Received: 18 July 2022 | Revised: 17 October 2022 | Accepted: 8 November 2022

DOI: 10.1002/ijc.34385

INNOVATIVE TOOLS AND METHODS

IJC INTERNATIONAL JOURNAL of CANCER

SOURCE beyond first-line: A survival prediction model for patients with metastatic esophagogastric adenocarcinoma after failure of first-line palliative systemic therapy

Steven C. Kuijper[1,2] | Marieke Pape[1,2,3] | Nadia Haj Mohammad[4] | Theo van Voorthuizen[5] | Rob H. A. Verhoeven[1,2,3] | Hanneke W. M. van Laarhoven[1,2]

[1]Amsterdam UMC Location University of Amsterdam, Medical Oncology, Amsterdam, The Netherlands

[2]Cancer Center Amsterdam, Cancer Treatment and Quality of Life, Amsterdam, The Netherlands

[3]Department of Research & Development, Netherlands Comprehensive Cancer Organisation (IKNL), Utrecht, The Netherlands

[4]Department of Medical Oncology, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

[5]Department of Medical Oncology, Rijnstate Hospital, Arnhem, The Netherlands

Correspondence
Steven C. Kuijper, Amsterdam UMC, Department of Medical Oncology, 1100DD Amsterdam, The Netherlands.
Email: s.c.kuijper@amsterdamumc.nl

Hanneke W. M. van Laarhoven, Amsterdam UMC, Department of Medical Oncology, Office D3-312, PO Box 22660, 1100DD Amsterdam, The Netherlands.
Email: h.vanlaarhoven@amsterdamumc.nl

Funding information
KWF Kankerbestrijding, Grant/Award Number: UVA 2014-7000

Abstract

Prior models have been developed to predict survival for patients with esophagogastric cancer undergoing curative treatment or first-line chemotherapy (SOURCE models). Comprehensive clinical prediction models for patients with esophagogastric cancer who will receive second-line chemotherapy or best supportive care are currently lacking. The aim of our study was to develop and internally validate a new clinical prediction model, called SOURCE beyond first-line, for survival of patients with metastatic esophagogastric adenocarcinoma after failure of first-line palliative systemic therapy. Patients with unresectable or metastatic esophageal or gastric adenocarcinoma (2015-2017) who received first-line systemic therapy (N = 1067) were selected from the Netherlands Cancer Registry. Patient, tumor and treatment characteristics at primary diagnosis and at progression of disease were used to develop the model. A Cox proportional hazards regression model was developed through forward and backward selection using Akaike's Information Criterion. The model was internally validated through 10-fold cross-validations to assess performance. Model discrimination (C-index) and calibration (slope and intercept) were used to evaluate performance of the complete and cross-validated models. The final model consisted of 11 patient tumor and treatment characteristics. The C-index was 0.75 (0.73-0.78), calibration slope 1.01 (1.00-1.01) and calibration intercept 0.01 (0.01-0.02). Internal cross-validation of the model showed that the model performed adequately on unseen data: C-index was 0.79 (0.77-0.82), calibration slope 0.93 (0.85-1.01) and calibration intercept 0.02 (−0.01 to 0.06). The SOURCE beyond first-line model predicted survival with fair discriminatory ability and good calibration.

KEYWORDS
best supportive care, esophageal cancer, gastric cancer, prediction model, second-line

Abbreviations: AIC, Akaike's Information Criterion; HER2, human epidermal growth factor receptor 2; IKNL, Netherlands Comprehensive Cancer Organization; LDH, lactate dehydrogenase; MICE, multiple imputations by chained equations; NCR, Netherlands Cancer Registry; OS, overall survival; PALGA, Dutch Nationwide Pathology Databank; TRIPOD, Transparent Reporting of a Multivariable Prediction model for Individual Prognosis or Diagnosis; WHO, World Health Organization.

What's new?

Patients with metastatic esophagogastric cancer generally have poor survival. Here, the authors present a model for predicting survival after failure of first-line palliative systemic treatment. They collected data from 1067 patients in the Netherlands Cancer Registry who had unresectable or metastatic esophageal or gastric adenocarcinoma. Unlike existing models, this model includes information about second-line chemotherapies and best supportive care, and was developed on a large number of patients. The final model included 11 different characteristics of the patient, tumor, and treatment, and could help patients and providers make informed decisions about starting second-line therapies.

