ISSN 1948-5204 (online)

World Journal of *Gastrointestinal Oncology*

World J Gastrointest Oncol 2024 June 15; 16(6): 2264-2866

Published by Baishideng Publishing Group Inc

World Journal of Gastrointestinal Oncology

Contents

Monthly Volume 16 Number 6 June 15, 2024

EDITORIAL

2264 Dual primary gastric and colorectal cancer: The known hereditary causes and underlying mechanisms
Azer SA

2271 Application of *Fusobacterium nucleatum* as a biomarker in gastrointestinal malignancies
Yu LC, Li YP, Xin YM, Mao M, Pan YX, Qu YX, Luo ZD, Zhang Y, Zhang X

2284 T1 colorectal cancer management in the era of minimally invasive endoscopic resection
Jiang SX, Zarrin A, Shahidi N

2295 Mixed neuroendocrine and adenocarcinoma of gastrointestinal tract: A complex diagnosis and therapeutic challenge
Shenoy S

2300 Advancements in breath-based diagnostics for pancreatic cancer: Current insights and future perspectives
Tez M, Şahingöz E, Martlı HF

2304 Colorectal cancer and dormant metastases: Put to sleep or destroy?
Senchukova MA

REVIEW

2318 Advances in targeted therapy for human epidermal growth factor receptor 2 positive in advanced gastric cancer
Jiang YK, Li W, Qiu YY, Yue M

2335 Research progress of ferroptosis regulating lipid peroxidation and metabolism in occurrence and development of primary liver cancer
Shu YJ, Lao B, Qiu YY

2350 Early monitoring values of oncogenic signalling molecules for hepatocellular carcinoma
Yao M, Fang RF, Xie Q, Xu M, Sai WL, Yao DF

MINIREVIEWS

2362 Therapeutic strategies targeting the epidermal growth factor receptor signaling pathway in metastatic colorectal cancer
Zhou Y, Wu S, Qu FJ

2380 Predicting the prognosis of hepatic arterial infusion chemotherapy in hepatocellular carcinoma
Wang QF, Li ZW, Zhou HF, Zhu KZ, Wang YJ, Wang YQ, Zhang YW

2394 Unraveling colorectal cancer prevention: The vitamin D - gut flora - immune system nexus

Zhan ZS, Zheng ZS, Shi J, Chen J, Wu SY, Zhang SY

ORIGINAL ARTICLE

Retrospective Cohort Study

2404 Unveiling the secrets of gastrointestinal mucous adenocarcinoma survival after surgery with artificial intelligence: A population-based study

Song J, Yan XX, Zhang FL, Lei YY, Ke ZY, Li F, Zhang K, He YQ, Li W, Li C, Pan YM

2419 Analysis of metabolic characteristics of metabolic syndrome in elderly patients with gastric cancer by non-targeted metabolomics

Zhang H, Shen WB, Chen L

Retrospective Study

2429 Predictive value of preoperative routine examination for the prognosis of patients with pT2N0M0 or pT3N0M0 colorectal cancer

Jing PF, Chen J, Yu ED, Miao CY

2439 Simplified liver imaging reporting and data system for the diagnosis of hepatocellular carcinoma on gadoxetic acid-enhanced magnetic resonance imaging

Lyu R, Hu WJ, Wang D, Wang J, Ye YB, Jia KF

2449 Efficacy comparison of fruquintinib, regorafenib monotherapy or plus programmed death-1 inhibitors for microsatellite stable metastatic colorectal cancer

An TQ, Qiu H, Zhou QB, Zong H, Hu S, Lian YG, Zhao RH

2463 Development of a diagnostic nomogram for alpha-fetoprotein-negative hepatocellular carcinoma based on serological biomarkers

He L, Zhang C, Liu LL, Huang LP, Lu WJ, Zhang YY, Zou DY, Wang YF, Zhang Q, Yang XL

2476 Drug-eluting bead transarterial chemoembolization as neoadjuvant therapy pre-liver transplantation for advanced-stage hepatocellular carcinoma

Ye ZD, Zhuang L, Song MC, Yang Z, Zhang W, Zhang JF, Cao GH

2487 Association between *Helicobacter pylori* infection, mismatch repair, HER2 and tumor-infiltrating lymphocytes in gastric cancer

Castaneda CA, Castillo M, Bernabe LA, Sanchez J, Fassan M, Tello K, Wistuba II, Chavez Passiuri I, Ruiz E, Sanchez J, Barreda F, Valdivia D, Bazan Y, Abad-Licham M, Mengoa C, Fuentes H, Montenegro P, Poquioma E, Alatrista R, Flores CJ, Taxa L

2504 Impact of baseline hepatitis B virus viral load on the long-term prognosis of advanced hepatocellular carcinoma treated with immunotherapy

Pan D, Liu HN, Yao ZY, Chen XX, Li YQ, Zhu JJ, Han ZX, Qin XB

2520 Prediction of pathological complete response and prognosis in locally advanced rectal cancer

Xu YJ, Tao D, Qin SB, Xu XY, Yang KW, Xing ZX, Zhou JY, Jiao Y, Wang LL

World Journal of Gastrointestinal Oncology

Monthly Volume 16 Number 6 June 15, 2024

Contents

Observational Study

2531 Extrahepatic cholestasis associated with paracoccidioidomycosis: Challenges in the differential diagnosis of biliopancreatic neoplasia

dos Santos JS, de Moura Arrais V, Rosseto Ferreira WJ, Ribeiro Correa Filho R, Brunaldi MO, Kemp R, Sankanrakutty AK, Elias Junior J, Bellissimo-Rodrigues F, Martinez R, Zangiacomi Martinez E, Ardengh JC

Clinical and Translational Research

2541 Development of a novel staging classification for Siewert II adenocarcinoma of the esophagogastric junction after neoadjuvant chemotherapy

Zhang J, Liu H, Yu H, Xu WX

2555 N6-methyladenosine methylation regulates the tumor microenvironment of Epstein-Barr virus-associated gastric cancer

Zhang Y, Zhou F, Zhang MY, Feng LN, Guan JL, Dong RN, Huang YJ, Xia SH, Liao JZ, Zhao K

2571 Hepatocellular carcinoma: An analysis of the expression status of stress granules and their prognostic value

Ren QS, Sun Q, Cheng SQ, Du LM, Guo PX

2592 Comprehensive analysis of clinical and biological value of *ING* family genes in liver cancer

Liu SC

2610 Epidemiology and prognostic nomogram for locally advanced gastric signet ring cell carcinoma: A population-based study

Yu ZH, Zhang LM, Dai ZQ, Zhang MN, Zheng SM

2631 Socioeconomic traits and the risk of Barrett's esophagus and gastroesophageal reflux disease: A Mendelian randomization study

Liu YX, Bin CL, Zhang L, Yang WT, An BP

Basic Study

2646 Complement factor I knockdown inhibits colon cancer development by affecting Wnt/β-catenin/c-Myc signaling pathway and glycolysis

Du YJ, Jiang Y, Hou YM, Shi YB

2663 Fine-needle aspiration technique under endoscopic ultrasound guidance: A technical approach for RNA profiling of pancreatic neoplasms

Seyfedinova SS, Freylikhman OA, Sokolnikova PS, Samochernykh KA, Kostareva AA, Kalinina OV, Solonitsyn EG

2673 Comprehensive analysis of gene mutations and mismatch repair in Chinese colorectal cancer patients

Chen H, Jiang RY, Hua Z, Wang XW, Shi XL, Wang Y, Feng QQ, Luo J, Ning W, Shi YF, Zhang DK, Wang B, Jie JZ, Zhong DR

2683 Action of circulating and infiltrating B cells in the immune microenvironment of colorectal cancer by single-cell sequencing analysis

Zhang JP, Yan BZ, Liu J, Wang W

2697 Bidirectional effects of the tryptophan metabolite indole-3-acetaldehyde on colorectal cancer

Dai Z, Deng KL, Wang XM, Yang DX, Tang CL, Zhou YP

2716 Sm-like 5 knockdown inhibits proliferation and promotes apoptosis of colon cancer cells by upregulating p53, CDKN1A and TNFRSF10B

Mo CJ, Deng XY, Ma RL, Zhu K, Shi L, Li K

2727 Shi-pi-xiao-ji formula suppresses hepatocellular carcinoma by reducing cellular stiffness through upregulation of acetyl-coA acetyltransferase 1

Jian HY, Liang ZC, Wen H, Zhang Z, Zeng PH

2742 Aspirin suppresses hepatocellular carcinoma progression by inhibiting platelet activity

Zhao LJ, Wang ZY, Liu WT, Yu LL, Qi HN, Ren J, Zhang CG

2757 Circ_0004592: An auxiliary diagnostic biomarker for gastric cancer

Kong S, Xu YH, Zheng M, Ju SQ, Shi HC

2769 N-glycosylation of Wnt3 regulates the progression of hepatocellular carcinoma by affecting Wnt/β-catenin signal pathway

Zhang XZ, Mo XC, Wang ZT, Sun R, Sun DQ

SYSTEMATIC REVIEWS

2781 Ferroptosis regulating lipid peroxidation metabolism in the occurrence and development of gastric cancer

Wang LM, Zhang WW, Qiu YY, Wang F

META-ANALYSIS

2793 Meta-analysis of transarterial chemoembolization combined with cryoablation *vs* transarterial chemoembolization alone for ≥ 5 cm hepatocellular carcinoma

Cheng JF, Sun QL, Tang L, Xu XJ, Huang XZ

2804 Dynamic contrast enhanced ultrasound in differential diagnosis of hepatocellular carcinoma: A systematic review and meta-analysis

Esposto G, Santini P, Termite F, Galasso L, Mignini I, Ainora ME, Gasbarrini A, Zocco MA

2816 Correlation analysis of interstitial maturity and prognosis of colorectal cancer: Meta-analysis

Liu ZJ, Zhang XW, Liu QQ, Wang SZ

SCIENTOMETRICS

2826 Visualization analysis of research hotspots and trends on gastrointestinal tumor organoids

Wang G, Liu T, He WT

2842 Trends and hotspots in gastrointestinal neoplasms risk assessment: A bibliometric analysis from 1984 to 2022

Fu QQ, Ma L, Niu XM, Zhao HX, Ge XH, Jin H, Yu DH, Yang S

World Journal of Gastrointestinal Oncology

Monthly Volume 16 Number 6 June 15, 2024

Contents

LETTER TO THE EDITOR

2862 New perspectives in prognostication of hepatocellular carcinoma: The role and clinical implications of transient receptor potential family genes

Guan SH, Hu WJ, Wang XY, Gu YX, Zhou DH

RETRACTION NOTE

2865 Retraction note to: RNA-binding protein CPSF6 regulates IBSP to affect pyroptosis in gastric cancer

Wang XJ, Liu Y, Ke B, Zhang L, Liang H

World Journal of Gastrointestinal Oncology

Monthly Volume 16 Number 6 June 15, 2024

Contents

ABOUT COVER

Editorial Board Member of *World Journal of Gastrointestinal Oncology*, Samy Azer, FACG, MD, PhD, Professor, Department of Medical Education, King Saud University College of Medicine, Riyadh 11461, Saudi Arabia. azer2000@optusnet.com.au

AIMS AND SCOPE

The primary aim of *World Journal of Gastrointestinal Oncology* (*WJGO*, *World J Gastrointest Oncol*) is to provide scholars and readers from various fields of gastrointestinal oncology with a platform to publish high-quality basic and clinical research articles and communicate their research findings online.

