

RESEARCH

Open Access



# Early prediction of response to palliative chemotherapy in patients with stage-IV gastric and esophageal cancer

Xiaoyuan Ma<sup>1</sup>, Eric Pierce<sup>2</sup>, Harsh Anand<sup>3</sup>, Natalie Aviles<sup>4</sup>, Paul Kunk<sup>2</sup> and Negin Alemazkooor<sup>3\*</sup>

## Abstract

**Background** The goal of therapy for many patients with advanced stage malignancies, including those with metastatic gastric and esophageal cancers, is to extend overall survival while also maintaining quality of life. After weighing the risks and benefits of treatment with palliative chemotherapy (PC) with non-curative intent, many patients decide to pursue treatment. It is known that a subset of patients who are treated with PC experience significant side effects without clinically significant survival benefits from PC.

**Methods** We use data from 150 patients with stage-IV gastric and esophageal cancers to train machine learning models that predict whether a patient with stage-IV gastric or esophageal cancers would benefit from PC, in terms of increased survival duration, at very early stages of the treatment.

**Results** Our findings show that machine learning can predict with high accuracy whether a patient will benefit from PC at the time of diagnosis. More accurate predictions can be obtained after only two cycles of PC (i.e., about 4 weeks after diagnosis). The results from this study are promising with regard to potential improvements in quality of life for patients near the end of life and a potential overall survival benefit by optimizing systemic therapy earlier in the treatment course of patients.

**Keywords** Metastatic cancer, Esophageal and gastric cancer, Response prediction, Machine learning, Palliative chemotherapy

## Introduction

It is established among medical professionals that what terminally ill patients need the most are: truth, touch, and time [1]. Specifically, they want their family and physicians to be truthful with them regarding their disease and its progress and treatments. Patients also want to be

touched and be reminded that they are loved and valuable [1, 2]. Most importantly, patients wish to have more time. They need time to accept their illness and losses and resolve various issues arising from their upcoming death [1]. Maintaining quality of life as symptom free as possible is essential to optimizing their remaining time [3]. Treatment with palliative chemotherapy (PC) for patients with metastatic gastric and esophageal cancers has traditionally been the standard of care. The treatment paradigm also includes Her-2 targeted therapy and immunotherapy based on PDL1 status. PC has resulted in relatively modest survival benefits with the risk of significant side effects including but not limited to hair loss, extreme weakness, nausea, and mucositis. Currently, there is no biomarker or other predictor for responses to

\*Correspondence:

Negin Alemazkooor  
na7fp@virginia.edu

<sup>1</sup> Department of Statistics, University of Virginia, Charlottesville, USA

<sup>2</sup> School of Medicine, University of Virginia, Charlottesville, USA

<sup>3</sup> System and Information Engineering, University of Virginia, Charlottesville, USA

<sup>4</sup> Department of Sociology, University of Virginia, Charlottesville, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

PC in patients with gastric and esophageal cancers [4]. In other words, only time shows whether PC is beneficial or not. In fact, the most common practice for evaluating the patients' response to chemotherapy is to perform imaging after approximately two to three months of the treatment. Signs of clinical improvement in addition to imaging surveillance are also used to assess for evidence of treatment response. Clinical signs of treatment benefit such as improved appetite, weight gain, improved energy, and less dysphagia are monitored. Although these signs are correlated with effectiveness of PC, one cannot guarantee the improved clinical signs will result in an overall survival improvement. Unfortunately, about 30% of the patients do not live more than three months after the initial diagnosis [5–7]. Consequently, early prognosis prediction for them is highly critical. In fact, many of these patients suffer from the toxicity of PC through their final days. If these patients had known their outcome ahead of time, they could have instead tried a potentially beneficial second-line treatment. Alternatively, they may have simply chosen to terminate the treatment and spend their remaining time more peacefully. The fact that these patients do not have enough time remaining for a trial and error approach to treatment makes early prognosis prediction for them crucial.