1 | INTRODUCTION

Survival of patients with metastatic esophagogastric cancer is poor.[1,2] First-line palliative systemic treatment for patients with metastatic esophagogastric cancer has the potential to extend survival, and to improve or sustain quality of life.[3-5] After failure of first-line systemic treatment, patients have the option to continue with second-line palliative systemic therapy or best-supportive care. Second-line treatment with paclitaxel and ramucirumab is considered standard of care for patients with esophagogastric adenocarcinoma.[6-8] In clinical practice, roughly a quarter of patients that received first-line systemic therapy continue with second-line systemic therapy and have a median overall survival (OS) of 5.4 months since start of second-line treatment.[9]

The emergence of prediction models has enabled physicians to improve communication of individualized information regarding life expectancy and can aid in shared decision making.[10] Recently, the SOURCE and SOURCE-PANC prediction models for patients with curable or incurable esophagogastric cancer and incurable pancreatic cancer, respectively, have shown good predictive performances and are important in informing patients about treatment outcomes.[11-13]

Currently, two prediction models exist for the survival after failure of first-line systemic treatment for patients with gastric cancer.[14,15] The first consisted of a prognostic model for patients with gastric cancer who received second-line chemotherapy.[14] However, this model lacks internal and external validation, and thus predictive performance on novel data cannot be assessed. Furthermore, at the time of publication second-line therapy ramucirumab and paclitaxel was not available and therefore not included in the model. Since second-line therapy with ramucirumab and paclitaxel has improved survival in recent years, the existing model has become less relevant for current clinical practice.[8] The second prediction model did internally and externally validated the model, but the model was trained on relatively small number of patients and only included patients that received second-line chemotherapy.[15] Patients that received best supportive care were not included. Finally, both models were developed for patients with gastric cancer only, and cannot be used for patients with esophageal cancer.

The aim of our study was to develop and internally validate a survival prediction model using nationwide population-based data of patients with esophagogastric adenocarcinoma after failure of first-line palliative systemic treatment for use in clinical practice with patient, tumor and treatment characteristics.

2 | METHODS

2.1 | Data collection

This manuscript is written according to the Transparent Reporting of a Multivariable Prediction model for Individual Prognosis or Diagnosis (TRIPOD) statement. Patients with synchronous metastatic adenocarcinoma (2015-2017) of the esophagus (C15.0-C15.9), gastroesophageal junction (GEJ)/cardia (C16.0) or stomach (C16.1-C16.9) and patients with a metachronous metastatic disease initially treated with curative intent (2015-2016) for a nonmetastatic esophageal, GEJ/cardia or gastric adenocarcinoma, who received first-line palliative systemic treatment in the Netherlands were selected from the Netherlands Cancer Registry (NCR). The NCR is a nationwide based registry that covers the total Dutch population of more than 17 million people. The NCR is linked to the pathology archive in the Netherlands (PALGA), which contains information from all newly diagnosed malignancies. Trained data managers routinely extract patient and treatment information from electronic medical records.

Follow-up information on tumor and treatment characteristics (including metachronous metastatic disease) was collected in the second half of 2019, with the exception of two hospitals due to logistical constraints. Data on vital status were obtained through annual linkage to the Dutch Personal Records Database and updated until February 2021. Metachronous metastatic disease was defined as diagnosis of metastases at least 5 days after end of treatment with curative intent for primary nonmetastatic disease for patients diagnosed in 2015 to 2016. Treatment with curative intent was defined as endoscopic resection, surgical resection or definitive chemoradiotherapy (chemotherapy with concurrent radiotherapy consisting of ≥28 fractions or total radiation dose of ≥50 Gy). For patients who developed metachronous metastases within 6 months after end of neoadjuvant chemotherapy, the neoadjuvant chemotherapy was considered as first-line systemic therapy. First-line palliative systemic therapy was defined as all chemotherapy or targeted agents that started within 3 days of each other, as described in more detail in a previous publication.[16] Second-line treatment was considered when a new agent of a different drug group was started that was not administered in first-line.[9] Patients with first-line treatment failure for other reasons than disease progression were excluded. Furthermore, patients were also excluded if patients first-line therapy despite progression or if a restart of the first-line was initiated after disease progression A comprehensive overview of included patients is available in Supplementary Figure 1.