WJGO mainly publishes articles reporting research results and findings obtained in the field of gastrointestinal oncology and covering a wide range of topics including liver cell adenoma, gastric neoplasms, appendiceal neoplasms, biliary tract neoplasms, hepatocellular carcinoma, pancreatic carcinoma, cecal neoplasms, colonic neoplasms, colorectal neoplasms, duodenal neoplasms, esophageal neoplasms, gallbladder neoplasms, *etc*.

INDEXING/ABSTRACTING

The *WJGO* is now abstracted and indexed in PubMed, PubMed Central, Science Citation Index Expanded (SCIE, also known as SciSearch®), Journal Citation Reports/Science Edition, Scopus, Reference Citation Analysis, China Science and Technology Journal Database, and Superstar Journals Database. The 2023 edition of Journal Citation Reports® cites the 2022 impact factor (IF) for *WJGO* as 3.0; IF without journal self cites: 2.9; 5-year IF: 3.0; Journal Citation Indicator: 0.49; Ranking: 157 among 241 journals in oncology; Quartile category: Q3; Ranking: 58 among 93 journals in gastroenterology and hepatology; and Quartile category: Q3. The *WJGO*'s CiteScore for 2023 is 4.2 and Scopus CiteScore rank 2023: Gastroenterology is 80/167; Oncology is 196/404.

RESPONSIBLE EDITORS FOR THIS ISSUE

Production Editor: *Si Zhao*; **Production Department Director:** *Xiang Li*; **Cover Editor:** *Jia-Ru Fan*.

NAME OF JOURNAL
World Journal of Gastrointestinal Oncology

ISSN
ISSN 1948-5204 (online)

LAUNCH DATE
February 15, 2009

FREQUENCY
Monthly

EDITORS-IN-CHIEF
Monjur Ahmed, Florin Burada

EDITORIAL BOARD MEMBERS
https://www.wjgnet.com/1948-5204/editorialboard.htm

PUBLICATION DATE
June 15, 2024

INSTRUCTIONS TO AUTHORS
https://www.wjgnet.com/bpg/gerinfo/204

GUIDELINES FOR ETHICS DOCUMENTS
https://www.wjgnet.com/bpg/GerInfo/287

GUIDELINES FOR NON-NATIVE SPEAKERS OF ENGLISH
https://www.wjgnet.com/bpg/gerinfo/240

PUBLICATION ETHICS
https://www.wjgnet.com/bpg/GerInfo/288

PUBLICATION MISCONDUCT
https://www.wjgnet.com/bpg/gerinfo/208

ARTICLE PROCESSING CHARGE
https://www.wjgnet.com/bpg/gerinfo/242

STEPS FOR SUBMITTING MANUSCRIPTS
https://www.wjgnet.com/bpg/GerInfo/239

ONLINE SUBMISSION
https://www.f6publishing.com

World Journal of Gastrointestinal Oncology

Submit a Manuscript: https://www.f6publishing.com

World J Gastrointest Oncol 2024 June 15; 16(6): 2404-2418

DOI: 10.4251/wjgo.v16.i6.2404

ISSN 1948-5204 (online)

ORIGINAL ARTICLE

Retrospective Cohort Study

Unveiling the secrets of gastrointestinal mucous adenocarcinoma survival after surgery with artificial intelligence: A population-based study

Jie Song, Xiang-Xiu Yan, Fang-Liang Zhang, Yong-Yi Lei, Zi-Yin Ke, Fang Li, Kai Zhang, Yu-Qi He, Wei Li, Chao Li, Yuan-Ming Pan

Specialty type: Oncology

Provenance and peer review: Unsolicited article; Externally peer reviewed.

Peer-review model: Single blind

Peer-review report's classification

Scientific Quality: Grade B
Novelty: Grade B
Creativity or Innovation: Grade B
Scientific Significance: Grade B

P-Reviewer: Govindarajan KK, India

Received: December 29, 2023
Revised: February 27, 2024
Accepted: April 3, 2024
Published online: June 15, 2024

Jie Song, Xiang-Xiu Yan, Department of Gastroenterology, Dongying People's Hospital, Dongying Hospital of Shandong Provincial Hospital Group, Dongying 257000, Shandong Province, China

Fang-Liang Zhang, Gastrointestinal Surgery Department, Suining Central Hospital, Suining 629000, Sichuan Province, China

Yong-Yi Lei, Obstetrical Department, Suining Central Hospital, Suining 629000, Sichuan Province, China

Zi-Yin Ke, School of Medicine, Shenzhen Campus of Sun Yat-sen University, Shenzhen 518107, Guangdong Province, China

Fang Li, Department of Pathology, Aerospace Center Hospital, Peking University Aerospace School of Clinical Medicine, Beijing 100049, China

Kai Zhang, General Department, Beijing Chest Hospital, Capital Medical University/Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China

Yu-Qi He, Department of Gastroenterology, Beijing Chest Hospital, Capital Medical University/Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China

Wei Li, Department of Thoracic Surgery, Sichuan Academy of Medical Sciences and Sichuan Provincial People's Hospital, University of Electronic Science and Technology of China, Chengdu 610072, Sichuan Province, China

Chao Li, Department of Gastroenterology, Aerospace Center Hospital, Peking University Aerospace School of Clinical Medicine, Beijing 100049, China

Yuan-Ming Pan, Cancer Research Center, Beijing Chest Hospital, Capital Medical University/Beijing Tuberculosis and Thoracic Tumor Research Institute, Beijing 101149, China

Co-first authors: Jie Song and Xiang-Xiu Yan.

Co-corresponding authors: Yuan-Ming Pan and Chao Li.

Corresponding author: Yuan-Ming Pan, PhD, Cancer Research Center, Beijing Chest Hospital,

Capital Medical University/Beijing Tuberculosis and Thoracic Tumor Research Institute, No. 9 Beiguan Street, Tongzhou District, Beijing 101149, China. peterfpan2020@mail.ccmu.edu.cn

Abstract

BACKGROUND

Research on gastrointestinal mucosal adenocarcinoma (GMA) is limited and controversial, and there is no reference tool for predicting postoperative survival.

AIM

To investigate the prognosis of GMA and develop predictive model.

METHODS

From the Surveillance, Epidemiology, and End Results database, we collected clinical information on patients with GMA. After random sampling, the patients were divided into the discovery (70% of the total, for model training), validation (20%, for model evaluation), and completely blind test cohorts (10%, for further model evaluation). The main assessment metric was the area under the receiver operating characteristic curve (AUC). All collected clinical features were used for Cox proportional hazard regression analysis to determine factors influencing GMA's prognosis.

RESULTS

This model had an AUC of 0.7433 [95% confidence intervals (95%CI): 0.7424-0.7442] in the discovery cohort, 0.7244 (GMA: 0.7234-0.7254) in the validation cohort, and 0.7388 (95%CI: 0.7378-0.7398) in the test cohort. We packaged it into Windows software for doctors' use and uploaded it. Mucinous gastric adenocarcinoma had the worst prognosis, and these were protective factors of GMA: Regional nodes examined [hazard ratio (HR): 0.98, 95%CI: 0.97-0.98, $P < 0.001$)] and chemotherapy (HR: 0.62, 95%CI: 0.58-0.66, $P < 0.001$).

CONCLUSION

The deep learning-based tool developed can accurately predict the overall survival of patients with GMA postoperatively. Combining surgery, chemotherapy, and adequate lymph node dissection during surgery can improve patient outcomes.

Key Words: Deep learning; Gastrointestinal mucous adenocarcinoma; Overall survival; Surgery; Clinical tool

Core Tip: After surgery, some patients can be diagnosed with gastrointestinal mucous adenocarcinoma (GMA) by pathology, a rare subtype cancer. However, research on GMA is limited and controversial, and there is no reference tool for their postoperative survival prediction. We searched Surveillance, Epidemiology, and End Results database and collected 11390 GMA patients' clinical information. Then we constructed a deep learning-based tool to predict GMA patients' overall survival after surgery, and the tool has been uploaded. After our analysis, combining surgery, chemotherapy, and adequate lymph node dissection during surgery can improve patient outcomes.

Citation: Song J, Yan XX, Zhang FL, Lei YY, Ke ZY, Li F, Zhang K, He YQ, Li W, Li C, Pan YM. Unveiling the secrets of gastrointestinal mucous adenocarcinoma survival after surgery with artificial intelligence: A population-based study. *World J Gastrointest Oncol* 2024; 16(6): 2404-2418

URL: https://www.wjgnet.com/1948-5204/full/v16/i6/2404.htm

DOI: https://dx.doi.org/10.4251/wjgo.v16.i6.2404

INTRODUCTION

Gastrointestinal cancer is one of the most common fatal tumors in the United States, and colorectal cancer is the third most frequent malignant tumor and the third most deadly tumor[1,2]. Surgery is one of the most popular therapies[3,4]. However, after surgery, some patients can be pathologically diagnosed with gastrointestinal mucous adenocarcinoma (GMA), a rare subtype represented by mucinous gastric adenocarcinoma (MGA), mucinous duodenal adenocarcinoma (MDA), and mucinous colorectal adenocarcinoma (MCA). Figure 1 shows typical endoscopic and pathological images of the GMA, including the MGA, MDA, and MCA. To further identify GMA, immunohistochemistry is used frequently, and common antibody combinations include MUC-2, CK-20, CDX-2, and CK-7[5,6]. Taking the MCA as an example, MUC-2, CK-20, and CDX-2 were positive, whereas CK-7 was negative (Figure 1C).