Several studies have attempted to improve our understanding of the prognosis of gastric and esophageal cancer. These studies mainly examine the correlation between survival time and patient or tumor characteristics, including but not limited to age [8], HER2 overexpression [9], sex [10], tumor size [11], metastasis sites [12–14], and laboratory variables [15]. Although these studies offer insights on what factors may lead to poor prognosis, they do not provide *response predictions* that can be used by oncologists or patients to make the decision to cease or continue PC. Few studies have attempted survival prediction for gastric and esophageal cancers, let alone cancers at stage-IV. In this work, we investigate the use of machine learning to predict survival for patients with stage-IV gastric and esophageal cancers. Machine learning tools, because of their ability to capture complex non-linearity within data [16], have proven to be capable in cancer risk prediction [17], detection [18, 19], classification [20, 21], and prognosis prediction [22, 23]. Nevertheless, only handful of studies have leveraged machine learning to predict prognosis in gastric and esophageal cancers. These studies mainly predict long-term survival (i.e., 3 or 5-year survival) based on serum markers, immunomarkers and clinicopathological parameters [24–27]. In fact, the literature lacks studies focusing on patients with the metastatic disease even though prognosis prediction is more vital for these patients given their limited time. Moreover, none of the existing research

examines the *response* to chemotherapy. Specifically, the available survival prediction models use the value of biomarkers at diagnosis time and do not consider their evolution after chemotherapy treatment.

This work addresses a major gap in the literature. It targets one of the most vulnerable groups of cancer patients and aims to improve the quality of their remaining life by predicting the response to chemotherapy and its benefit in terms of survival extension. To this end, we investigate the possibility of predicting survival at the time of diagnosis and after two cycles of PC treatment. To the authors' best knowledge, this is the first attempt in the literature to predict survival based on blood biomarkers for stage IV-patients in early stages of treatment.

We find that, with the information on metastasis sites and blood biomarkers at the time of diagnosis, machine learning models can predict whether a patient would survive beyond 6 month or 9 months, if they go through PC, with more than 75% accuracy. The accuracy of prognosis predictions increases to more than 85% when updated blood markers after two cycles of PC are included in the prediction models. Our hope is not only to leverage information about the evolution of blood markers for making early decisions on continuing, altering, or ceasing PC for end-stage gastric and esophageal cancers, but also to provide a blueprint for similar investigations in other types of cancers.

## Methods

### Data description

In this work, we used data from 150 patients with stage-IV gastric and esophageal cancers who were treated with standard PC at the University of Virginia Hospital over a ten year period from 2010 to 2020. For all patients, information on age, gender, body mass index (BMI), diabetes diagnosis, pathology and CT scan reports, as well as comprehensive metabolic panel and complete blood count panel tests (performed every two weeks before each chemotherapy treatment) are available. Specifically, the blood biomarkers included in this study are Hematocrit, albumin, platelets, WBC, creatinine, and total bilirubin. More specifically, we only included those blood biomarkers for which data for all patients was available. Other blood biomarkers such as CEA, CA19-9, and ctDNA were only available for small portion of patients and hence were excluded from the study. Additionally, pathology reports were searched to identify patients' Her-2 and PDL1 status. Lastly, data for the primary tumor location and metastasis sites were extracted from CT scan reports. Specifically, we defined binary variables to indicate whether the primary tumor was located in the esophagus or stomach (and when the binary variable is one for both, it indicates the existence of the tumor in

the junction). We also established four additional binary variables to determine whether cancer had spread to lymph nodes, liver, lungs, or unusual metastatic sites, which were included in the study. Notably, any metastatic site other than lymph nodes, liver, and lungs (e.g., brain or bone) was considered within the unusual metastatic site category. This categorization resulted from the relatively small number of patients with these uncommon metastatic sites, making it impractical to create separate variables for each one. Figure 1 shows a schematic view of different components of the dataset. The available bi-monthly lab data provides a unique opportunity to predict the prognosis given changes in blood markers after receiving PC for a short period. We have made the anonymous data set publicly available (see data availability statement), and interested readers can refer to it for more details about the data.

**Poor prognosis versus good prognosis**

The goal of therapy for the majority of Stage-IV gastric and esophageal cancer patients is non-curative. This raises the question of what should be considered as a good prognosis. Note that the expected overall survival for these patients is 4.3 months if they refuse PC treatment [28]. Figure 2 shows the distribution of overall survival of patients treated with PC. Even with PC treatment, a third of the patients will pass away during the first three months after diagnosis, half of patients die within six months, and only 37% patients survive beyond 9 months. We, therefore, developed different models that predict whether a patient would live beyond 6 or 9 months.

