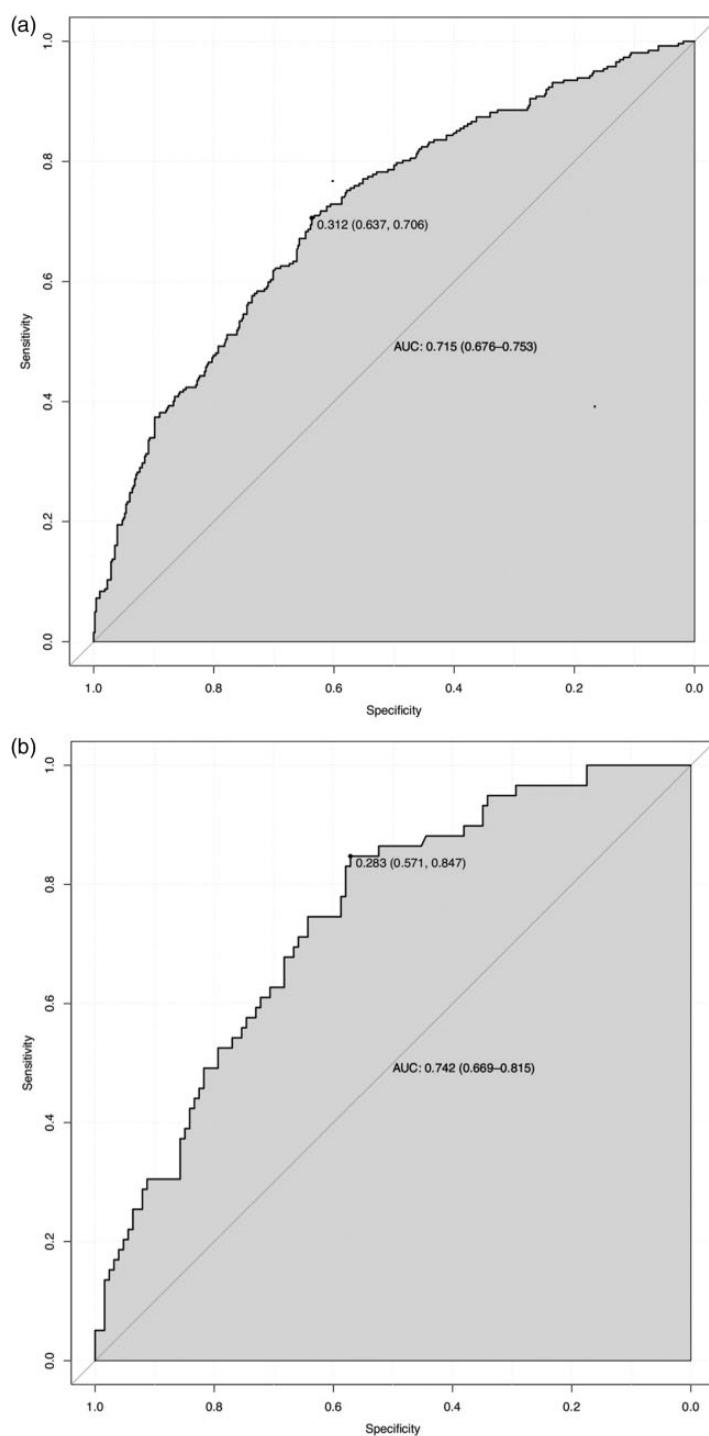


Figure 3. Receiver operating characteristic curves for the probability of colorectal sessile serrated lesions in the training (a) and validation (b) sets. AUC: area under the curve.