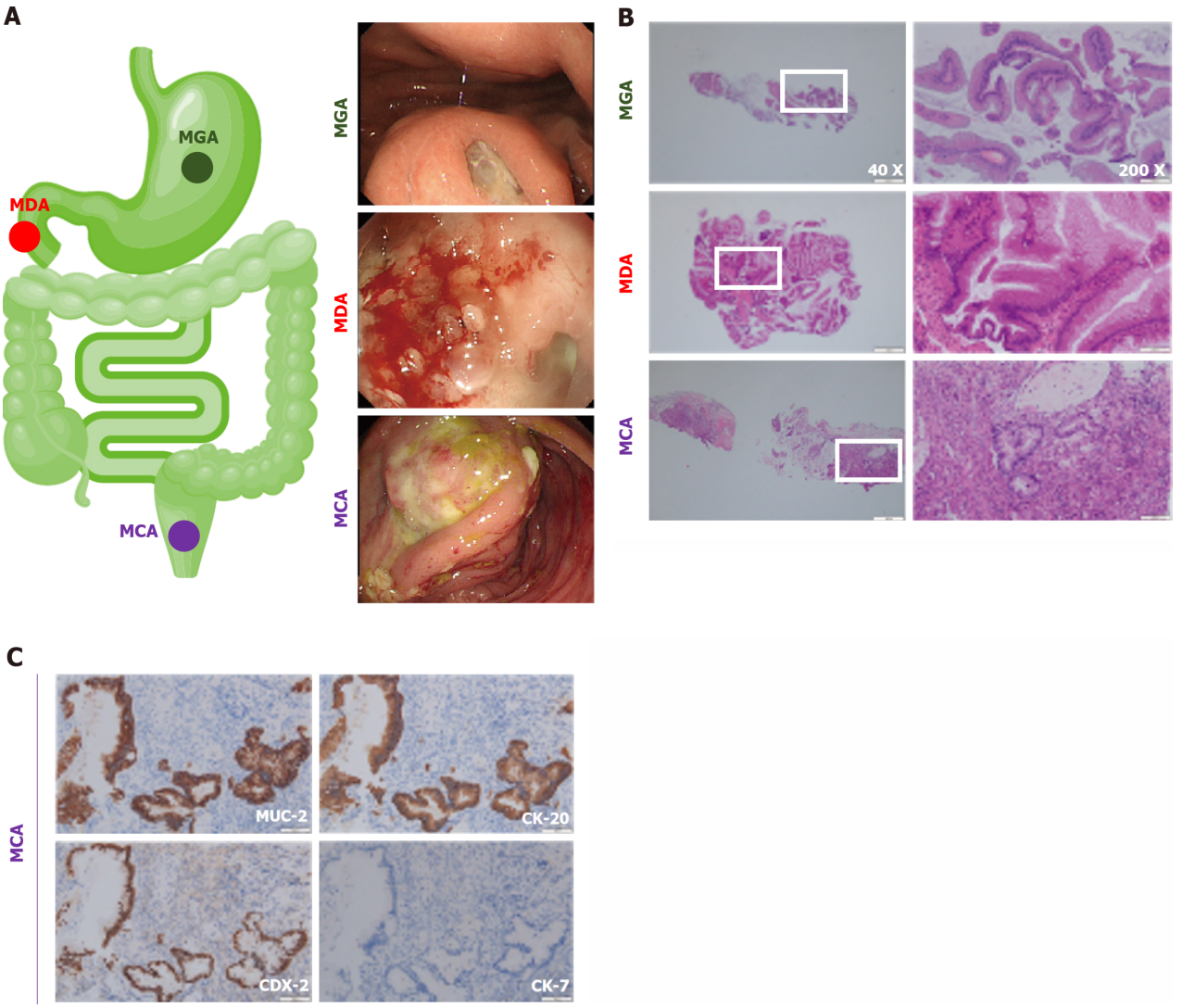


Figure 1 The endoscopic, hematoxylin-eosin staining, and immunohistochemistry photos of gastrointestinal mucous adenocarcinoma. A: The endoscopic photos of gastrointestinal mucous adenocarcinoma (GMA), including mucinous gastric adenocarcinoma (MGA), mucinous duodenal adenocarcinoma (MDA), and mucinous colorectal adenocarcinoma (MCA). MGA, ulcer lesions on the posterior wall of the gastric body, with white coating and jelly-like material on the surface. MDA, duodenal bulb ulcer bleeding with accumulation of transparent jelly-like material. MCA, the electronic colonoscope was inserted 90 cm through the anus. There was a raised lesion with a diameter of about 5 cm, and the surface was covered with yellow and white coating; B: Hematoxylin-eosin staining of GMA; C: Immunohistochemistry results of MCA, MUC-2, CK-20, and CDX-2 were positive while CK-7 was negative. MGA: Mucinous gastric adenocarcinoma; MDA: Mucinous duodenal adenocarcinoma; MCA: Mucinous colorectal adenocarcinoma.

Research on the GMA remains limited, and some conclusions from related studies are contradictory[7,8]. For example, there is conflicting information in the literature regarding the prognosis and overall survival (OS) of patients with MCA in the literature[7]. Consequently, awareness of GMA among doctors and researchers is limited, including some necessary expertise, such as a dearth of pertinent research to support additional preoperative or postoperative treatment for GMA. Large-scale clinical data analyses are required, particularly in randomized controlled clinical trials with high levels of evidence. The postoperative prognosis is another matter that concerns doctors, patients, and their families. Prognostic information currently available for GMA is scarce, especially because an individualized survival prediction system is lacking.

The Surveillance, Epidemiology, and End Results (SEER) database is the largest tumor database in the United States, with over 50 years of history. It covers approximately 48.0% of the American population. It is especially well-suited for studies on uncommon illnesses and cancer epidemiology surveys because of its wide coverage and authority.

In this study, we searched the SEER database, retrospectively analyzed the clinical data of patients with GMA using a large amount of clinical data, developed an OS prediction model for patients with GMA based on deep learning algorithms, and packaged it for simple usage by clinicians. In addition, we conducted statistical analyses and reviewed studies on the GMA to identify the risk and protective factors related to prognosis.

MATERIALS AND METHODS

Data source

We searched the SEER database and collected the clinical information of patients with GMA. Data originated from SEER Research Plus Data, 18 Registries, Nov 2020 Sub (2000-2018) sub-database, which covers approximately 27.8% of the American population. Detailed inclusion and exclusion criteria were as follows: (1) ICD-O-3 Hist/behav was 8480/3: Mucinous adenocarcinoma; (2) primary sites were gastrointestinal tract (C16.0-Cardia, NOS, C16.1-Fundus of stomach, C16.2-Body of stomach, C16.3-Gastric antrum, C16.4-Pylorus, C16.5-Lesser curvature of stomach NOS, C16.6-Greater curvature of stomach NOS, C16.8-Overlapping lesion of stomach, C16.9-Stomach, NOS, C17.0-Duodenum, C17.1-Jejunum, C17.2-Ileum, C17.8-Overlapping lesion of small intestine, C17.9-Small intestine, NOS, C18.0-Cecum, C18.1-Appendix, C18.2-Ascending colon, C18.3-Hepatic flexure of colon, C18.4-Transverse colon, C18.5-Splenic flexure of colon, C18.6-Descending colon, C18.7-Sigmoid colon, C18.8-Overlapping lesion of colon, C18.9-Colon, NOS, C19.9-Rectosigmoid junction or C20.9-Rectum, and NOS); (3) patients have gotten surgery; (4) complete American Joint Committee on Cancer TNM stage and other clinical features needed; and (5) no missing values (Table 1).

Study design

This retrospective study was designed for diagnostic testing. After screening according to the inclusion and exclusion criteria, all patients were randomly assigned to the discovery (70%), validation (20%), and test (10%) cohorts. The discovery cohort was used to train the deep learning survival model, which was evaluated in the validation cohort and another completely blind test cohort. The primary outcome was the OS of the patients with GMA (Figure 2).

The data for this research came from the publicly accessible SEER database, and patients' information was anonymized and untraceable. Consequently, this study was exempt from ethical approval and written permission.

Predictive variables

Age, sex, tumor site, history of malignant tumors, and TNM stage are potential risk factors for gastrointestinal cancer[7,9-11]. A larger tumor diameter or more positive lymph nodes generally indicates a more advanced tumor stage, and additional lymph node examinations can help determine this stage. Therefore, they are also considered conceivable predictors. Radiotherapy and chemotherapy are the most commonly used treatment strategies in addition to surgery.

The variables listed above were entered into the least absolute shrinkage and selection operator (LASSO) regression with 10-fold cross-validation to find the lowest lambda value. Clinical features with nonzero coefficients in the regression model were selected as final predictor variables based on this lambda value.

According to SEER rules, tumors with a diameter of 989 mm or larger are still recorded as 989 mm. Patients older than 100 years were listed as such. The tumor sites were merged according to their records. Those who survived for less than one month were still regarded as one month.

Model training, evaluation, and packaging

The training process was completed in python 3.9 (using Pytorch, Torchtuples, Sklearn, Pandas, Numpy, and Pycox). Unlike the typical classification, survival prediction has two variables: Survival time and status. This model was built based on DeepSuvr theory[12]. To obtain a better training effect, we transformed categorical clinical features (sex, malignant tumor history, tumor site, T, N, M, and stage, radiotherapy, and chemotherapy) to number labels (Supplementary Table 1). In contrast, numerical clinical features (age, tumor size, regional nodes positive, and regional nodes examined) were standardized (also known as the z-score, calculated by subtracting the population mean from an individual raw score and then dividing the difference by the population standard deviation; Supplementary Table 2). Batch training (using 2048 samples for training per epoch) was performed to obtain a better fit. A batch normalization layer and dropout layer were used to avoid overfitting. The Adam optimizer was adopted by setting the learning rate at 0.05. An early stopping function was used, which could terminate training automatically if the model had been trained for numerous rounds (setting 30); however, its performance improved slightly. The ultimate output of the model was a group of numbers (no bias). After sigmoid conversion, the values were between 0 and 1, the predicted survival probabilities for different months.

Model performance was evaluated using the area under the receiver operating characteristic curve (AUC). The closer the AUC is to 1.0, the better the model performance. The closer the AUC is to 0.5, the more inclined the model is toward random guessing. The bootstrap method was used to obtain the AUC and 95% confidence interval (CI). The model was truncated at 1, 3, and 5 years to obtain a more comprehensive assessment. Other evaluation metrics included specificity, sensitivity, accuracy, positive predictive value (PPV), and negative predictive value (NPV). We also used a Cox proportional hazards (CPH) model using the same clinical features for comparison.

Finally, the model was packaged into a Windows tool that doctors could use more conveniently. This process was completed in Pycharm, using the pyside6 and pyinstaller package.

