ABSTRACT

Objective To explore the influencing factors of survival in intestinal-type gastric adenocarcinoma (IGA) and set up prediction model for the prediction of survival of patients diagnosed with IGA.

Design A retrospective cohort study.

Setting and participants A total of 2232 patients with IGA who came from the Surveillance, Epidemiology, and End Results database.

Primary and secondary outcome measures Patients’ overall survival (OS) rate and cancer-specific survival (CSS) at the end of follow-up.

Results Of the total population, 25.72% survived, 54.93% died of IGA and 19.35% died of other causes. The median survival time of patients was 25 months. The result showed that age, race, stage group, T stage, N stage, M stage, grade, tumour size, radiotherapy, number of lymph nodes removed and gastrectomy were independent prognostic factors of OS risk for patients with IGA; age, race, stage group, T stage, N stage, M stage, grade, radiotherapy and gastrectomy were associated with CSS risk for patients with IGA. In view of these prognostic factors, we developed two prediction models for predicting the OS and CSS risk for patients with IGA separately. For the developed OS-related prediction model, the C-index was 0.750 (95% CI: 0.740 to 0.760) in the training set, corresponding to 0.785 (95% CI: 0.770 to 0.793) in the testing set. Likewise, for the developed CSS-related prediction model, the C-index was 0.781 (95% CI: 0.770 to 0.793) in the training set, corresponding to 0.785 (95% CI: 0.766 to 0.803) in the testing set. The calibration curves of the training set and testing set revealed a good agreement between model predictions in the 1-year, 3-year and 5-year survival for patients with IGA and actual observations.

Conclusion Combining demographic and clinicopathological features, two prediction models were developed to predict the risk of OS and CSS in patients with IGA, respectively. Both models have good predictive performance.

INTRODUCTION

Gastric cancer was considered as the fifth most common cancer and the second most common cancer-related death worldwide. More than 95% of gastric cancers are gastric adenocarcinoma. To make matters worse, most patients with gastric adenocarcinoma were diagnosed at an advanced stage and lacked effective surgical treatment. A study pointed out that these patients with gastric adenocarcinoma had a poor prognosis, with a 5-year survival rate of less than 25%.

Intestinal-type gastric adenocarcinoma (IGA) was considered the most common subtype of gastric adenocarcinoma according to the Lauren’s classification, which was characterised by cohesive cells arranged in glandular structures. A study has pointed out that the overall 5-year mortality rate for IGA reached 42.3%. In addition, Tang et al reported that the patients with IGA had a median survival time of 27 months in a cohort study. These findings suggested that it is important to focus on the prognosis of patients with IGA. However, to the best of our knowledge, most studies to date only focused on the influence of treatment on the prognosis of IGA or the survival of gastric adenocarcinoma. For example, Chu et al expounded that radiotherapy played a survival benefit for patients with IGA. It is beneficial for the management of patients to explore the related factors affecting the prognosis of IGA and establish the predictive model for predicting the survival of patients. In the current study, we aimed to analyse the influencing factors of survival, and set up prediction models related to survival in different populations.
patients diagnosed with IGA. Simultaneously, we present prediction models in the form of online tools, which are more convenient for clinical use.

METHODS
Patient and public involvement statement
Not applicable.

Data extraction and patients’ selection
All data of patients were extracted from the Surveillance, Epidemiology, and End Results (SEER) Regs 18 Custom Data (with Calculated Months Fields) between 2004 and 2015 in this retrospective cohort study. SEER, as a data-accessible and publicly free database, collected data on the epidemiology of various malignancies.11 The database contains patients’ demographics, primary tumour data, regional nodal data, vital status and survival.12

For this retrospective cohort study, we extracted 2753 patients with primary gastric cancer (C16.0–C16.9) who belonged to IGA (histological type: 8144/3) according to the International Classification of Diseases for Oncology (third edition). The exclusion criteria were as follows: (1) patients younger than 20 years old (n=2); (2) patients diagnosed as IGA outside the period 2004–2015 (n=105); (3) patients with incomplete follow-up dates (n=120); (4) patients with unknown survival time (n=34); (5) patients with some missing data (n=108). In addition, there were 152 patients lost to follow-up. After excluding these patients, a total of 2232 patients with IGA were included for analysis (figure 1).

Participant’s data was extracted from the database, including age, sex, race, marital status, stage group, tumour (T) stage, node (N) stage, metastasis (M) stage, grade, tumour size, primary site, number of lymph nodes removed, radiation, chemotherapy, sequence of radiotherapy with surgery, gastrectomy. The primary site was recorded as cardia, fundus of stomach, body of stomach, gastric antrum, pylorus, lesser curvature of stomach, overlapping lesion of stomach and stomach. TNM stage was assessed according to the American Joint Committee on Cancer’s staging (eighth edition) definition for gastric cancer.13 14 T stage was classified into T1 (tumour invades lamina propria or muscularis mucosae or submucosa), T2 (tumour invades muscularis propria), T3 (tumour penetrates subserosal connective tissue), T4 (tumour invades the serosa or adjacent structures) and TX (unknown). N stage was divided into N0 (no lymph nodes positive), N1 (1–2 nodes positive), N2 (3–6 nodes positive), N3 (>7 nodes positive) and NX (unknown). M stage was divided into M0 (no distant metastasis), M1 (distant metastasis) and MX (unknown).