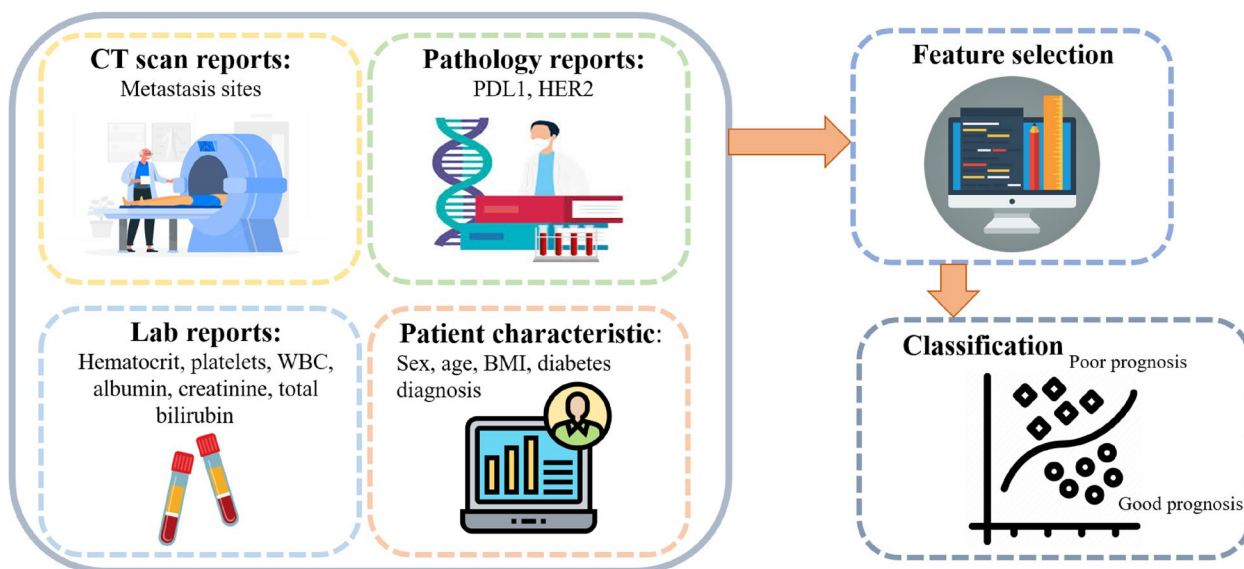
**Feature selection and prediction models**

Cox proportional hazards regression (Cox model) is the most widely used approach for survival analysis in cancer research [29–31]. Cox model investigates the association between covariates and the survival time of patients. Despite its ability to identify factors that influence survival time, the Cox model cannot directly predict survival time since it is a hazard model (i.e., it is not a prediction model). We exploit the Cox model for feature selection. Specifically, we only pass those covariates that are found to impact survival time to the machine learning models which are then trained to predict prognosis. In particular, we investigated five widely used machine learning classifiers: Random Forest, Support Vector Machine, Quadratic Discriminant Analysis, Naive Bayes, and Neural Networks. All classifiers were trained and calibrated using cross-validation with 75% of the data. The classifiers were tested on the remaining 25% of the data and their accuracy is reported in terms of Area Under the Curve (AUC), which is widely used as the accuracy measure in cancer research or any type of classification problem in general [32, 33].