Survival analysis

We compared the prognosis of GMA at different sites using Kaplan-Meier curves and log-rank tests. All collected clinical features were utilized to conduct multivariate CPH regression to identify the protective and risk factors for GMA. Some clinical features (T, N, M, and stage) were reintegrated before this process.

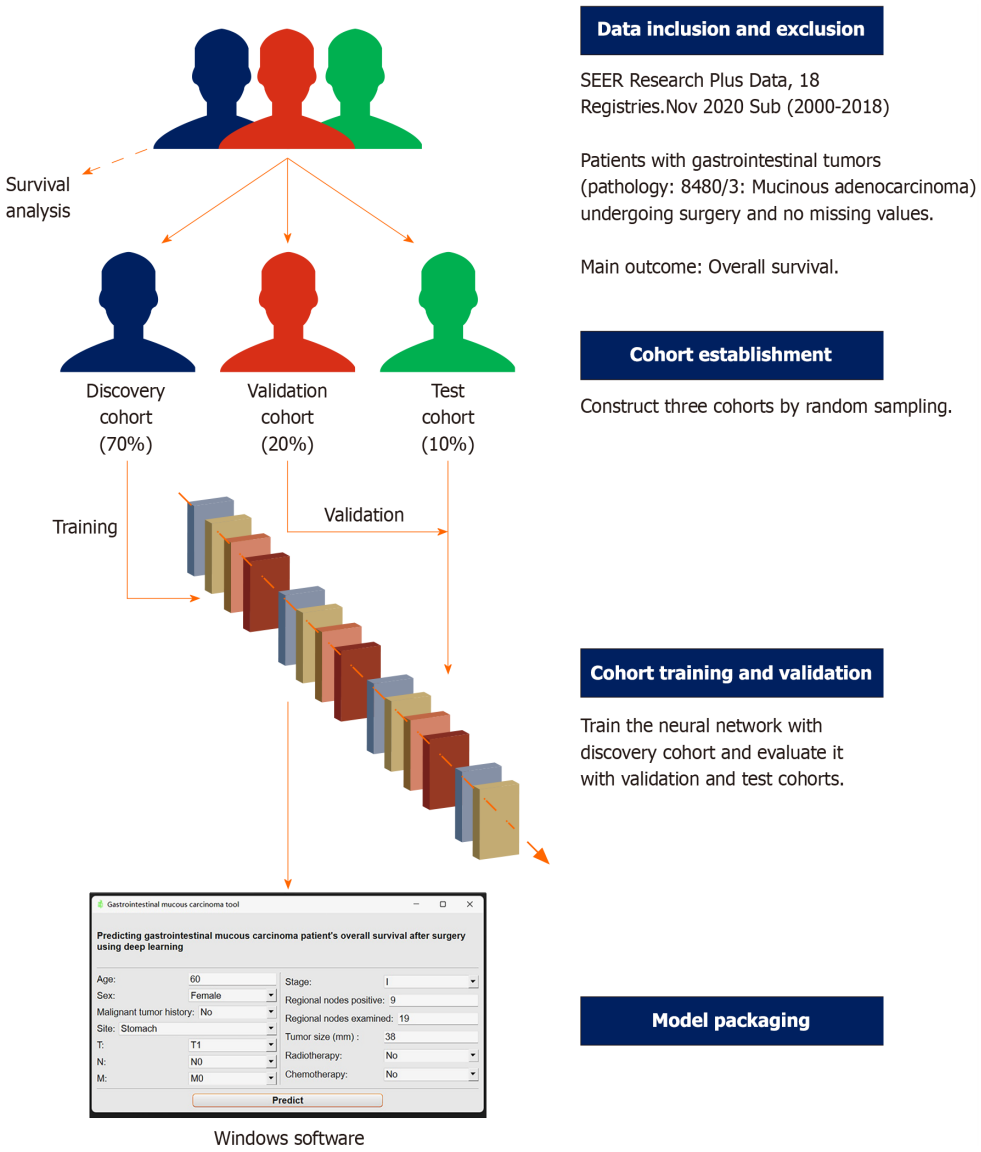


Figure 2 Flowchart for this research. SEER: Surveillance, Epidemiology, and End Results.

Statistical analysis

All statistical analyses were performed using R version 4.2.0. The Chi-square test was used for categorical variables, and the Kruskal-Wallis test was used for continuous variables with a non-normal distribution. A two-sided *P* value less than 0.05 was considered statistically significant. The following R packages were used for data analysis and visualization: glmnet, pROC, ggsci, ggplot2, survminer, survival, forest model, epiDisplay, circlize, and ggridge.

RESULTS

Demographic characteristics

Ultimately, 11390 patients were included in the study. They were then randomly assigned to one of the three cohorts. There were 7972 patients in the discovery cohort, 2392 in the validation cohort, and 1026 in the test cohort. There were no significant differences among the three cohorts. The median ages of the discovery and test cohorts were 69 years, whereas that of the validation cohort was 70 years. Among the three cohorts, most patients with GMA were female among three cohorts and had no history of malignant tumors. In the three cohorts, the most common GMA tumor sites were the other parts of the colon (not the rectum and rectosigmoid junction, or the cecum and appendix). Most patients were evaluated as T3, N0, or M0; therefore, most patients were staged as IIA. The median tumor size was 53.5 mm in the discovery cohort, 53.0 mm in the validation cohort, and 51.0 mm in the test cohort. The median number of positive regional nodes in the three cohorts was 0. The median number of regional nodes examined was 18 in the discovery and validation cohorts and 17 in the test cohort. Most patients did not receive radiotherapy or chemotherapy. The median survival time in the discovery and validation cohorts was 45 months, while test cohort patients had a 48-month median survival time. Almost 50% of the patients in the three groups were alive at the end of the follow-up period (Table 2).

Table 1 Data filtering condition

Option in SEER	Value
Database	SEER Research Plus Data, 18 registries. Nov 2020 Sub (2000-2018)
ICD-O-3 Hist/behav	8480/3: Mucinous adenocarcinoma
Primary site-labeled	C16.0-Cardia, NOS
	C16.1-Fundus of stomach
	C16.2-Body of stomach
	C16.3-Gastric antrum
	C16.4-Pylorus
	C16.5-Lesser curvature of stomach NOS
	C16.6-Greater curvature of stomach NOS
	C16.8-Overlapping lesion of stomach
	C16.9-Stomach, NOS
	C17.0-Duodenum
	C17.1-Jejunum
	C17.2-Ileum
	C17.8-Overlapping lesion of small intestine
	C17.9-Small intestine, NOS
	C18.0-Cecum
	C18.1-Appendix
	C18.2-Ascending colon
	C18.3-Hepatic flexure of colon
	C18.4-Transverse colon
	C18.5-Splenic flexure of colon
	C18.6-Descending colon
	C18.7-Sigmoid colon
	C18.8-Overlapping lesion of colon
	C18.9-Colon, NOS
	C19.9-Rectosigmoid junction
	C20.9-Rectum, NOS
Other	Receive surgery and records without missing value

SEER: The Surveillance, Epidemiology, and End Results; ICD-O-3: The International Classification of Diseases for Oncology, Third Edition; NOS: Not otherwise specified.

The characteristics of all three cohorts of patients were visually displayed in Figure 3, including categorical (Figure 3A) and numerical variables (Figure 3B). They described the sources and general distribution of GMA at different tumor sites.

Variable filtering process

LASSO Cox regression was used to filter the collected clinical features. After 10-fold cross-validation, the minimum lambda value was 0.0031 (Supplementary Figure 1A). The model's variable coefficients were examined with this lambda value, and none was equal to zero (Supplementary Table 3). This means that age, sex, malignant tumor history, tumor site, TNM stage, tumor size, regional lymph node positivity, regional lymph nodes examined, radiotherapy, and chemotherapy could predict the OS of patients with GMA. Therefore, they were all used in subsequent modeling.

Model training and performance

After 100 epochs, the early stopping function abruptly terminated training. The training curves are presented in Supplementary Figure 1B. Finally, the deep learning model had 14 layers. It included a linear layer (13 × 32), an activation layer

Table 2 Clinical features of patients, *n* (%)

	Discovery cohort (*n* = 7972)	Validation cohort (*n* = 2392)	Test cohort (*n* = 1026)	Statistical test method	*P* value
Age				Kruskal-Wallis	0.5211
Median (IQR)	69 (58, 79)	70 (59, 79)	69 (58, 80)		
Sex				Chi-square	0.4781
Female	4029 (50.54)	1219 (50.96)	539 (52.53)		
Male	3943 (49.46)	1173 (49.04)	487 (47.47)		
Malignant tumor history				Chi-square	0.2368
No	6217 (77.99)	1830 (76.51)	807 (78.65)		
Yes	1755 (22.01)	562 (23.49)	219 (21.35)		
Site				Chi-square	0.6989
Stomach	180 (2.26)	46 (1.92)	18 (1.75)		
Small intestine	105 (1.32)	37 (1.55)	11 (1.07)		
Cecum and appendix	2312 (29.00)	729 (30.48)	293 (28.56)		
Rectum and rectosigmoid junction	1003 (12.58)	295 (12.33)	129 (12.57)		
Other colon	4372 (54.84)	1285 (53.72)	575 (56.04)		
T				Chi-square	0.3557
T1	265 (3.32)	75 (3.14)	27 (2.63)		
T1a	2 (0.03)	2 (0.08)	0 (0.00)		
T1b	16 (0.20)	4 (0.17)	1 (0.10)		
T2	883 (11.08)	301 (12.58)	107 (10.43)		
T3	4505 (56.51)	1331 (55.64)	591 (57.60)		
T4	54 (0.68)	20 (0.84)	8 (0.78)		
T4a	1323 (16.60)	356 (14.88)	173 (16.86)		
T4b	924 (11.59)	303 (12.67)	119 (11.60)		
N				Chi-square	0.6443
N0	4123 (51.72)	1284 (53.68)	543 (52.92)		
N1	260 (3.26)	74 (3.09)	22 (2.14)		
N1a	832 (10.44)	235 (9.82)	96 (9.36)		
N1b	920 (11.54)	267 (11.16)	126 (12.28)		
N1c	137 (1.72)	29 (1.21)	16 (1.56)		
N2	104 (1.30)	31 (1.30)	13 (1.27)		
N2a	654 (8.20)	208 (8.70)	99 (9.65)		
N2b	898 (11.26)	255 (10.66)	107 (10.43)		
N3	11 (0.14)	4 (0.17)	2 (0.19)		
N3a	20 (0.25)	3 (0.13)	1 (0.10)		
N3b	13 (0.16)	2 (0.08)	1 (0.10)		
M				Chi-square	0.4620
M0	6689 (83.91)	2022 (84.53)	873 (85.09)		
M1	96 (1.20)	22 (0.92)	10 (0.97)		
M1a	563 (7.06)	162 (6.77)	79 (7.70)		