TABLE 1 Patient, disease and treatment characteristics at progression of disease after failure of first-line

	All patients (N = 1067)
Median survival (95% CI), months	3.55 (3.29-3.84)
Variables at primary diagnosis	
Sex	
Male	835 (78.3%)
Female	232 (21.7%)
cT	
1	2 (0.2%)
1A	2 (0.2%)
1B	1 (0.1%)
2	386 (36.2%)
3	355 (33.3%)
4A	36 (3.4%)
4B	48 (4.5%)
X	237 (22.2%)
cN	
0	205 (19.2%)
1	359 (33.6%)
2	359 (33.6%)
3	91 (8.5%)
X	53 (5.0%)
Primary tumor location	
Esophagus	563 (52.8%)
Stomach	317 (29.7%)
GE-junction/cardia	187 (17.5%)
Tumor differentiation	
Well	21 (2.0%)
Moderate	256 (24.0%)
Poorly	417 (39.1%)
Unknown	373 (35.0%)
Variables at progression of disease	
Age	
Mean (SD)	63.38 (10.00)
Albumin (g/L)	
Mean (SD)	35.43 (7.13)
Missing (N)	390
LDH (U/L)	
Mean (SD)	412.04 (695.25)
Missing (N)	162
Neutrophile count ($\times 10^9$/L)	
Mean (SD)	5.76 (3.786)
Missing (N)	384
HER2 status	
Negative	678 (63.5%)
Positive	192 (18.0%)
Unknown	197 (18.5%)
WHO performance status	
0	116 (10.9%)
1	310 (29.1%)
2	110 (10.3%)
>2	97 (9.1%)
Missing (N)	431 (40.5%)
Duration first line therapy (mo)	
Mean (SD)	8.80 (8.223)
Type of metastatic disease	
Metachronous metastases	185 (17.3%)
Synchronous metastases	882 (82.7%)
Number of metastatic sites	
Mean (SD)	2.57 (1.39)
First-line therapy	
Monotherapy	51 (4.8%)
Doublet therapy	627 (58.8%)
Triplet therapy	212 (19.9%)
Trastuzumab-containing regimen	164 (15.4%)
Nontrastuzumab targeted therapy-containing regimen	13 (1.2%)
Type of second-line treatment	
Paclitaxel and ramucirumab	232 (21.7%)
Monochemotherapy	114 (10.7%)
Doublet or triplet chemotherapy	81 (7.6%)
Best supportive care	640 (60.0%)

2.2 | Model development and validation

Characteristics of patients included in our study were summarized with mean and standard deviation (SD) for continuous variables, frequencies for categorical variables, and median for overall survival estimates. An initial variable selection was performed to select predictors that were available for at least 50% of patients. After the selection, all variables that were available at primary diagnosis and at progression after the first-line treatment were used for the modeling procedure. Potential predictors included patient, tumor and treatment characteristics. Type of treatment (including best supportive care) after first-line systemic therapy was a mandatory variable and was forced to be in the model, since the model's primary aim is predicting treatment effects.

Next, a multivariable Cox proportional hazard model was fitted. Through back- and forward variable selection, the final set of predictors was determined and fitted as the final model. Predictor selection was performed based on Akaike's Information Criterion (AIC). Multiple imputations by chained equations (MICE) with 10 iterations was used to handle missing data,[17] with the exception of cN-stage, cT-stage, differentiation grade and HER2-status. These variables could not be

TABLE 2 Hazard ratios (HR) of overall survival and 95% confidence intervals of predictors in the model

	HR (95% CI)	*P*-value
Variables at primary diagnosis		
cN		
0	Reference	
1	0.75 (0.63-0.90)	.002
2	0.81 (0.68-0.97)	.022
3	1.05 (0.81-1.36)	.734
X	1.06 (0.78-1.45)	.703
Tumor differentiation grade		
Well	Reference	
Moderate	1.19 (0.76-1.88)	.449
Poorly	1.41 (0.90-2.21)	.129
Unknown	1.16 (0.74-1.82)	.520
Variables at progression of the disease		
WHO performance status?		
0	Reference	
1	1.33 (1.11-1.59)	.002
2	1.38 (1.10-1.73)	.006
>2	3.06 (2.41-3.87)	<.001
Albumin (g/L)	0.99 (0.98-1.00)	.014
LDH (U/L)	1.0002 (1.0001-1.0003)	<.001
Neutrophils count (10^9/L)	1.04 (1.02-1.06)	<.001
HER2 status		
Negative	Reference	
Positive	0.80 (0.67-0.94)	.008
Unknown	0.96 (0.82-1.14)	.650
Duration first-line systemic therapy (months)	0.96 (0.95-0.97)	<.001
Type of metastatic disease		
Metachronous	Reference	
Synchronous	0.65 (0.53-0.81)	<.001
Number of metastatic sites after first-line therapy	1.21 (1.16-1.26)	<.001
Treatment after first-line therapy		
Paclitaxel and Ramucirumab	Reference	
Monochemotherapy	1.23 (0.97-1.54)	.082
Doublet or triplet chemotherapy	1.21 (0.93-1.56)	.149
Best supportive care	2.65 (2.24-3.14)	<.001

assumed to be missing at random and the fact that they are missing was likely to have predictive information. We therefore included a separate category "unknown" in these variables, which was attributed to patients whose data was missing on that variable.