Outcomes
The primary outcome of this study was defined as patients’ overall survival (OS) rate and cancer-specific survival (CSS) rate. OS was considered as the time from the diagnosis to the date of death from any cause or date of last follow-up, and CSS as the time from the diagnosis to the date of IGA-specific death or date of last follow-up. The follow-up deadline is until the end of 2019. The loss of follow-up rate during the follow-up period was ‘6.10%’ (n=152).

Development and validation of prediction model
First, we used univariate and multivariate Cox analyses to screen out predictors related to the survival of patients with IGA. Then, we randomly divided all eligible patients

Figure 1 Flow diagram of the selected patients with IGA. IGA, intestinal-type gastric adenocarcinoma; SEER, Surveillance, Epidemiology, and End Results.
with IGA in two groups. Overall, 70% of patients (n=1563) were selected as the training set for the construction of prediction model, and 30% (n=609) were assigned to the testing set for the validation of the model. The prediction model was established by using these predictors in the training group. Lastly, C-index and calibration curves were adopted to evaluate the performance of the constructed model.

Statistical analysis
The categorical variable adopted the number of cases and composition ratio n (%) to describe, and the difference of groups were compared by \( \chi^2 \) test. The continuous variables were presented as mean±SD (normal distribution) or the median and quartile spacing (M (Q1, Q3), non-normal distribution). The comparison between the two groups was carried out by the student’s t-test or Mann-Whitney U test, respectively.

First, possible factors related to OS and CSS in patients with IGA were identified separately by univariate Cox analysis. These possible factors were then included in multivariate Cox analysis and the predictors were obtained by stepwise regression. We developed two prediction models to predict OS and CSS in the patients with IGA, respectively. Internal validation of two models was performed in the testing set. The HR and 95% CI were calculated. The statistical analyses were performed by using SAS V.9.4 and R V.4.2.0. All statistical tests were two-sided, and p<0.05 was considered to be statistically significant.

RESULTS
Patient characteristics
Overall, 2232 patients with IGA were analysed for this present study. The median survival time of patients was 25 months. As shown in online supplemental table 1, the mean age of patients with IGA was 71.88±12.34 years, 822 (n=36.83%) were women and 1410 (63.17%) were men. Of the total population, 25.72% survived, 54.93% died of IGA and 19.35% died of other causes. In addition, online supplemental table 1 presents the comparison of patients’ characteristics between the training set and testing set. Detailed baseline information was given in online supplemental table 1.

Development and validation of prediction model
Some possible factors (p<0.05, online supplemental table 2) were included in multivariate Cox analysis, and the predictors were obtained by stepwise regression. As shown in online supplemental table 3, the findings found that 11 variables were associated with the OS of patients with IGA. In the training set, a prediction model integrating age, race, stage group, T stage, N stage, M stage, grade, tumour size, radiotherapy, number of lymph nodes removed and gastrectomy was developed to predict the risk of OS for patients with IGA (an online prediction nomogram: https://weibinxu.shinyapps.io/DynNomapp/). Similarly, online supplemental table 3 also indicated that age, race, stage group, T stage, N stage, M stage, grade, radiotherapy and gastrectomy were predictors, which were used to develop a prediction model to predict the risk of CSS for patients with IGA (an online prediction nomogram: https://evianhong.shinyapps.io/DynNomapp/).

In the current study, we validated the predictive performance of two prediction models. For the developed OS-related prediction model, the C-index was 0.750 (95% CI: 0.740 to 0.760) in the training set, corresponding to 0.753 (95% CI: 0.736 to 0.770) in the testing set. Likewise, for the developed CSS-related prediction model, the C-index was 0.781 (95% CI: 0.770 to 0.793) in the training set, corresponding to 0.785 (95% CI: 0.766 to 0.803) in the testing set. Furthermore, the validation curves of the training set and testing set (figures 2 and 3) also revealed a good agreement between model predictions in the 1-year, 3-year and 5-year survival for patients with IGA and actual observations.

DISCUSSION
In this retrospective cohort study, we collected the data of 2232 patients with IGA who came from the SEER database and evaluated the risk factors that in connection with the OS and CSS of patients. The result showed that age, race, stage group, T stage, N stage, M stage, grade, tumour size, radiotherapy, number of lymph nodes removed and gastrectomy were independent prognostic factors of OS risk for patients with IGA; age, race, stage group, T stage, N stage, M stage, grade, radiotherapy and gastrectomy were associated with CSS risk for patients with IGA. In view of these prognostic factors, we developed and validated two prediction models for predicting the OS and CSS risk for patients with IGA separately.