M1b	624 (7.83)	186 (7.78)	64 (6.24)		
Stage				Chi-square	0.8700
I	883 (11.08)	292 (12.21)	102 (9.94)		
IA	10 (0.13)	3 (0.13)	0 (0.00)		
IB	21 (0.26)	3 (0.13)	2 (0.19)		
II	3 (0.04)	1 (0.04)	0 (0.00)		
IIA	2236 (28.05)	675 (28.22)	305 (29.73)		
IIB	405 (5.08)	112 (4.68)	52 (5.07)		
IIC	303 (3.80)	111 (4.64)	46 (4.48)		
III	2 (0.03)	1 (0.04)	0 (0.00)		
IIIA	222 (2.78)	67 (2.80)	28 (2.73)		
IIIB	1742 (21.85)	506 (21.15)	227 (22.12)		
IIIC	862 (10.81)	251 (10.49)	111 (10.82)		
IV	96 (1.20)	22 (0.92)	10 (0.97)		
IVA	523 (6.56)	145 (6.06)	73 (7.12)		
IVB	603 (7.56)	188 (7.86)	63 (6.14)		
IVC	61 (0.77)	15 (0.63)	7 (0.68)		
Tumor size (mm)				Kruskal-Wallis	0.4812
Median (IQR)	53.5 (38.0, 72.0)	53.0 (38.0, 75.0)	51.0 (40.0, 70.0)	Median (IQR)	
Regional nodes positive				Kruskal-Wallis	0.4721
Median (IQR)	0 (0, 3)	0 (0, 3)	0 (0, 3)		
Regional nodes examined				Kruskal-Wallis	0.4691
Median (IQR)	18 (13, 24)	18 (13, 24)	17 (13, 23)		
Radiotherapy				Chi-square	0.4074
No	7231 (90.70)	2180 (91.14)	943 (91.91)		
Yes	741 (9.30)	212 (8.86)	83 (8.09)		
Chemotherapy				Chi-square	0.6545
No	4615 (57.89)	1378 (57.61)	608 (59.26)		
Yes	3357 (42.11)	1014 (42.39)	418 (40.74)		
Survival time				Kruskal-Wallis	0.2472
Median (IQR)	45 (20.00, 69.00)	45 (20.75, 69.00)	48 (21.00, 71.75)		
Status				Chi-square	0.8302
Alive	4034 (50.60)	1195 (49.96)	522 (50.88)		
Dead	3938 (49.40)	1197 (50.04)	504 (49.12)		

IQR: Interquartile range.

(ReLU), a batch normalization layer, a dropout layer (10%), a second linear layer (32 × 8), a second activation layer (ReLU), second batch normalization layer, second dropout layer (10%), third linear layer (8 × 4), third activation layer (ReLU), third batch normalization layer, third dropout layer (10%), fourth linear layer (4 × 1), and fourth activation layer (Sigmoid) (Figure 4A). The final output was a GMA patient's OS probability for the next 1-107 months. The model parameters are shown in Supplementary Figure 2.

This model had a 0.7433 (95%CI: 0.7424-0.7442) AUC in the discovery cohort, 0.7244 (95%CI: 0.7234-0.7254) AUC in the validation cohort, 0.7388 (95%CI: 0.7378-0.7398) AUC in the test cohort (Table 3). The receiver operating characteristic curves are shown in Figure 4B.

In comparison, the same variables and the CPH method were used to fit the data. It only had an AUC of 0.7155 (95%CI: 0.7145-0.7166) in the discovery cohort, 0.6942 (95%CI: 0.6932-0.6953) in the validation cohort, and 0.7178 (95%CI: 0.7168-0.7188) in the test cohort (Supplementary Table 4). Regardless of the mean or 95%CI of the AUC, it is evident that the

Table 3 Overall performance of the model

	AUC	AUC, 95%CI	
	Mean	Low	High
Discovery cohort	0.7433	0.7424	0.7442
Validation cohort	0.7244	0.7234	0.7254
Test cohort	0.7388	0.7378	0.7398

AUC: Area under receiver operating characteristic curve; CI: Confidence interval.

deep-learning-based model performs better than the CPH.

After 1, 3, and 5 years, we thoroughly assessed the performance of the deep-learning-based model. In discovery cohort, it had 0.7953 AUC (95%CI: 0.7817-0.8090), 0.7688 specificity, 0.6742 sensitivity, 0.7536 accuracy, 0.9250 NPV and 0.3581 PPV in 1-year OS prediction; 0.8034 AUC (95%CI: 0.7933-0.8136), 0.7675 specificity, 0.6963 sensitivity, 0.7421 accuracy, 0.8200 NPV and 0.6243 PPV in the 3-year OS prediction; 0.7971 AUC (95%CI: 0.7873-0.8069), 0.7985 specificity, 0.6595 sensitivity, 0.7365 accuracy, 0.7441 NPV and 0.7253 PPV in 5-year OS prediction. The validation cohort showed 0.7757 AUC (95%CI: 0.7501-0.8012), 0.6493 specificity, 0.7775 sensitivity, 0.6697 accuracy, 0.9388 NPV, and 0.2964 PPV in the 1-year OS prediction; 0.7843 AUC (95%CI: 0.7650-0.8036), 0.7260 specificity, 0.7242 sensitivity, 0.7253 accuracy, 0.8234 NPV and 0.5987 PPV in the 3-year OS prediction; 0.7772 AUC (95%CI: 0.7587-0.7958), 0.7586 specificity, 0.6746 sensitivity, 0.7203 accuracy, 0.7355 NPV and 0.7010 PPV in the 5-year OS prediction. The test cohort showed 0.7938 AUC (95%CI: 0.7566-0.8310), 0.7182 specificity, 0.7386 sensitivity, 0.7212 accuracy, 0.9400 NPV and 0.3148 PPV in 1-year OS prediction; 0.7888 AUC (95%CI: 0.7603-0.8173), 0.6869 specificity, 0.7507 sensitivity, 0.7086 accuracy, 0.8424 NPV and 0.5527 PPV in 3-year OS prediction; 0.7871 AUC (95%CI: 0.7597-0.8146), 0.7296 specificity, 0.7127 sensitivity, 0.7222 accuracy, 0.7655 NPV and 0.6723 PPV in 5-year OS prediction (Table 4).

Model packaging and usage

For convenience, we packaged the model into Windows software. After unzipping, users can double-click *Main.exe* to start. After inputting the GMA patient's clinical characteristics, click *Predict* to run the built-in pre-trained neural network. After the calculation, the prediction results were automatically drawn into a survival curve (Kaplan-Meier curve). The horizontal axis represents a certain month, and the vertical axis represents the OS probability that the predicted patient is still alive in that month. The curve can be zoomed in or out using the mouse, and a specific value is displayed when hovering (Figure 4C).

Survival analysis

Overall, the incidence rate of GMA is declining, about 1.7% (1.9% in male and 1.5% in female) (Supplementary Table 5). Moreover, the 1-year survival rate of patients with GMA is about 84% (95%CI: 83%-85%), the 3-year survival rate of them is about 64% (95%CI: 63%-65%) and the 5-year survival rate of them is about 53% (95%CI: 52%-54%) (Supplementary Table 6).

Survival analysis showed that patients with GMA in the stomach had the worst prognosis ($P < 0.0001$) (Figure 5A). Multivariate CPH regression displayed that these clinical features were risk factors: older age [hazard ratio (HR): 1.03; 95%CI: 1.03-1.03, $P < 0.001$), male (HR: 1.09, 95%CI: 1.03-1.15, $P = 0.002$), malignant tumor history (HR: 1.22, 95%CI: 1.14-1.29, $P < 0.001$), rectum and rectosigmoid junction (HR: 1.25, 95%CI: 1.12-1.38, $P < 0.001$), small intestine (HR: 1.41, 95%CI: 1.13-1.75, $P < 0.002$), stomach (HR: 1.66, 95%CI: 1.36-2.02, $P < 0.001$), other colon sites (HR: 1.18, 95%CI: 1.11-1.25, $P < 0.001$), T3 (HR: 1.59, 95%CI: 1.26-2.02, $P < 0.001$), T4 (HR: 2.47, 95%CI: 1.94-3.13, $P < 0.001$), N1 (HR: 1.70, 95%CI: 1.48-1.96, $P < 0.001$), N2 (HR: 2.02, 95%CI: 1.74-2.35, $P < 0.001$), N3 (HR: 1.60, 95%CI: 1.07-2.39, $P = 0.021$), M1 (HR: 2.47, 95%CI: 1.96-3.11, $P < 0.001$), larger tumor size (HR: 1.00, 95%CI: 1.00-1.00, $P < 0.001$), regional nodes positive (HR: 1.05, 95%CI: 1.05-1.06, $P < 0.001$). These were protective factors: Regional nodes examined (HR: 0.98, 95%CI: 0.97-0.98, $P < 0.001$) and chemotherapy (HR: 0.62, 95%CI:0.58-0.66, $P < 0.001$) (Figure 5B).

DISCUSSION

After surgical resection, some patients may be pathologically diagnosed with GMA, with approximately 1%-20% in the colorectum and 7% in the stomach[13-16]. GMA is distinguished by the presence of many mucinous components that account for approximately 50% of the tumor volume[7,17]. More mucinous components may indicate a poor prognosis [18]. Several factors, including younger age, advanced tumor stage, female sex, microsatellite instability (MSI), and molecular mutations (such as KRAS and BRAF), have been linked to the development of GMA according to earlier investigations[15,16,18-20]. It is still debatable whether GMA and common gastrointestinal tumors have similar OS, as previous studies have produced conflicting reports[7,21]. For example, Warschkow *et al*[22] observed that MCA had a similar prognosis to other colorectal cancers. Hugen argued that stage III mucinous rectal adenocarcinoma instead of MCA had a worse prognosis. However, more studies, especially a retrospective analysis with a larger sample size (222256 patients),

Table 4 Model's performance at 1-, 3- and 5-years' overall survival prediction

	Discovery cohort			Validation cohort			Test cohort		
	1 yr	3 yr	5 yr	1 yr	3 yr	5 yr	1 yr	3 yr	5 yr
AUC	0.7953	0.8034	0.7971	0.7757	0.7843	0.7772	0.7938	0.7888	0.7871
AUC, 95%CI	0.7817-0.8090	0.7933-0.8136	0.7873-0.8069	0.7501-0.8012	0.7650-0.8036	0.7587-0.7958	0.7566-0.8310	0.7603-0.8173	0.7597-0.8146
Specificity	0.7688	0.7675	0.7985	0.6493	0.7260	0.7586	0.7182	0.6869	0.7296
Sensitivity	0.6742	0.6963	0.6595	0.7775	0.7242	0.6746	0.7386	0.7507	0.7127
Accuracy	0.7536	0.7421	0.7365	0.6697	0.7253	0.7203	0.7212	0.7086	0.7222
NPV	0.9250	0.8200	0.7441	0.9388	0.8234	0.7355	0.9400	0.8424	0.7655
PPV	0.3581	0.6243	0.7253	0.2964	0.5987	0.7010	0.3148	0.5527	0.6723

AUC: Area under receiver operating characteristic curve; CI: Confidence interval; NPV: Negative predictive value; PPV: Positive predictive value.

demonstrated that MCA increased mortality risk by 2%-8%[16,22-25]. Similarly, Rokutan *et al*[15] noticed that MGA was related to poor outcomes, but Hsu stated the opposite conclusion[26]. Most studies have reported that the prognosis of GMA is poor, although additional research and attention are needed.