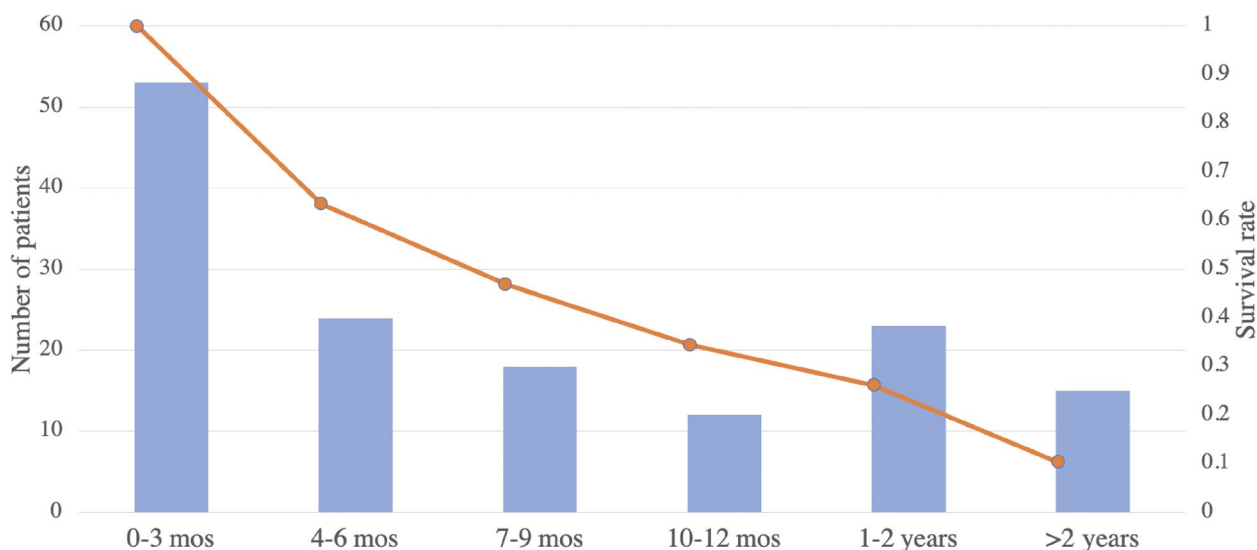
**Results**

**Selected features**

Tables 1 and 2 summarize the coefficients of Cox model for the selected features at the time of diagnosis and after two cycles of PC. Since the Cox model is a hazard model, an increment in covariates with positive coefficients is expected to increase hazard, thereby reducing survival time.



**Fig. 1** The schematic diagram of the data and framework used in this study



**Fig. 2** Overall survival of stage-IV gastric and esophageal cancer under PC treatment

**Table 1** Coefficients of Cox model at the time of diagnosis

Feature	Coef	z	p
Gender (0 for males, 1 for females)	0.399	1.894	0.058
Albumin at the time of diagnosis	-0.599	-3.648	<0.001
WBC at the time of diagnosis	0.076	3.094	0.002
With unusual metastasis sites	0.769	3.482	<0.001

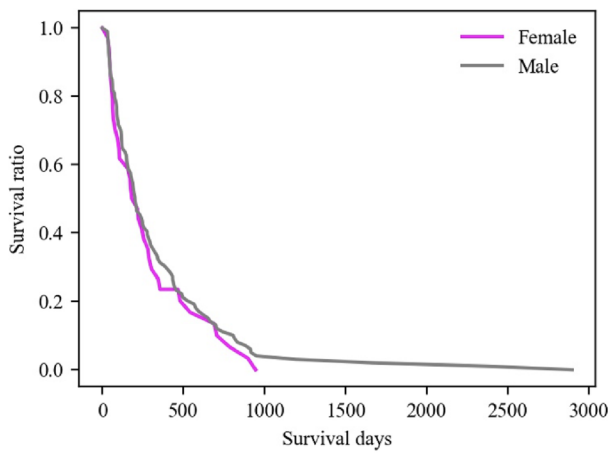
**Table 2** Coefficients of Cox model after two PC cycles

Feature	Coef	z	p
WBC after two PC cycles	0.093	4.13	<0.001
Albumin after two PC cycles	-1.035	-5.962	<0.001
Creatinine after two PC cycles	0.342	2.464	0.014
Total bilirubin after two PC cycles	0.098	1.895	0.058
With unusual metastasis sites	0.646	2.878	0.004

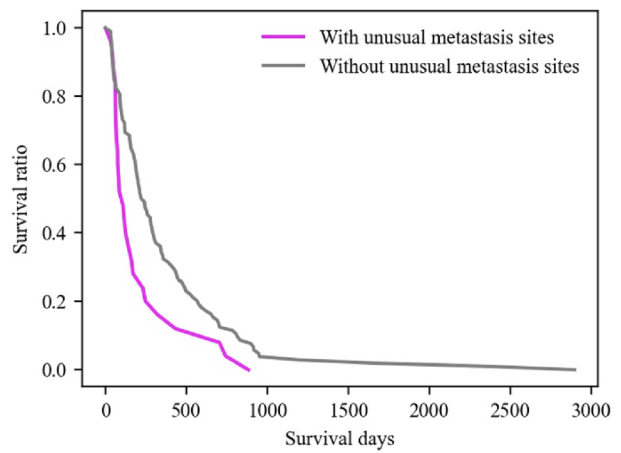
One interesting observation is that the location of the primary tumor is not among the variables that significantly impact survival. This finding indicates that patients with gastric and esophageal cancer have similar survival patterns, which are primarily determined by the value of blood biomarkers (that can be influenced by the location of metastasis sites). To provide an insight into the survival impact of each of the two binary covariates selected by the Cox model, Fig. 3 shows the survival plots for patients grouped on these binary variables. Specifically, Fig. 3a shows survival plots for males and females, with male patients having slightly better survival. The

observed slight difference, however, can be a result of small data size and may not be a general trend. Moreover, gender is not selected by the Cox model as a significant feature when blood markers after two PC cycles are included in the model. In other words, it is possible that the lack of enough significant variables in the initial Cox model has led to selection of Gender in the model. On the other hand, as shown in Fig. 3b, the survival plots for patients with and without unusual metastasis sites are significantly different. This is intuitive and in line with the literature.