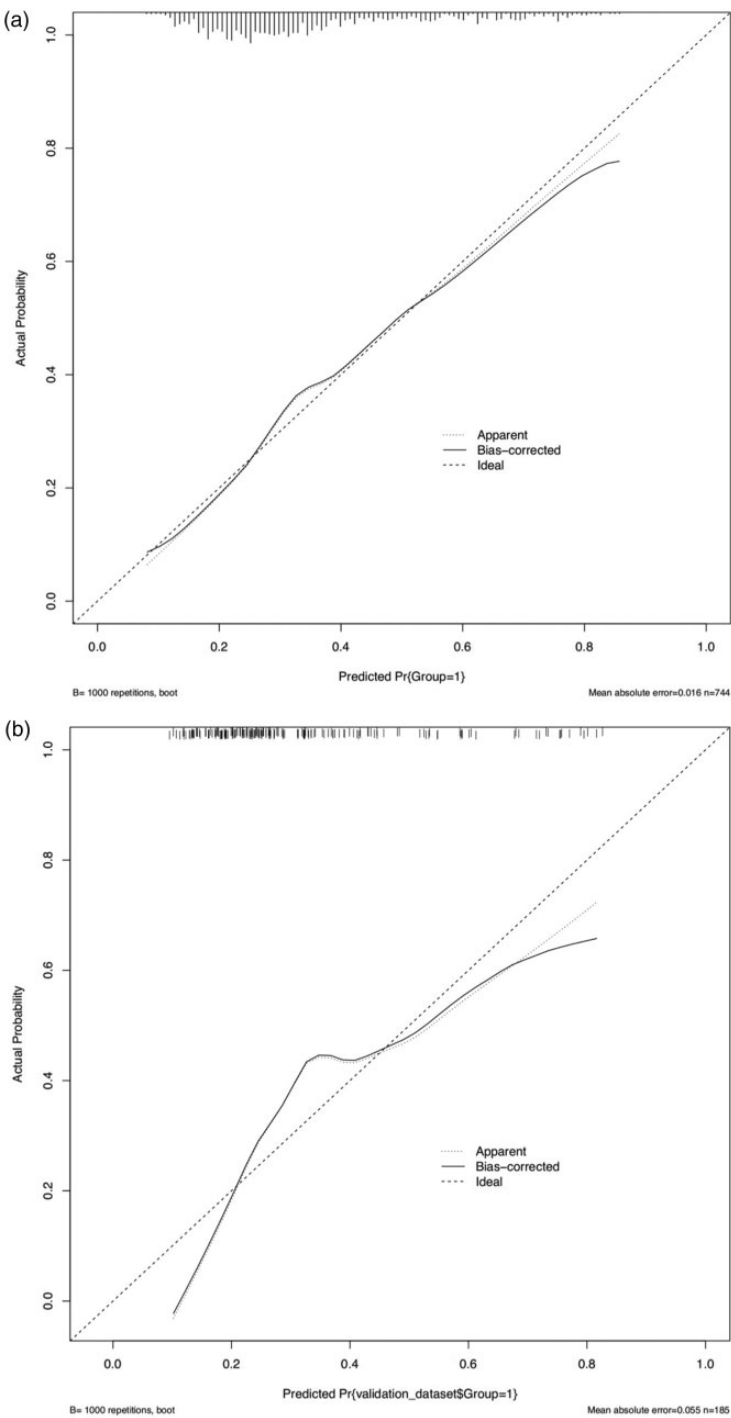


Figure 4. Calibration curves for the risk nomogram model of colorectal sessile serrated lesions in the training (a) and validation (b) sets.