The predictive performance of the final model was evaluated with the concordance index (C-index), calibration slope and calibration intercept. The C-index is a measure for model discrimination and ranges from 0.5 (random chance) to 1.0 (perfect discrimination).[18] C-indices of 0.60 to 0.69 are typically interpreted as poor discrimination, 0.70 to 0.79 fair discrimination, 0.80 to 0.89 good discrimination and 0.90 to 1.0 excellent discrimination.[19] The calibration slope and intercept refer to the accordance between predicted and observed survival outcomes.[20] For each prediction, we calculated the partial chi-squared statistic minus the predictor degree of freedom which quantified the relative importance of each variable.[21] Higher values correspond with higher relative variable importance.

To assess model performance on unseen data, 10-fold cross-validation was performed.[20] With this method, the data is randomly shuffled and split into 10 equal parts called folds. The model was then trained in nine folds and tested in the remaining fold. This process is repeated 10 times so that every patient is included in the train and test fold at least once. The C-index, calibration slope and intercept across cross-validations were evaluated with a meta-analysis to obtain pooled performance estimates similarly to previously published SOURCE and SOURCE-PANC models.[11,12]

3 | RESULTS

3.1 | Predictors

We identified 1067 patients with metastatic esophagogastric adenocarcinoma with failure on first-line palliative systemic treatment (Table 1). Median OS of all patients since progression was 3.6 (95% CI: 3.2-3.8) months (Table 1; Supplementary Figure 2). After back and forward predictor selection, the final model contained 11 patient tumor and treatment characteristics (Table 2). A significant predictor at primary diagnosis was cN-stage. Although tumor differentiation grade remained in the final model, its hazard ratios were not significant compared to the reference. Significant predictors after progression on first line systemic therapy were WHO performance status, albumin (g/L), lactate dehydrogenase (LDH) (U/L), neutrophils count (10^9/L), human epidermal growth factor receptor 2 (HER2) status, duration of first-line systemic therapy (months), type of metastatic disease (synchronous or metachronous), number of metastatic sites and type of treatment after failure of first-line therapy (including best supportive care).

At primary diagnosis, cN1 and cN2 were associated with higher OS compared to cN0. At progression, poorer WHO performance status, higher LDH concentrations, higher neutrophils count, a higher number of metastatic sites was associated with lower OS. Patients with HER2 positive tumors, higher albumin concentrations, synchronous metastatic disease, and a longer duration of first-line therapy were associated with higher OS.

3.2 | Model performance

The final model had a C-index of 0.75 (0.73-0.78), calibration slope of 1.01 (1.00-1.01) and calibration intercept of 0.01 (0.01-0.02) (Figure 1). 10-Fold cross-validation showed similar point estimates, C-

