1. Compared with the existing researches of the same type, what is the innovation of this paper?

Response: Thank you for your suggestion. Our study analyzed the independent risk factors for LNM in EGC patients based on the SEER database. Based on a 7:3 ratio, 2217 EGC patients were randomly divided into training and testing sets, respectively, which means that 1550 people were in the training set and 667 people were in the testing set. Compared with the existing researches of the same type, our study had a larger sample size and built a clinical predictive model and validate them internally and externally.

2. The included data are from 2010 to 2015, and it is suggested to include the data of recent years.

Response: Thank you for your suggestion. We analyzed and discussed the independent risk factors for LNM in EGC patients based on data from 2010 to 2015 in this study. We want to know the results of five years and then compared with another fives years(2016-2020) and then calculated the ten years. In the future, we will include the data of recent years in the analysis.

3. As mentioned in the main text, whether unclear data can be deleted can improve data accuracy.

Response: Thank you for your suggestion. The data from 2010 to 2015 in this study were carefully screened and processed. Although there are some variables with unclear data, these data accounted for a small percentage and had almost no effect on the subsequent analysis. If a few unclear data are removed for each variable, multiple variables cumulatively will have a large portion of the data removed, resulting in a significant reduction in the sample size, besides, we build a model with unclear data, so we can know the points and analyze the risk of LNM of some patient with uncomplete data.

Manuscript review "Nomogram established using risk factors of early gastric cancer for predicting the lymph node metastasis"

Metastases are the main cause of death in cancer patients. In early gastric cancer, assessing the risk of lymphogenous metastasis is important, as it determines not only the prognosis of the disease, but also the type and extent of surgical intervention. The authors analyzed the independent risk factors for LNM in EGC patients and based on the above risk factors, the nomogram was established for predicting risk of LNM in EGC patients, which might be beneficial for selecting a more precise treatment.

Despite the relevance of the study, there are a number of comments on the manuscript that should be corrected:

1. In accordance with the requirements of the journal, it is necessary to submit a Core Tip;

Response: Thank you for your suggestion. We have been added a Core Tip in our manuscript that meet the requirements of the journal.

2. The appropriateness of using "Surveillance, Epidemiology, and End Results (SEER)" in the keywords is questionable

Response: Thank you for your suggestion. We have been removed "Surveillance, Epidemiology, and End Results (SEER)" from the Key words section.

Introduction

1. Incorrect wording: "In previous research, LNM, as the independent risk factor for EGC, had good predictive value(12).". LNM is a prognostic factor for stomach cancer.

Response: Thank you for your suggestion. We have revised according to your correction.

Results

Characteristics of patients
Authors should provide age limits for patients included in the study

Response: Thank you for your suggestion. We have added the range of age of enrolled patients in "characteristics of patients" section. Because prognostic models need to be constructed for scoring in the nomograph, a stratified score and quantitative analysis of age - the continuous variable - is needed to calculate the score and assess the hazard level. Although the range of the patients' age is 18 to 100, it is necessary to quantitatively score all age groups. So, the stratification parameter is set to 10, so it is 10 to 100 for hierarchical scoring.

Construction of the Prediction Model for EGC Patients
Correct -   ranged from 0 to 400, is the sum of the points of all variable.

Response: Thank you for your suggestion. We have revised according to your correction.

Figure 2. The age line needs to be corrected. It should correspond to the age limits of the patients included in the study.

Response: Thank you for your suggestion. Because prognostic models need to be constructed for scoring in the nomograph, a stratified score and quantitative analysis of age - the continuous variable - is needed to calculate the score and assess the hazard level. Although the range of the patients' age is 18 to 100, it is necessary to quantitatively score all age groups. So, the stratification parameter is set to 10, so it is 10 to 100 for hierarchical scoring.