This study considered the age as a risk factor for patients with IGA; namely, the older patients with IGA had a more negative prognosis. Similarly, Chen et al found that age was significantly correlated with the OS in gastric cancer. Additionally, our result revealed that T stage, N stage and M stage were associated with the survival of patients with IGA. Previous studies have shown that the prognosis of gastric adenocarcinoma was related to the TNM stage at the time of diagnosis. Although TNM stage has been recognised as a prognostic indicator of cancer, TNM only considers the extent of invasion of the primary tumour, lymph node metastatic status and distant spread. Additionally, patients with IGA at the same stage may have a different prognosis and survival. These results also imply that the TNM stage may be not optimal for clinical prognosis prediction. Currently, surgery, systemic chemotherapy, radiotherapy and targeted therapy have been proven to have significant survival benefits for patients with gastric adenocarcinoma. We also found that radiotherapy and gastrectomy can improve OS and CSS in patients with IGA in this study. Our results are consistent with previous studies. Additionally, lymphadenectomy is also regarded as the cornerstone in the treatment of gastric...
adenocarcinoma. It has previously been suggested that the number of lymph nodes removed in gastric adenocarcinoma may have prognostic significance. Biffi et al have shown that removal of at least 15 lymph nodes may be important to improve survival of patients with node-negative gastric cancer. However, the results of multivariate Cox analysis showed that the number of lymph nodes removed was not associated with the prognosis of IGA in this study. Considering the possible impact of the number of lymph nodes removed on patient outcomes, we included the number of lymph nodes removed in the prediction model.

To date, several studies have developed predictive models for predicting the prognosis of patients with gastric adenocarcinoma. Zhao et al successfully established a prognostic model based on metabolism-related genes to predict OS of patients with gastric adenocarcinoma. A study from China constructed a nomogram model to predict the prognosis of patients with stage II/III gastric adenocarcinoma after radical gastrectomy by combining clinicopathological features, baseline inflammatory parameters and tumour markers, where older age and later stage were considered to be significantly associated with lower OS. In fact, most of these studies tend to...
focus more on the prognosis of diffuse type gastric adenocarcinoma, and very few to focus on the prognosis of IGA. In general, IGA is the dominant type of gastric adenocarcinoma and has a better prognosis compared with the diffuse type gastric adenocarcinoma. However, some studies also pointed out that the 5-year survival rate of IGA was only about 60%. Therefore, the development of prognostic models to identify patients with IGA at high risk would be helpful for individual risk determination, clinical decision-making and prognosis improvement. Notably, our study constructed two prediction models that integrated patients’ demographic and clinicopathological characteristics to predict the OS and CSS with respect to patients with IGA, respectively. A total of 11 variables were included in the OS-related prediction model. Similar variables were included in the CSS-related prediction model. Importantly, this study presents two prediction models in the form of two online tools, which are more convenient for clinical use (OS-related prediction model: https://weibinxu.shinyapps.io/DynNomapp/; CSS-related prediction model: https://evianhong.shinyapps.io/DynNomapp/). A good predictive performance was also verified by C-index and the calibration curves.

The advantage of this study should be pointed out: (1) our model was established based on the data in the SEER database, and since the SEER database contains different races, which means that our model may be suitable for the prognosis prediction in different populations; (2) the proposed model contains some parameters that were easy to obtain, and the internal verification data also showed a good predictive performance, which supporting a potential application of our model in clinical practice. Nevertheless, some limitations of this study were as follows: (1) because all data derived from the SEER database, and we excluded some patients who had the missing data and lost follow-up during the data collection process. We could not be sure if these patients may affect our results. (2) The sample size of this study was relatively small, which might affect the prediction power of this model. (3) Regarding the patient’s treatment information, the database only shows the sequence of radiotherapy with surgery, as well as information on whether surgery, chemotherapy and radiotherapy. Other details, such as treatment regimen, lymphadenectomy (D1, D2), chemotherapy sequence, neoadjuvant, adjuvant chemotherapy and concurrent chemotherapy–radiotherapy, were not collected in the SEER database. (4) Although an internal validation of the model was performed and showed a good predictive performance, this study still lacked an external validation. Therefore, our results should be interpreted with caution, and more studies will be used in the future to verify the validity and applicability of the constructed model.

CONCLUSION
In conclusion, our study found the risk factors that might be related to the OS and CSS of patients with IGA. Combining demographic and clinicopathological features, two prediction models were developed to predict the risk of OS and CSS in patients with IGA, respectively. Both models have good predictive performance. Future studies will focus on the predictive value and applicability of these constructed models.

Contributors WX is responsible for the overall content as the guarantor. JH and WX designed the study. JH wrote the manuscript. YC and XG collected, analysed and interpreted the data. WX critically reviewed, edited and approved the manuscript. All authors read and approved the final manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained directly from patient(s).

Ethics approval This study involves human participants. The requirement of ethical approval for this was waived by the institutional review board of Rudong People’s Hospital, because the data was accessed from the SEER (a publicly available database). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data sharing not applicable as no datasets generated and/or analysed for this study. There are no additional data.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD Weibing Xu http://orcid.org/0000-0002-8513-8003

REFERENCES