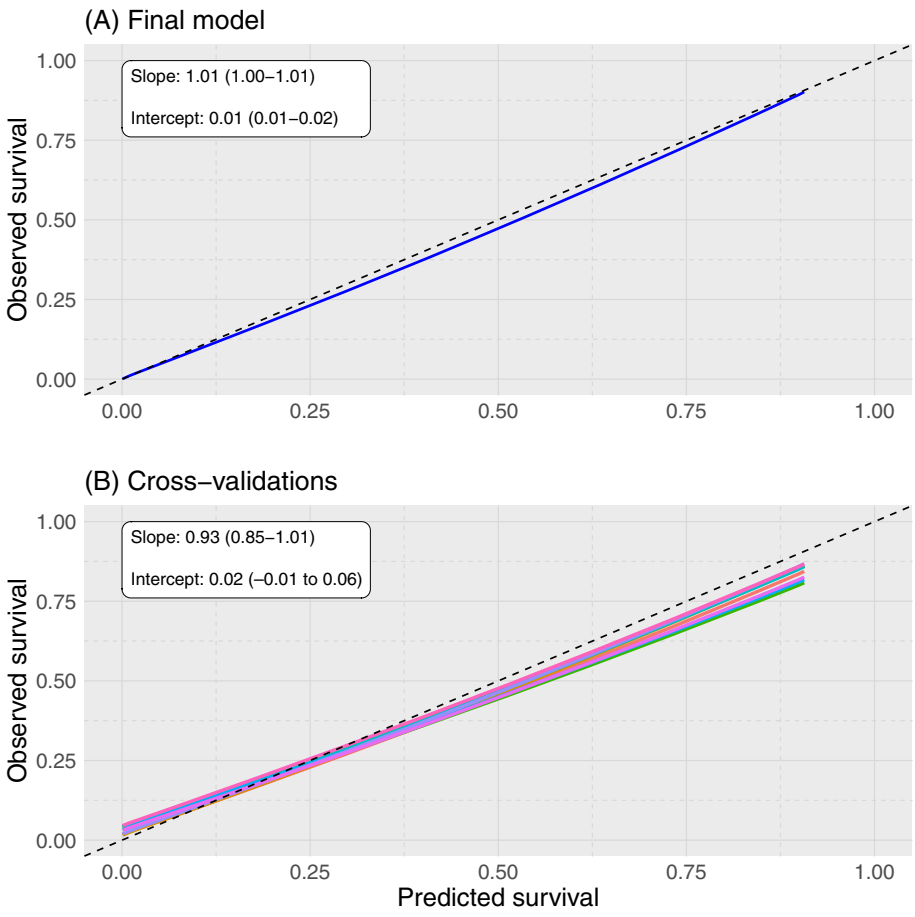

FIGURE 1 Calibration plot of complete model and cross-validations. This shows the accordance between the observed survival and the predicted survival of the final model (A) and across 10-fold cross-validations (B). The colored lines represent 10 different validation folds on which the trained model was tested. Perfect values are a slope of 1 and an intercept of 0

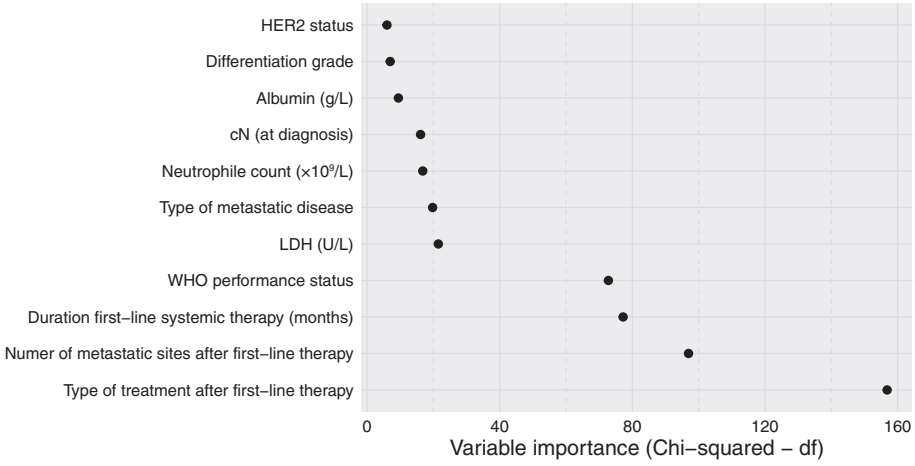

FIGURE 2 Relative variable importance. Variables with higher values correspond with a larger relative importance in predicting survival in the final Cox-regression model. Relatively, type of treatment after first-line systemic therapy had the most predictive capability. df, degrees of freedom; LDH, lactate dehydrogenase; WHO, World Health Organization

index of 0.79 (0.77-0.82), calibration slope 0.93 (0.85-1.01) and calibration intercept of 0.02 (−0.01 to 0.06) (Figure 1). In the final model, the type of treatment after first-line systemic therapy was the most predictive for survival, followed by the number of metastatic sites and duration of the first-line therapy (Figure 2).

A nomogram of the model predicting 6 month and 1-year survival is available in Supplementary Figure 3. Predictions can be made by adding the points of each variable, and finding the corresponding probability to the total amount of points.

4 | DISCUSSION

Our study developed the first population-based prediction model for survival of patients with metastatic esophagogastric adenocarcinoma after failure of first-line palliative systemic therapy. The model showed fair discrimination (0.75) and good accordance between predicted and observed overall survival. This indicates that the SOURCE beyond first-line model can be valuable for shared decision making between patient and physician when

considering second-line palliative systemic therapy or best-supportive care.