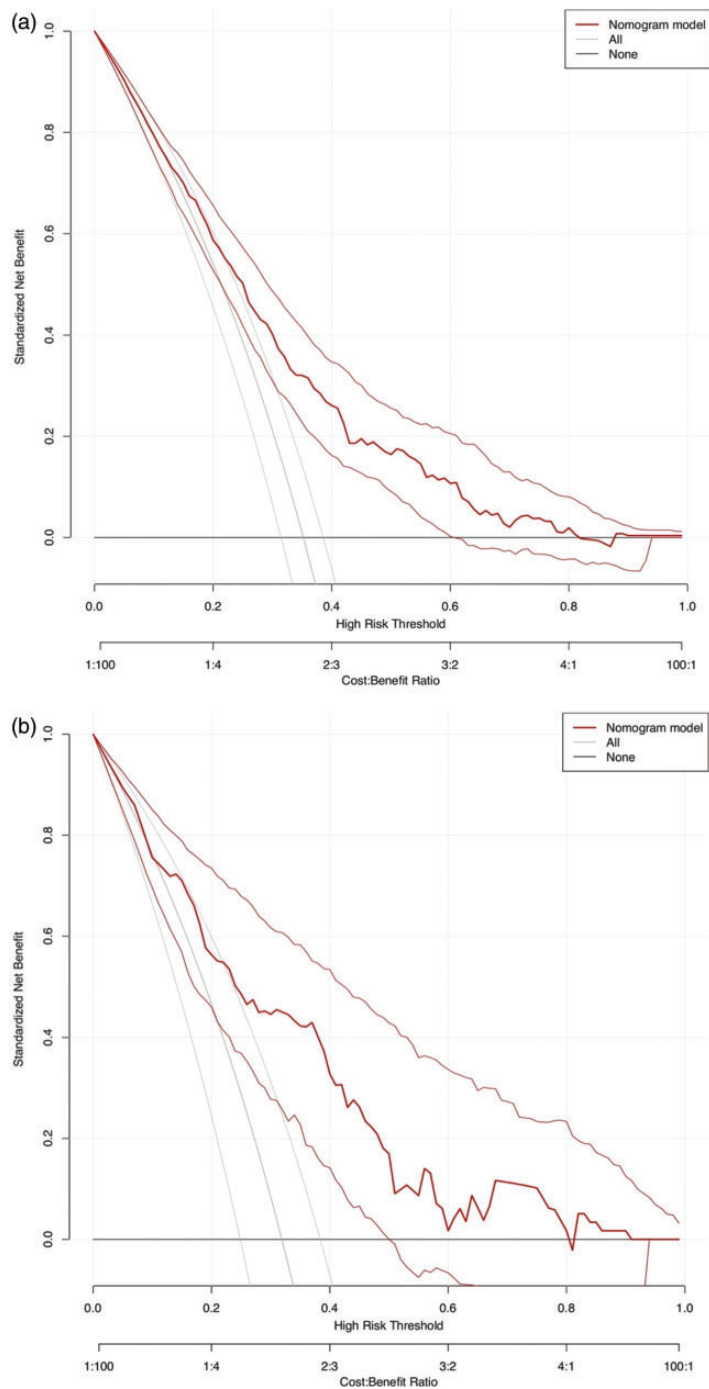


Figure 5. Decision curve analysis for the risk nomogram model of colorectal sessile serrated lesions in the training (a) and validation (b) sets.

It is currently believed that early detection and treatment of colorectal precancerous lesions is one of the effective ways to reduce the incidence and mortality of CRC. The science popularization and screening of CRC have been conducted as early as possible in European and American countries as well as in Japan and South Korea. Therefore, CRC screening is generally more popular in these countries, with a 5-year survival rate of $\geq 64\%$. However, in China, CRC screening started late, the science popularization work has been inadequate, and the population acceptance is generally low. Therefore, the early diagnosis rate of CRC is low in China, with a 5-year survival rate of only 57.6%.^{22,23} Colonoscopy is currently the most important method for screening colorectal SSLs. However, colonoscopy resources in China are scarce and unevenly distributed. High-quality endoscopy requires good intestinal preparation, sufficient withdrawal time, pigment endoscopy, and complete theoretical knowledge of endoscopists. Risk stratification has not yet been achieved, exposing some low-risk groups to risks caused by unnecessary invasive examinations. Hence, endoscopic screening strategies may not be as cost-effective as expected.

First, high-risk patients with colorectal SSLs are screened by non-invasive screening techniques based on questionnaires and blood, urine, stool samples, etc. Then, the targeted colonoscopy is more suitable for Chinese conditions.²⁴ Therefore, this study analyzed the risk factors for colorectal SSLs by comparing the clinical data of patients with colorectal SSLs with the healthy population. A prediction model was constructed to assist clinicians in identifying high-risk patients as early as possible and arranging for further colonoscopy. This may avoid overmedication and save medical resources to a certain extent.