Since GMA is mainly diagnosed during postoperative pathological examination and there is currently no effective prognostic model, we searched the SEER database and constructed a deep learning algorithm. In the medical field, classic survival prediction is based on the CPH. However, the biggest shortcoming of this theory is that it assumes that the impact of covariates on survival is linear. Although it is simple and easy to implement, this ideal assumption is unsuitable for the intricate changes in the real world. Machine learning, especially deep learning, has been gradually applied in medicine in recent years, including clinical data, medical imaging data, pathological slides, and genomics[27,28]. Due to the presence of a time series and the surviving state, survival prediction is neither a typical classification nor a regression problem. Katzman *et al*[12] proposed the DeepSurv method to solve this problem, which has been applied in some tumor prognosis studies, such as lung cancer and head and neck cancer[12,29,30]. Our previous study also demonstrated that it is better than traditional algorithms such as CPH[31]. Therefore, we built a deep-learning-based model based on the DeepSurv algorithm.

In this study, we collected the clinical data of 11390 patients. We divided them into three cohorts (7972 patients in the discovery cohort for model training, 2392 in the validation cohort, and 1026 in the test cohort for model evaluation) to predict the OS of patients with GMA after surgery. This model had a 0.7433 (95%CI: 0.7424-0.7442) AUC in the discovery cohort, 0.7244 (95%CI: 0.7234-0.7254) AUC in the validation cohort, 0.7388 (95%CI: 0.7378-0.7398) AUC in the test cohort, which showed predictive value to prognosis and was packaged into a Windows tool. Multivariate survival analysis revealed that chemotherapy and more regional lymph nodes examined were protective factors for GMA, which means that clinicians should consider a combination therapy of surgery and chemotherapy and perform adequate lymph node dissection during surgery.

According to previous studies, the diagnosis of GMA, including MCA and MGA[7,8]. This is consistent with the findings of this study. The median tumor size was approximately 51-53 mm, similar to previous report[26]. Stage II-III was most common when GMA was diagnosed in both our and previous studies, but Hsu and Rokutan found that most GMA had lymph node metastasis, which was different from ours[15,26]. Chemotherapy is a protective factor against GMA, which is consistent with the findings of previous studies[18,32]. Although Reynolds *et al*[20] thought that patients with GMA might respond poorly to chemoradiotherapy, they believed patients could benefit from regimens containing fluorouracil. We found that older age was associated with a worse prognosis, which was not significant in the study of Yan *et al*[18]. The potential reason may be that they only included limited younger patients (41 patients < 60 years but 371 patients ≥ 60 years). We observed that an advanced N stage was related to a worse prognosis, as mentioned in a previous study[18].

Currently, no GMA patient survival model is available. Existing prognostic models mostly focus on common pathological subtypes of gastrointestinal tumors, including the use of clinical features and genomics-based bioinformatics analyses, *etc*[33-36]. The deep learning-based tool developed in this study focused on GMA patient prognosis after surgery, filled the gaps in related fields, and may help assist in clinical decision-making.

Some researchers have reported that new targeted drugs for GMA are in progress[7,37,38]. Simultaneously, as GMA usually has a higher MSI, immunotherapy may bring better efficacy to GMA[7]. These factors are expected to enhance GMA the prognosis of patients with GMA.

This study had some limitations. It has been observed that some gene mutations are related to GMA prognosis, which was not considered at this time[15]. Besides, other potential factors like co-morbidity, immunohistochemistry, family history/genetic syndromes, and type of surgery (open/min access) may have potential influence on GMA, but not recorded in SEER database. This retrospective study inevitably has selection bias and information bias. And more Asian data and prospective data can validate our model better. Subsequent researchers may consider further improvements in the above areas.

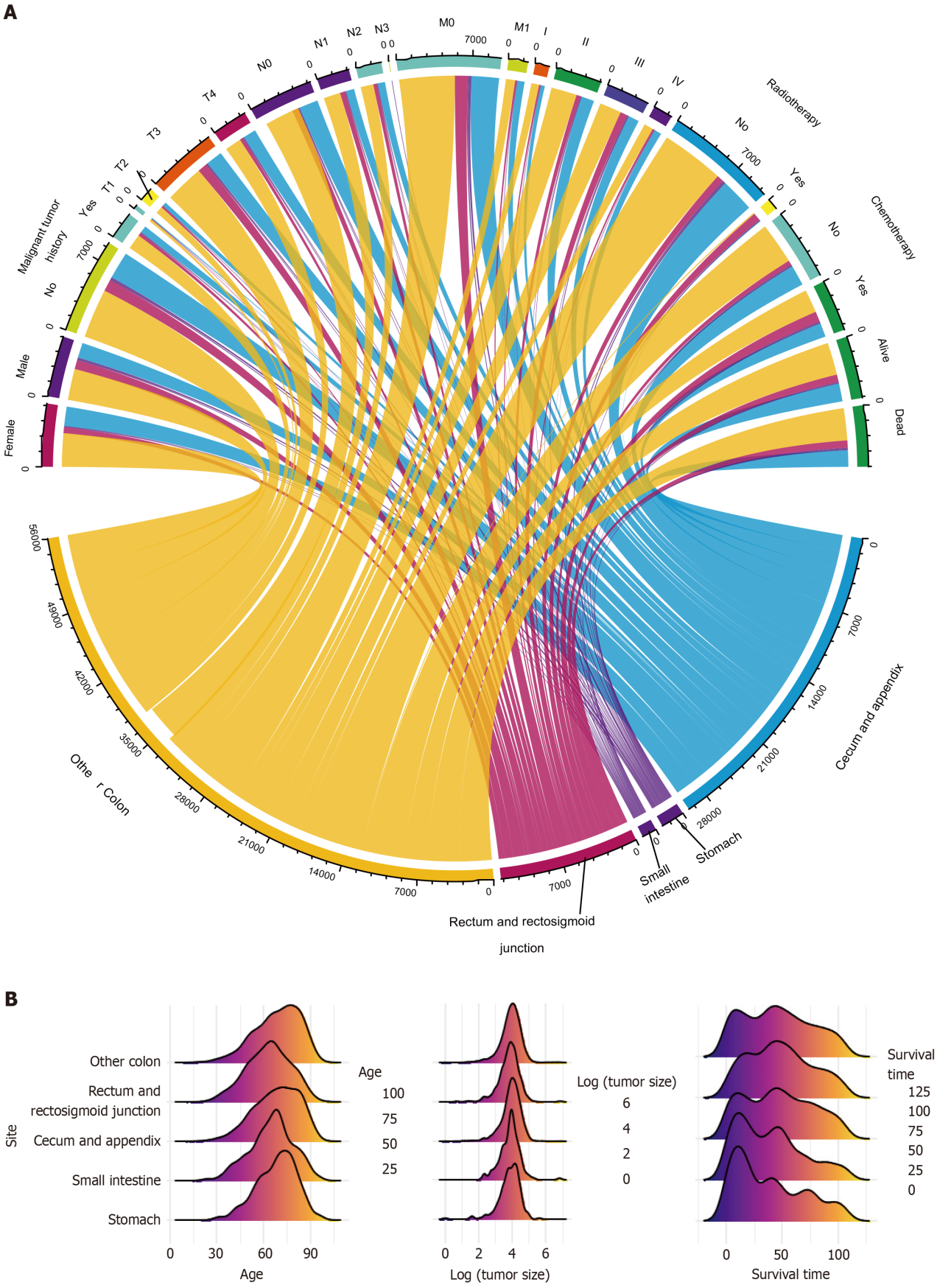


Figure 3 Visual presentation of collected patients' data. A: The flow of categorical variables; B: The distribution of numerical variables.