The predictive performance of the prediction model was fairly similar to the previously developed SOURCE models for esophagogastric cancer where the C-indexes ranged from 0.73 to 0.78, and calibration estimates were alike.[11,12] In line with the previous SOURCE model for patients with metastatic esophagogastric cancer, WHO performance status, albumin, LDH, HER2 status, cN stage and number of metastatic sites were predictive for overall survival. In the final SOURCE beyond-first line model, poorer WHO performance status, a larger number of metastatic sites and higher LDH concentrations were predictive of lower OS. Higher albumin concentrations, positive HER2-status, cN1 and cN2 compared to cN0 were predictive of higher OS. Unlike previous SOURCE models for esophagogastric cancer, patient characteristics such as age, sex and body mass index were not predictive for survival after failure of first-line systemic therapy due to progression.[11] Since survival of these patients is generally poor (median survival of around 4 months in our study), variables that reflect patients' fitness may be more predictive than general patient characteristics.[22]

Novel predictor variables included the neutrophils count, duration of the first-line systemic therapy and whether a patient had synchronous or metachronous metastatic disease. Higher neutrophils count were predictive for lower OS, which is consistent with earlier findings.[23,24] It is suggested that increased numbers of neutrophils can reduce anticancer activity and increase tumor growth.[25-27] Furthermore, longer first-line therapy was predictive for higher OS, which showed that patients that respond well to first-line chemotherapy have a better OS. Synchronous metastases were predictive for higher OS compared to patients with metachronous metastases. Finally, compared to paclitaxel and ramucirumab best supportive care was predictive of a lower OS. Monochemotherapy and doublet or triplet chemotherapy were not predictive of a different OS compared to the reference treatment paclitaxel and ramucirumab.

Furthermore, although WHO performance status after failure of the first-line due to progression was predictive for survival, it was not the most predictive variable as this was the type of treatment after first-line therapy. Performance status should be accounted for in the decision to start or forgo second-line treatment, however our results show that variability of survival among patients cannot be solely accounted for by patients' performance status.[28,29] It should be noted, that 40% of performance status scores were missing.

The robustness and generalizability of the models was assessed and tested with an internal-external 10-fold cross validation scheme. With this method it can be assessed how the model performs on data that was not used for training the model. In development of prior SOURCE models, a temporal cross-validation scheme was employed which mimics real-world practice of testing the model on a new sample of patients.[11] However, follow-up of patients diagnosed between 2015 and 2017 was obtained in 2019. Hence, the follow-up time for patients diagnosed in 2017 was shorter compared to patients diagnosed in 2015. Survival estimates from these cohorts may therefore be different. To counter this potential source of bias, we created the folds using a random patient sample rather than consecutive cohorts of patients. Additionally, missing data of continuous variables were handled through multiple imputation with chained equations (MICE), which reduces bias due to missing data and is preferred over complete case analysis.[30] Missing differentiation grade, cN, cT, and HER2-status variables were handled by including "unknown" as a separate category. The combination of missing indicators with multiple imputation has been found to be a valid method to handle missing data that are not missing at random.[31] This is also useful in clinical practice since not all variables can always be known for some patients.

Our study has several limitations. First, this prediction model is only developed using population-based data from the Netherlands which could affect the generalizability to other populations of patient with esophagogastric cancer. Second, health related quality of life is an important prognostic factor for survival in patients with metastatic esophagogastric cancer, but was not available to use in our study.[32]

A strength of our study is that the prediction model was developed on data from the population-based Netherlands cancer registry which is directly linked to the national pathology archive. Additionally, steps were taken to increase the robustness and generalizability of the results. Finally, this is the first model predicting survival for patients with esophagogastric cancer after failure of first-line treatment due to progression which includes paclitaxel and ramucirumab as second-line therapy and can be used as a treatment decision aid.

5 | CONCLUSION

Our study presented a prediction model for patients with esophagogastric adenocarcinoma that receive second-line systemic therapy or best-supportive care after failure of the first-line due to progression. The SOURCE beyond first-line model predicted survival with fair discriminatory ability and good calibration. In the future this model will be integrated in an online decision support tool to be used in clinical practice.

AUTHOR CONTRIBUTIONS

Conceptualization, methodology: Steven C. Kuijper, Marieke Pape, Rob H. A. Verhoeven, Hanneke W. M. van Laarhoven. *Formal analysis*: Steven C. Kuijper. *Writing—original draft*: Steven C. Kuijper, Marieke Pape. *Writing—review and editing*: Steven C. Kuijper, Marieke Pape, Nadia Haj Mohammad, Theo van Voorthuizen, Rob H. A. Verhoeven, Hanneke W. M. van Laarhoven. The work reported in the paper has been performed by the authors, unless clearly specified in the text.

ACKNOWLEDGEMENTS

The authors thank the registration team of the Netherlands Comprehensive Cancer Organization (IKNL) for the collection of data for the Netherlands Cancer Registry. The authors thank all participating hospitals in the Netherlands.

FUNDING INFORMATION

This work was supported by The Dutch Cancer Society, grant number UVA 2014-7000. The funder had no role in the design and conduct of the study; analysis and interpretation of the data; and preparation of the manuscript.