This study found that age, history of smoking, history of alcohol consumption,

and TyG index were independent predictors of colorectal SSLs. This result is partially consistent with those of previous studies. With aging, the risk of colorectal SSLs increases. Anwar et al.²⁵ compared the results of initial follow-up colonoscopy in 6297 patients and found that older age (≥ 75 years) was independently associated with high-risk sessile serrated adenomas at follow-up. This provides evidence regarding whether the relevant population will undergo subsequent monitoring colonoscopy. IJspeert et al.²⁶ concluded that patients aged < 50 years had a lower risk of SSLs than those aged > 50 years, but the risk did not increase significantly with aging. More importantly, it was also found that smoking could greatly increase the risk of colorectal SSLs. Choe et al.²⁷ suggested that smoking is the only important risk factor for the progression of colorectal SSLs (OR: 1.394, 95% CI: 1.012–1.764). Bailie et al.²⁸ reported that smoking increased the risk of colorectal SSLs by more than 3-fold (OR: 3.40, 95% CI: 1.90–6.07). Recently, a meta-analysis²⁹ showed that smoking increased the risk of CRC in a dose-dependent manner, and the duration and intensity of smoking were directly proportional to CRC. Therefore, smoking may increase the risk of CRC through the microsatellite instability pathway (characterized by microsatellite instability, CpG island methylator phenotype, and *BRAF* mutation), which is one of the molecular mechanism characteristics of colorectal SSLs.⁶ Several studies have also shown an association between alcohol consumption and colorectal SSLs. High levels of alcohol consumption are associated with an increased risk of colorectal SSLs, with an OR of approximately 1.8 compared with non-drinkers.^{28,30} Furthermore, the results of this study suggest that the TyG index is an independent risk factor for colorectal SSLs. Insulin resistance (IR) is a known risk factor for CRC.^{31,32} The TyG index is

a reliable surrogate for assessing IR.^{33,34} Relevant studies have also demonstrated that the elevated TyG index is a risk factor for CRC.^{12–14} However, there are only a few studies on the direct correlation between the TyG index and colorectal SSLs, and it remains uncertain whether there exists a correlation between them. TyG index is a surrogate index of IR that is closely related to metabolic syndrome.³⁵ A large number of previous studies have revealed that metabolic-related factors such as abnormal blood lipid levels, obesity, and diabetes are closely related to SSLs, which may suggest an association between the TyG index and colorectal SSLs.^{36–39} These results suggest that colorectal SSLs can be prevented to some extent by altering modifiable risk factors (such as improving lifestyle, reducing smoking and drinking, and maintaining a healthy metabolic state through dietary control and exercise). Future studies may further explore the mediating mechanisms of these risk factors and provide a scientific basis for the development of personalized preventive measures.

This study comprehensively included the risk factors for colorectal SSLs. Multiple covariates were considered in the multivariate logistic regression analysis to more accurately assess the association between variables and colorectal SSLs. This may reduce the risk of possible confounding effects in the univariate analysis and thus generate results with greater explanatory power and clinical significance. Then, a nomogram was constructed to determine the risk of colorectal SSLs. According to the AUC values and calibration curve results, the nomogram showed good discrimination and calibration ability in predicting the risk of colorectal SSLs, which indicated its high clinical value. The decision curve further showed that this model provided positive net benefits within a certain threshold probability range. These findings highlight the potential of the

nomogram as a practical tool for clinicians to identify and stratify the risk of colorectal SSLs, which may contribute to timely preventive interventions.

There are still certain limitations in this study. First, this was a retrospective study. Multivariate analysis revealed that a history of smoking and history of alcohol consumption are independent predictors of colorectal SSLs. The medical history mainly depends on the self-report of the participants, which may cause bias. Moreover, it is difficult to further analyze the correlation between the duration and amount of smoking and alcohol consumption and colorectal SSLs. Second, this is a single-center study with a small sample size and lacks external tests from other institutions. Therefore, the risk prediction effect of other populations outside this medical institution remains unknown. Future studies should conduct a multicenter prospective study to verify the effectiveness of this prediction model. Finally, although data on a large number of variables have been collected, some predictors (such as dietary composition) that may be related to the occurrence and development of colorectal SSLs are not included in this study. Future studies need to continue the collection of relevant data to comprehensively assess the risk factors for colorectal SSLs.

In conclusion, age, history of smoking, history of alcohol consumption, and TyG index are independent predictors of colorectal SSLs. The prediction model of colorectal SSLs constructed in this study demonstrates high diagnostic efficacy, which is helpful for clinical risk assessment and decision-making.

Acknowledgements

We thank Professor Zhaolin Lu from the School of Information and Control Engineering at China University of Mining and Technology who shared theoretical knowledge regarding artificial intelligence.

Author contributions

Sihui Huang: Conceptualization, Methodology, Data curation, and Writing—Original draft preparation. Shiyu Liu: Methodology and Data curation. Fang Tan & Hu Chen: Methodology. Guangxia Chen: Conceptualization, Supervision, and Writing—Reviewing and editing. All authors have read and approved the final manuscript.