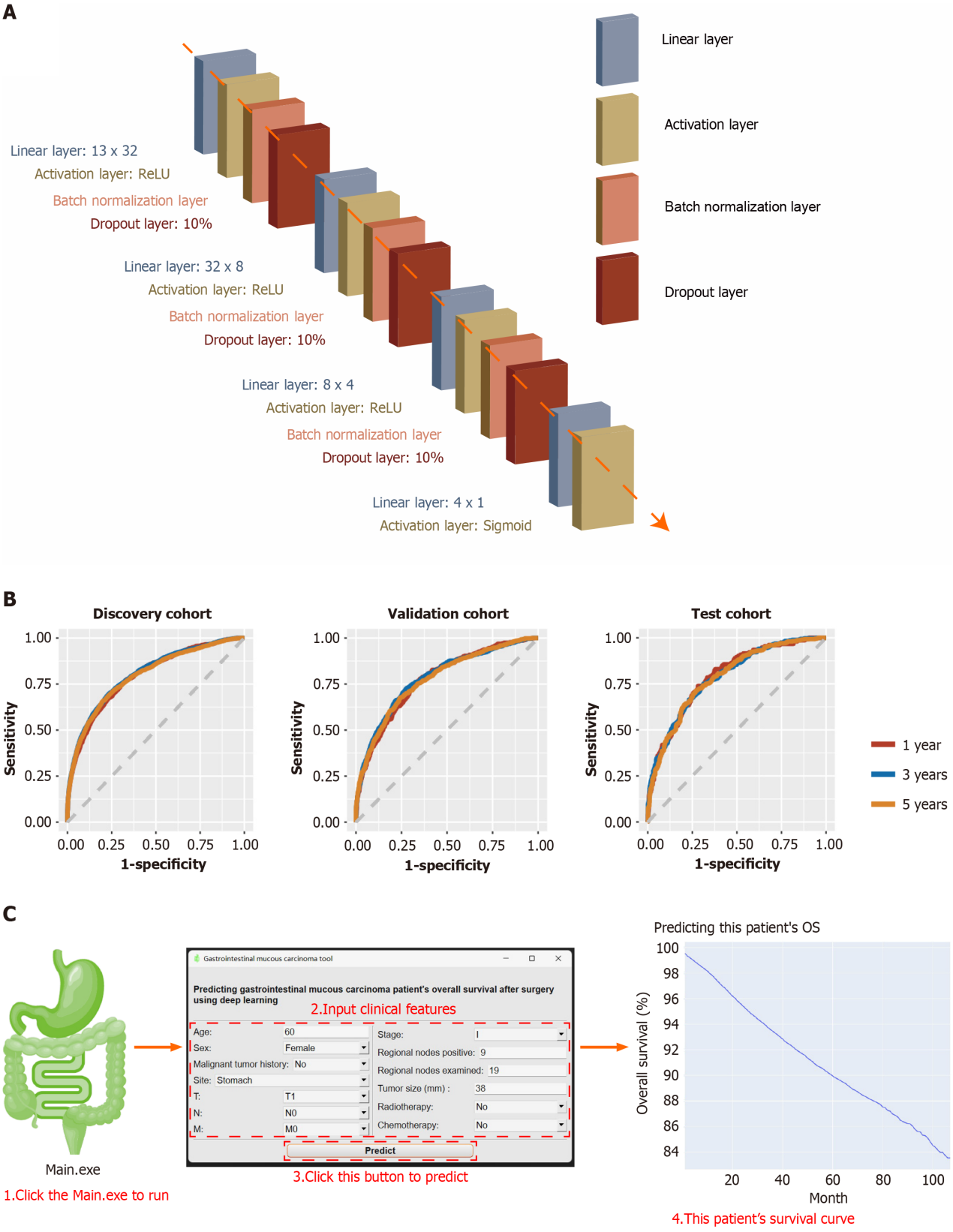
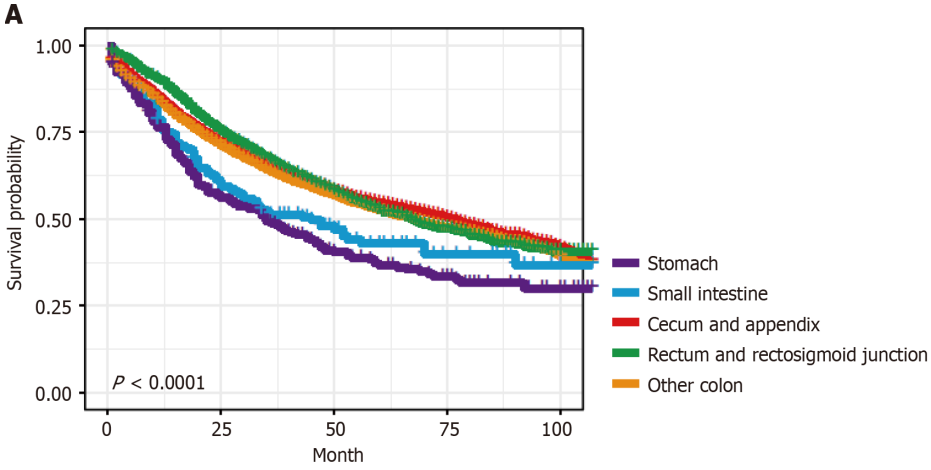


Figure 4 Deep learning-based tool to predict gastrointestinal mucous carcinoma patients' overall survival after surgery. A: The structure of neural network; B: The receiver operating characteristic curves of it; C: The instruction for its use. OS: Overall survival.

CONCLUSION

The deep learning-based tool developed in this study can accurately predict the OS of patients with gastrointestinal mucous carcinoma after surgery. Combining surgery, chemotherapy, and adequate lymph node dissection during surgery can improve patient outcomes.

A



B

		N	Hazard ratio	*P* value
Age		11390	1.03 (1.03, 1.03)	< 0.001[c]
Sex	Female	5787	Reference	
	Male	5603	1.09 (1.03, 1.15)	0.002[b]
Malignant tumor history	No	8854	Reference	
	Yes	2536	1.22 (1.14, 1.29)	< 0.001[c]
Site	Cecum and appendix	3334	Reference	
	Other colon	6232	1.18 (1.11, 1.25)	< 0.001[c]
	Rectum and rectosigmoid junction	1427	1.25 (1.12, 1.38)	< 0.001[c]
	Small intestine	153	1.41 (1.13, 1.75)	0.002[b]
	Stomach	244	1.66 (1.36, 2.02)	< 0.001[c]
T	T1	392	Reference	
	T2	1291	1.13 (0.92, 1.38)	0.232
	T3	6427	1.59 (1.26, 2.02)	< 0.001[c]
	T4	3280	2.47 (1.94, 3.13)	< 0.001[c]
N	N0	5950	Reference	
	N1	3014	1.70 (1.48, 1.96)	< 0.001[c]
	N2	2369	2.02 (1.74, 2.35)	< 0.001[c]
	N3	57	1.60 (1.07, 2.39)	0.021[a]
M	M0	9584	Reference	
	M1	1806	2.47 (1.96, 3.11)	< 0.001[c]
Stage	I	1316	Reference	
	II	4249	0.92 (0.76, 1.13)	0.426
	III	4019	0.83 (0.65, 1.04)	0.110
	IV	1806	Insufficient sample	
Tumor size		11390	1.00 (1.00, 1.00)	< 0.001[c]
Regional nodes positive		11390	1.05 (1.05, 1.06)	< 0.001[c]
Regional nodes examined		11390	0.98 (0.97, 0.98)	< 0.001[c]
Radiotherapy	No	10354	Reference	
	Yes	1036	1.04 (0.93, 1.16)	0.530
Chemotherapy	No	6601	Reference	
	Yes	4789	0.62 (0.58, 0.66)	< 0.001[c]

Likelihood ratio test = 3940 on 21 degree of freedom. Model *P* < 2.2e-16.
n = 11390, number of events = 5639

Figure 5 Survival analysis of gastrointestinal mucous adenocarcinoma patients. A: Kaplan-Meier curve was used to compare the prognosis of different sites' gastrointestinal mucous adenocarcinoma (GMA); B: The protective and risk factors of GMA, revealed by Cox proportional hazard regression. [a]*P* < 0.05, [b]*P* < 0.01, [c]*P* < 0.001.

ACKNOWLEDGEMENTS

The authors thank the Aerospace Center Hospital, Peking University Aerospace School of Clinical Medicine, for providing endoscopic and pathological photographs.

FOOTNOTES

Author contributions: Song J and Yan XX contributed equally to this study; Song J and Yan XX designed the research study; Li W, Zhang FL, Lei YY, Ke ZY, and Zhang K collected the data; Li F, Zhang K, and He YQ performed statistical analysis; Li F and Li C checked and interpreted endoscopic image; Pan YM and Li W designed the prediction tool; Song J and Yan XX wrote the manuscript; and all authors have read and approve the final manuscript.

Institutional review board statement: This study was reviewed and approved by the Ethics Committee of Beijing Chest Hospital affiliated to Capital Medical University (Approval No. LW-2024-004).

Informed consent statement: All study participants, or their legal guardian, provided informed written consent prior to study enrollment.

Conflict-of-interest statement: There are no conflicts of interest to report.

Data sharing statement: The code used in this study is available from the corresponding author upon request. The deep learning-based tool used to predict the OS of patients with gastrointestinal mucous carcinoma after surgery is also available from the corresponding author. The raw data are saved in Supplementary Table 5.

STROBE statement: The authors have read the STROBE Statement—checklist of items, and the manuscript was prepared and revised according to the STROBE Statement—checklist of items.

Country/Territory of origin: China

ORCID number: Wei Li 0000-0003-0471-2382.

S-Editor: Chen YL
L-Editor: A
P-Editor: Zhao YQ

REFERENCES

1 **Tong Y**, Gao H, Qi Q, Liu X, Li J, Gao J, Li P, Wang Y, Du L, Wang C. High fat diet, gut microbiome and gastrointestinal cancer. *Theranostics* 2021; **11**: 5889-5910 [PMID: 33897888 DOI: 10.7150/thno.56157]
2 **Siegel RL**, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin* 2023; **73**: 17-48 [PMID: 36633525 DOI: 10.3322/caac.21763]
3 **Kuipers EJ**, Grady WM, Lieberman D, Seufferlein T, Sung JJ, Boelens PG, van de Velde CJ, Watanabe T. Colorectal cancer. *Nat Rev Dis Primers* 2015; **1**: 15065 [PMID: 27189416 DOI: 10.1038/nrdp.2015.65]
4 **Machlowska J**, Baj J, Sitarz M, Maciejewski R, Sitarz R. Gastric Cancer: Epidemiology, Risk Factors, Classification, Genomic Characteristics and Treatment Strategies. *Int J Mol Sci* 2020; **21** [PMID: 32512697 DOI: 10.3390/ijms21114012]
5 **Chu PG**, Chung L, Weiss LM, Lau SK. Determining the site of origin of mucinous adenocarcinoma: an immunohistochemical study of 175 cases. *Am J Surg Pathol* 2011; **35**: 1830-1836 [PMID: 21881489 DOI: 10.1097/PAS.0b013e3182299c25]
6 **Shin JH**, Bae JH, Lee A, Jung CK, Yim HW, Park JS, Lee KY. CK7, CK20, CDX2 and MUC2 Immunohistochemical staining used to distinguish metastatic colorectal carcinoma involving ovary from primary ovarian mucinous adenocarcinoma. *Jpn J Clin Oncol* 2010; **40**: 208-213 [PMID: 19926591 DOI: 10.1093/jjco/hyp150]
7 **Luo C**, Cen S, Ding G, Wu W. Mucinous colorectal adenocarcinoma: clinical pathology and treatment options. *Cancer Commun (Lond)* 2019; **39**: 13 [PMID: 30922401 DOI: 10.1186/s40880-019-0361-0]
8 **Meng NL**, Wang YK, Wang HL, Zhou JL, Wang SN. Research on the Histological Features and Pathological Types of Gastric Adenocarcinoma With Mucinous Differentiation. *Front Med (Lausanne)* 2022; **9**: 829702 [PMID: 35308509 DOI: 10.3389/fmed.2022.829702]
9 **Zuo D**, Li C, Liu T, Yue M, Zhang J, Ning G. Construction and validation of a metabolic risk model predicting prognosis of colon cancer. *Sci Rep* 2021; **11**: 6837 [PMID: 33767290 DOI: 10.1038/s41598-021-86286-z]
10 **Li X**, Wen D, Li X, Yao C, Chong W, Chen H. Identification of an Immune Signature Predicting Prognosis Risk and Lymphocyte Infiltration in Colon Cancer. *Front Immunol* 2020; **11**: 1678 [PMID: 33013820 DOI: 10.3389/fimmu.2020.01678]
11 **Xiaobin C**, Zhaojun X, Tao L, Tianzeng D, Xuemei H, Fan Z, Chunyin H, Jianqiang H, Chen L. Analysis of Related Risk Factors and Prognostic Factors of Gastric Cancer with Bone Metastasis: A SEER-Based Study. *J Immunol Res* 2022; **2022**: 3251051 [PMID: 35211630 DOI: 10.1155/2022/3251051]
12 **Katzman JL**, Shaham U, Cloninger A, Bates J, Jiang T, Kluger Y. DeepSurv: personalized treatment recommender system using a Cox proportional hazards deep neural network. *BMC Med Res Methodol* 2018; **18**: 24 [PMID: 29482517 DOI: 10.1186/s12874-018-0482-1]
13 **Glasgow SC**, Yu J, Carvalho LP, Shannon WD, Fleshman JW, McLeod HL. Unfavourable expression of pharmacologic markers in mucinous colorectal cancer. *Br J Cancer* 2005; **92**: 259-264 [PMID: 15655543 DOI: 10.1038/sj.bjc.6602330]
14 **Leopoldo S**, Lorena B, Cinzia A, Gabriella DC, Angela Luciana B, Renato C, Antonio M, Carlo S, Cristina P, Stefano C, Maurizio T, Luigi R,