CONFLICT OF INTEREST

Hanneke W. M. van Laarhoven has served as a consultant for BMS, Dragonfly, Eli Lilly, MSD, Nordic Pharma and Servier and as a speaker for Astellas and Novartis. She has received research funding and/or medication supply from Bayer, BMS, Celgene, Janssen, Incyte, Eli Lilly, MSD, Nordic Pharma, Philips, Roche, Servier. Rob H. A. Verhoeven reports grants from Roche, BMS and has served as consultant for Daiichi Sankyo. All funding was paid to the institution and was not related to the current study. Nadia Haj Mohammad has served as consultant for BMS, MSD, Astra Zenaca, Eli Lilly, Servier and has received funding from Servier. Steven C. Kuijper, Theo van Voorthuizen and Marieke Pape have no disclosures to declare.

DATA AVAILABILITY STATEMENT

Data are available via the Netherlands Cancer Registry (https://iknl.nl/). Further information is available from the corresponding author upon request.

ORCID

Steven C. Kuijper https://orcid.org/0000-0001-7035-7863
Marieke Pape https://orcid.org/0000-0001-9054-7541
Nadia Haj Mohammad https://orcid.org/0000-0002-4688-2921

REFERENCES

1. van Putten M, de Vos-Geelen J, Nieuwenhuijzen GAP, et al. Long-term survival improvement in oesophageal cancer in The Netherlands. *Eur J Cancer*. 2018;94:138-147. doi:10.1016/j.ejca.2018.02.025
2. Bernards N, Creemers GJ, Nieuwenhuijzen GAP, Bosscha K, Pruijt JFM, Lemmens VEPP. No improvement in median survival for patients with metastatic gastric cancer despite increased use of chemotherapy. *Ann Oncol*. 2013;24(12):3056-3060. doi:10.1093/annonc/mdt401
3. Al-Batran SE, Ajani JA. Impact of chemotherapy on quality of life in patients with metastatic esophagogastric cancer. *Cancer*. 2010;116(11):2511-2518. doi:10.1002/cncr.25064
4. ter Veer E, Mohammad NH, van Valkenhoef G, et al. The efficacy and safety of first-line chemotherapy in advanced esophagogastric cancer: a network meta-analysis. *J Natl Cancer Inst*. 2016;108(10):1-13. doi:10.1093/jnci/djw166
5. Janmaat VT, Steyerberg EW, van der Gaast A, et al. Palliative chemotherapy and targeted therapies for esophageal and gastroesophageal junction cancer. *Cochrane Database Syst Rev*. 2017;2017(11):CD004063. doi:10.1002/14651858.CD004063.pub4
6. Lordick F, Mariette C, Haustermans K, Obermannová R, Arnold D, the ESMO Guidelines Committee. Oesophageal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27:v50-v57. doi:10.1093/annonc/mdw329
7. Smyth EC, Verheij M, Allum W, et al. Gastric cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2016;27:v38-v49. doi:10.1093/annonc/mdw350
8. Wilke H, Muro K, Van Cutsem E, et al. Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *Lancet Oncol*. 2014;15(11):1224-1235. doi:10.1016/S1470-2045(14)70420-6
9. Dijksterhuis WPM, Verhoeven RHA, Pape M, et al. Hospital volume and beyond first-line palliative systemic treatment in metastatic oesophagogastric adenocarcinoma: a population-based study. *Eur J Cancer*. 2020;139:107-118. doi:10.1016/j.ejca.2020.08.010
10. Vickers AJ. Prediction models in cancer care. *CA Cancer J Clin*. 2011;61(5):315-326. doi:10.3322/caac.20118
11. van den Boorn HG, Abu-Hanna A, Haj Mohammad N, et al. SOURCE: prediction models for overall survival in patients with metastatic and potentially curable esophageal and gastric cancer. *J Natl Compr Canc Netw*. 2021;19(4):403-410. doi:10.6004/jnccn.2020.7631
12. van den Boorn HG, Dijksterhuis WPM, van der Geest LGM, et al. SOURCE-PANC: a prediction model for patients with metastatic pancreatic ductal adenocarcinoma based on nationwide population-based data. *J Natl Compr Cancer Netw*. 2021;19(9):1045-1053. doi:10.6004/jnccn.2020.7669
13. Van De Water LF, Van Den Boorn HG, Hoxha F, et al. Informing patients with esophagogastric cancer about treatment outcomes by using a web-based tool and training: development and evaluation study. *J Med Internet Res*. 2021;23(8):e27824. doi:10.2196/27824
14. Kanagavel D, Pokataev IA, Fedyanin MY, et al. A prognostic model in patients treated for metastatic gastric cancer with second-line chemotherapy. *Ann Oncol*. 2010;21(9):1779-1785. doi:10.1093/annonc/mdq032
15. Pietrantonio F, Barretta F, Fanotto V, et al. Estimating survival probabilities of advanced gastric cancer patients in the second-line setting: the gastric life nomogram. *Oncologia*. 2018;95(6):344-352. doi:10.1159/000491753
16. Dijksterhuis WPM, Verhoeven RHA, Slingerland M, et al. Heterogeneity of first-line palliative systemic treatment in synchronous metastatic esophagogastric cancer patients: a real-world evidence study. *Int J Cancer*. 2020;146(7):1889-1901. doi:10.1002/ijc.32580
17. van Buuren S, Groothuis-Oudshoorn K. mice: multivariate imputation by chained equations in R. *J Stat Softw*. 2011;45(3):1-67. doi:10.18637/jss.v045.i03
18. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology*. 2010;21(1):128-138. doi:10.1097/EDE.0b013e3181c30fb2
19. van den Boorn HG, Engelhardt EG, van Kleef J, et al. Prediction models for patients with esophageal or gastric cancer: a systematic review and meta-analysis. *PLoS One*. 2018;13(2):1-20. doi:10.1371/journal.pone.0192310
20. Steyerberg EW. *Clinical Prediction Models*. 2nd ed. Switzerland: Springer Nature; 2019.
21. Harrell FE. *Regression Modeling Strategies*. Switzerland: Springer International Publishing; Vol 45; 2003. doi:10.1198/tech.2003.s158
22. Abraham P, Gricar J, Zhang Y, Shankaran V. Real-world treatment patterns and outcomes in patients receiving second-line therapy for advanced/metastatic esophageal squamous cell carcinoma. *Adv Ther*. 2020;37(7):3392-3403. doi:10.1007/s12325-020-01394-y
23. Fuchs CS, Muro K, Tomasek J, et al. Prognostic factor analysis of overall survival in gastric cancer from two phase iii studies of second-line ramucirumab (REGARD and RAINBOW) using pooled patient data. *J Gastric Cancer*. 2017;17(2):132-144. doi:10.5230/jgc.2017.17.e16
24. Xu J, Li Y, Fan Q, et al. Clinical and biomarker analyses of sintilimab versus chemotherapy as second-line therapy for advanced or metastatic esophageal squamous cell carcinoma: a randomized, open-label phase 2 study (ORIENT-2). *Nat Commun*. 2022;13(1):1-12. doi:10.1038/s41467-022-28408-3