Consent to participate

Written informed consent was obtained from the participants before starting colonoscopy. Patient data were not shared with third parties.

Data availability statement

Not applicable.

Declaration of conflicting interest

The authors declare that there is no conflict of interest.

Ethical considerations

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Affiliated Xuzhou Municipal Hospital of Xuzhou Medical University Research Ethics Committee (approval no. 2023-KY-059) in August 2023.

Funding

This study was supported by the Xuzhou Key R&D Program (Social Development) Project (No.KC22095), the Xuzhou Municipal Health and Health Commission Medical Leading Talents Training Program (No.XWRCHT20210025), and Xuzhou Key Research and Development Program General Project (No. KC23167).

ORCID iDs

Sihui Huang  <https://orcid.org/0009-0009-3708-4475>

Guangxia Chen  <https://orcid.org/0000-0002-3540-3382>

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN

- estimates of incidence and mortality world-wide for 36 cancers in 185 countries. *CA Cancer J Clin* 2024; 74: 229–263.
2. Leggett B and Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* 2010; 138: 2088–2100.
3. Kim SY and Kim TI. Serrated neoplasia pathway as an alternative route of colorectal cancer carcinogenesis. *Intest Res* 2018; 16: 358–365.
4. Nguyen LH, Goel A and Chung DC. Pathways of colorectal carcinogenesis. *Gastroenterology* 2020; 158: 291–302.
5. Nagtegaal ID, Odze RD, Klimstra D; WHO Classification of Tumours Editorial Board, et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology* 2020; 76: 182–188.
6. Crockett SD and Nagtegaal ID. Terminology, molecular features, epidemiology, and management of serrated colorectal neoplasia. *Gastroenterology* 2019; 157: 949–966.e4.
7. Von Elm E, Altman DG, Egger M; STROBE Initiative, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Ann Intern Med* 2007; 147: 573–577.
8. Carethers JM and Jung BH. Genetics and genetic biomarkers in sporadic colorectal cancer. *Gastroenterology* 2015; 149: 1177–1190. e3.
9. Balkwill F and Mantovani A. Inflammation and cancer: back to Virchow? *Lancet* 2001; 357: 539–545.
10. Xu S and Dou ZL. Diagnostic value of NLR, PLR, MLR, and SII in colorectal cancer. *Med Inf* 2024; 37: 85–89.
11. Karadeniz FÖ, Sancaktepe EA and Karadeniz Y. High triglyceride-glucose index is associated with poor prognosis in patients with acute coronary syndrome in long-term follow-up. *Angiology* 2023; 74: 139–148.
12. Okamura T, Hashimoto Y, Hamaguchi M, et al. Triglyceride-glucose index (TyG index) is a predictor of incident colorectal cancer: a population-based longitudinal study. *BMC Endocr Disord* 2020; 20: 113.