Cesare B. Two subtypes of mucinous adenocarcinoma of the colorectum: clinicopathological and genetic features. *Ann Surg Oncol* 2008; **15**: 1429-1439 [PMID: 18301950 DOI: 10.1245/s10434-007-9757-1]

15 **Rokutan H**, Hosoda F, Hama N, Nakamura H, Totoki Y, Furukawa E, Arakawa E, Ohashi S, Urushidate T, Satoh H, Shimizu H, Igarashi K, Yachida S, Katai H, Taniguchi H, Fukayama M, Shibata T. Comprehensive mutation profiling of mucinous gastric carcinoma. *J Pathol* 2016; **240**: 137-148 [PMID: 27313181 DOI: 10.1002/path.4761]

16 **Verhulst J**, Ferdinande L, Demetter P, Ceelen W. Mucinous subtype as prognostic factor in colorectal cancer: a systematic review and meta-analysis. *J Clin Pathol* 2012; **65**: 381-388 [PMID: 22259177 DOI: 10.1136/jclinpath-2011-200340]

17 **Nagtegaal ID**, Odze RD, Klimstra D, Paradis V, Rugge M, Schirmacher P, Washington KM, Carneiro F, Cree IA; WHO Classification of Tumours Editorial Board. The 2019 WHO classification of tumours of the digestive system. *Histopathology* 2020; **76**: 182-188 [PMID: 31433515 DOI: 10.1111/his.13975]

18 **Yan C**, Yang H, Chen L, Liu R, Shang W, Yuan W, Yang F, Sun Q, Xia L. Clinical significance of mucinous component in colorectal adenocarcinoma: a propensity score-matched study. *BMC Cancer* 2021; **21**: 1286 [PMID: 34852768 DOI: 10.1186/s12885-021-09031-9]

19 **Chew MH**, Yeo SA, Ng ZP, Lim KH, Koh PK, Ng KH, Eu KW. Critical analysis of mucin and signet ring cell as prognostic factors in an Asian population of 2,764 sporadic colorectal cancers. *Int J Colorectal Dis* 2010; **25**: 1221-1229 [PMID: 20686777 DOI: 10.1007/s00384-010-1033-3]

20 **Reynolds IS**, Furney SJ, Kay EW, McNamara DA, Prehn JHM, Burke JP. Meta-analysis of the molecular associations of mucinous colorectal cancer. *Br J Surg* 2019; **106**: 682-691 [PMID: 30945755 DOI: 10.1002/bjs.11142]

21 **Kanda M**, Oba K, Aoyama T, Kashiwabara K, Mayanagi S, Maeda H, Honda M, Hamada C, Sadahiro S, Sakamoto J, Saji S, Yoshikawa T; Japanese Foundation for Multidisciplinary Treatment of Cancer. Clinical Signatures of Mucinous and Poorly Differentiated Subtypes of Colorectal Adenocarcinomas by a Propensity Score Analysis of an Independent Patient Database from Three Phase III Trials. *Dis Colon Rectum* 2018; **61**: 461-471 [PMID: 29521828 DOI: 10.1097/DCR.0000000000001022]

22 **Warschkow R**, Tarantino I, Huttner FJ, Schmied BM, Guller U, Diener MK, Ulrich A. Predictive value of mucinous histology in colon cancer: a population-based, propensity score matched analysis. *Br J Cancer* 2016; **114**: 1027-1032 [PMID: 26977857 DOI: 10.1038/bjc.2016.57]

23 **Hugen N**, Verhoeven RH, Radema SA, de Hingh IH, Pruijt JF, Nagtegaal ID, Lemmens VE, de Wilt JH. Prognosis and value of adjuvant chemotherapy in stage III mucinous colorectal carcinoma. *Ann Oncol* 2013; **24**: 2819-2824 [PMID: 24057984 DOI: 10.1093/annonc/mdt378]

24 **Mekenkamp LJ**, Heesterbeek KJ, Koopman M, Tol J, Teerenstra S, Venderbosch S, Punt CJ, Nagtegaal ID. Mucinous adenocarcinomas: poor prognosis in metastatic colorectal cancer. *Eur J Cancer* 2012; **48**: 501-509 [PMID: 22226571 DOI: 10.1016/j.ejca.2011.12.004]

25 **Kim SH**, Shin SJ, Lee KY, Kim H, Kim TI, Kang DR, Hur H, Min BS, Kim NK, Chung HC, Roh JK, Ahn JB. Prognostic value of mucinous histology depends on microsatellite instability status in patients with stage III colon cancer treated with adjuvant FOLFOX chemotherapy: a retrospective cohort study. *Ann Surg Oncol* 2013; **20**: 3407-3413 [PMID: 23943026 DOI: 10.1245/s10434-013-3169-1]

26 **Hsu JT**, Wang CW, Le PH, Wu RC, Chen TH, Chiang KC, Lin CJ, Yeh TS. Clinicopathological characteristics and outcomes in stage I-III mucinous gastric adenocarcinoma: a retrospective study at a single medical center. *World J Surg Oncol* 2016; **14**: 123 [PMID: 27112436 DOI: 10.1186/s12957-016-0886-5]

27 **Kang M**, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform* 2022; **23** [PMID: 34791014 DOI: 10.1093/bib/bbab454]

28 **Jiang Y**, Yang M, Wang S, Li X, Sun Y. Emerging role of deep learning-based artificial intelligence in tumor pathology. *Cancer Commun (Lond)* 2020; **40**: 154-166 [PMID: 32277744 DOI: 10.1002/cac2.12012]

29 **Howard FM**, Kochanny S, Koshy M, Spiotto M, Pearson AT. Machine Learning-Guided Adjuvant Treatment of Head and Neck Cancer. *JAMA Netw Open* 2020; **3**: e2025881 [PMID: 33211108 DOI: 10.1001/jamanetworkopen.2020.25881]

30 **She Y**, Jin Z, Wu J, Deng J, Zhang L, Su H, Jiang G, Liu H, Xie D, Cao N, Ren Y, Chen C. Development and Validation of a Deep Learning Model for Non-Small Cell Lung Cancer Survival. *JAMA Netw Open* 2020; **3**: e205842 [PMID: 32492161 DOI: 10.1001/jamanetworkopen.2020.5842]

31 **Li W**, Lin S, He Y, Wang J, Pan Y. Deep learning survival model for colorectal cancer patients (DeepCRC) with Asian clinical data compared with different theories. *Arch Med Sci* 2023; **19**: 264-269 [PMID: 36817685 DOI: 10.5114/aoms/156477]

32 **Wang L**, Hirano Y, Heng G, Ishii T, Kondo H, Hara K, Obara N, Asari M, Kato T, Yamaguchi S. Mucinous Adenocarcinoma as a High-risk Factor in Stage II Colorectal Cancer: A Propensity Score-matched Study from Japan. *Anticancer Res* 2020; **40**: 1651-1659 [PMID: 32132070 DOI: 10.21873/anticanres.14115]

33 **Cai C**, Yang L, Zhuang X, He Y, Zhou K. A five-lncRNA model predicting overall survival in gastric cancer compared with normal tissues. *Aging (Albany NY)* 2021; **13**: 24349-24359 [PMID: 34751670 DOI: 10.18632/aging.203685]

34 **Yue T**, Chen S, Zhu J, Guo S, Huang Z, Wang P, Zuo S, Liu Y. The aging-related risk signature in colorectal cancer. *Aging (Albany NY)* 2021; **13**: 7330-7349 [PMID: 33658390 DOI: 10.18632/aging.202589]

35 **Hung YS**, Chang SC, Liu KH, Hung CY, Kuo YC, Tsai CY, Hsu JT, Yeh TS, Chen JS, Chou WC. A prognostic model based on lymph node metastatic ratio for predicting survival outcome in gastric cancer patients with N3b subclassification. *Asian J Surg* 2019; **42**: 85-92 [PMID: 29248301 DOI: 10.1016/j.asjsur.2017.10.001]

36 **Miao Y**, Xu Z, Feng W, Zheng M, Gao H, Li W, Zhang Y, Zong Y, Lu A, Zhao J. Platelet infiltration predicts survival in postsurgical colorectal cancer patients. *Int J Cancer* 2022; **150**: 509-520 [PMID: 34551136 DOI: 10.1002/ijc.33816]

37 **Pereira de Sousa I**, Cattoz B, Wilcox MD, Griffiths PC, Dalgliesh R, Rogers S, Bernkop-Schnürch A. Nanoparticles decorated with proteolytic enzymes, a promising strategy to overcome the mucus barrier. *Eur J Pharm Biopharm* 2015; **97**: 257-264 [PMID: 25661320 DOI: 10.1016/j.ejpb.2015.01.008]

38 **Marxen E**, Mosgaard MD, Pedersen AML, Jacobsen J. Mucin dispersions as a model for the oromucosal mucus layer in in vitro and ex vivo buccal permeability studies of small molecules. *Eur J Pharm Biopharm* 2017; **121**: 121-128 [PMID: 28974436 DOI: 10.1016/j.ejpb.2017.09.016]

Published by **Baishideng Publishing Group Inc**
7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-3991568
E-mail: office@baishideng.com
Help Desk: https://www.f6publishing.com/helpdesk
https://www.wjgnet.com