25. Shau HY, Kim A. Suppression of lymphokine-activated killer induction by neutrophils. *J Immunol*. 1988;141(12):4395-4402.
26. Gregory AD, Houghton AMG. Tumor-associated neutrophils: new targets for cancer therapy. *Cancer Res*. 2011;71(7):2411-2416. doi:10.1158/0008-5472.CAN-10-2583
27. An X, Ding PR, Li YH, et al. Elevated neutrophil to lymphocyte ratio predicts survival in advanced pancreatic cancer. *Biomarkers*. 2010;15(6):516-522. doi:10.3109/1354750X.2010.491557
28. Orditura M, Galizia G, Sforza V, et al. Treatment of gastric cancer. *World J Gastroenterol*. 2014;20(7):1635-1649. doi:10.3748/wjg.v20.i7.1635
29. Wesolowski R, Lee C, Kim R. Is there a role for second-line chemotherapy in advanced gastric cancer? *Lancet Oncol*. 2009;10(9):903-912. doi:10.1016/S1470-2045(09)70136-6
30. Sterne JAC, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009;338:b2393. doi:10.1136/bmj.b2393
31. Sperrin M, Martin GP. Multiple imputation with missing indicators as proxies for unmeasured variables: simulation study. *BMC Med Res Methodol*. 2020;20(1):1-11. doi:10.1186/s12874-020-01068-x
32. van Kleef JJ, Dijksterhuis WPM, van den Boorn HG, et al. Prognostic value of patient-reported quality of life for survival in oesophagogastric cancer: analysis from the population-based POCOP study. *Gastric Cancer*. 2021;24:1203-1212. doi:10.1007/s10120-021-01209-1

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kuijper SC, Pape M, Haj Mohammad N, van Voorthuizen T, Verhoeven RHA, van Laarhoven HWM. SOURCE beyond first-line: A survival prediction model for patients with metastatic esophagogastric adenocarcinoma after failure of first-line palliative systemic therapy. *Int J Cancer*. 2023;152(6):1202-1209. doi:10.1002/ijc.34385