13. Fritz J, Bjørge T, Nagel G, et al. The triglyceride-glucose index as a measure of insulin resistance and risk of obesity-related cancers. *Int J Epidemiol* 2020; 49: 193–204.
14. Liu T, Zhang Q, Wang Y, et al. Association between the TyG index and TG/HDL-C ratio as insulin resistance markers and the risk of colorectal cancer. *BMC Cancer* 2022; 22: 1007.
15. Morson BCMD. Some peculiarities in the histology of intestinal polyps. *Dis Colon Rectum* 1962; 5: 337–344.
16. Anderson JC, Rex DK, Mackenzie TA, et al. Higher serrated polyp detection rates are associated with lower risk of postcolonoscopy colorectal cancer: data from the new Hampshire colonoscopy registry. *Am J Gastroenterol* 2023; 118: 1927–1930.
17. Bettington M, Walker N, Rosty C, et al. Clinicopathological and molecular features of sessile serrated adenomas with dysplasia or carcinoma. *Gut* 2017; 66: 97–106.
18. Bouwens MWE, Riedl RG, Bosman FT, et al. Large proximal serrated polyps: natural history and colorectal cancer risk in a retrospective series. *J Clin Gastroenterol* 2013; 47: 734–735.
19. Holme Ø, Bretthauer M, Eide TJ, et al. Long-term risk of colorectal cancer in individuals with serrated polyps. *Gut* 2015; 64: 929–936.
20. Lash RH, Genta RM and Schuler CM. Sessile serrated adenomas: prevalence of dysplasia and carcinoma in 2139 patients. *J Clin Pathol* 2010; 63: 681–686.
21. Omori K, Yoshida K, Tamiya S, et al. Endoscopic observation of the growth process of a right-side sessile serrated adenoma/polyp with cytological dysplasia to an invasive submucosal adenocarcinoma. *Case Rep Gastrointest Med* 2016; 2016: 6576351.
22. Basu P, Ponti A, Anttila A, et al. Status of implementation and organization of cancer screening in The European Union Member States-summary results from the second European screening report. *Int J Cancer* 2018; 142: 44–56.
23. Allemani C, Matsuda T, Di Carlo V; CONCORD Working Group, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37513025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 2018; 391: 1023–1075.
24. National Clinical Research Center for Digestive Diseases (Shanghai), Chinese Society of Digestive Endoscopy, Cancer Endoscopy Professional Committee of China Anti-Cancer Association, et al. Expert consensus on the management strategies for precancerous lesions and precancerous conditions of colorectal cancer in China. *Chin J Dig Endosc* 2022; 39: 1–18.
25. Anwar S, Cock C, Young J, et al. Features associated with high-risk sessile serrated polyps at index and follow-up colonoscopy. *J Gastroenterol Hepatol* 2021; 36: 1620–1626.
26. IJspeert JEG, Bevan R, Senore C, et al. Detection rate of serrated polyps and serrated polyposis syndrome in colorectal cancer screening cohorts: a European overview. *Gut* 2017; 66: 1225–1232.
27. Choe AR, Song EM, Seo H, et al. Different modifiable risk factors for the development of non-advanced adenoma, advanced adenomatous lesion, and sessile serrated lesions, on screening colonoscopy. *Sci Rep* 2024; 14: 16865.
28. Bailie L, Loughrey MB and Coleman HG. Lifestyle risk factors for serrated colorectal polyps: a systematic review and meta-analysis. *Gastroenterology* 2017; 152: 92–104.
29. Botteri E, Borroni E, Sloan EK, et al. Smoking and colorectal cancer risk, overall and by molecular subtypes: a meta-analysis. *Am J Gastroenterol* 2020; 115: 1940–1949.
30. He X, Wu K, Ogino S, et al. Association between risk factors for colorectal cancer and risk of serrated polyps and conventional adenomas. *Gastroenterology* 2018; 155: 355–373. e18.
31. Farahani H, Mahmoudi T, Asadi A, et al. Insulin resistance and colorectal cancer risk: the role of elevated plasma resistin levels. *J Gastrointest Cancer* 2020; 51: 478–483.
32. Cirillo F, Catellani C, Sartori C, et al. Obesity, insulin resistance, and colorectal cancer: could miRNA dysregulation play a role? *Int J Mol Sci* 2019; 20: 2922.

33. Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. *J Clin Endocrinol Metab* 2010; 95: 3347–3351.
34. Sánchez-García A, Rodríguez-Gutiérrez R, Mancillas-Adame L, et al. Diagnostic accuracy of the triglyceride and glucose index for insulin resistance: a systematic review. *Int J Endocrinol* 2020; 2020: 4678526.
35. Tahapary DL, Pratisthita LB, Fitri NA, et al. Challenges in the diagnosis of insulin resistance: focusing on the role of HOMA-IR and triglyceride/glucose index. *Diabetes Metab Syndr* 2022; 16: 102581.
36. Fliss-Isakov N, Zelber-Sagi S, Webb M, et al. Distinct metabolic profiles are associated with colorectal adenomas and serrated polyps. *Obesity (Silver Spring)* 2017; 25: S72–S80.
37. Tang C, Li J, Yang Z, et al. Different lipid metabolic profiles and their associated genes in sessile serrated adenoma or polyps compared to hyperplastic polyps. *Am J Cancer Res* 2022; 12: 1982–1994.
38. Wang X, Zou Y, Zhang R, et al. The relationship between serum lipid levels and colorectal serrated lesions: a systematic review and meta-analysis. *Front Physiol* 2022; 13: 984586.
39. Zhang R, Ni Y, Guo CL, et al. Risk factors for sessile serrated lesions among Chinese patients undergoing colonoscopy. *J Gastroenterol Hepatol* 2023; 38: 1468–1473.