Figure 4 depicts the box plots of WBC and albumin, the selected blood markers in the Cox model at the time of diagnosis, for patients with good and poor prognosis, where good prognosis indicates surviving longer than 9 months. Figure 5 depicts the box plots of WBC, albumin, creatinine and total bilirubin, the selected blood markers in the Cox model after 2 PC cycles, for patients with good and poor prognosis. Box plots show “minimum”, first quartile, median, third quartile, and “maximum” of blood markers in each group, where the “minimum” and “maximum” are defined as first quartile minus the 1.5 interquartile range and third quartile plus 1.5 interquartile range, respectively. Individual values of data are also shown along the box plots to provide an insight on the percentage of patients with normal and abnormal blood markers within the two groups with poor and good prognoses. It can be seen that the distribution of different blood markers is substantially different between patients with good and poor prognoses. Specifically, it is observed that larger percentages of those patients that survive less than 9 months have abnormal blood markers.

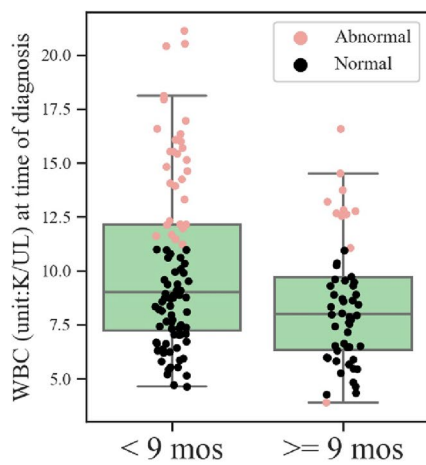


(a) Survival plots for males and females

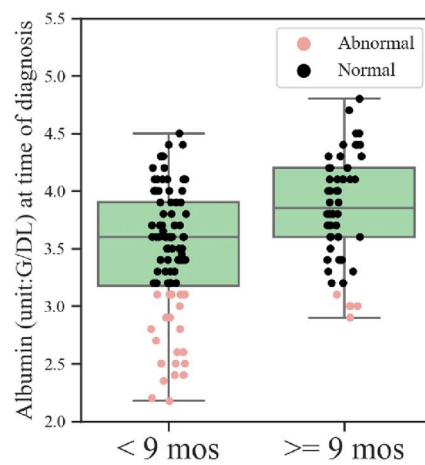


(b) Survival plots for patients with and without unusual metastasis sites

**Fig. 3** Survival plots of patients with stage-IV gastric and esophageal cancers categorized by gender and metastasis sites

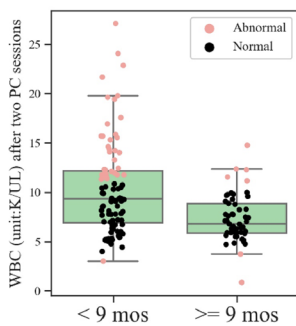


(a) WBC

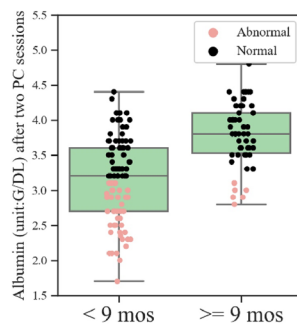


(b) Albumin

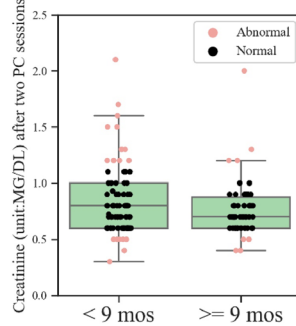
**Fig. 4** Distribution of WBC and Albumin for patients with poor and good prognoses at the time of diagnosis



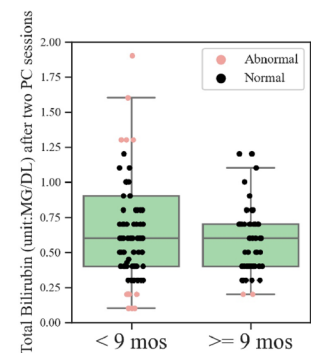
(a) WBC



(b) Albumin



(c) Creatinine



(d) Total Bilirubin

**Fig. 5** Distribution of blood markers for patients with poor and good prognoses after two PC cycles

For example, at the time of diagnosis, 32% of those with poor prognosis have abnormal Albumin while only 13% of patients with good prognosis have abnormal Albumin. Interestingly, it is observed that this difference becomes even larger after two PC cycles. Specifically, after two PC cycles, 52% and 11% of patients with poor and good prognoses have abnormal Albumin, respectively. Similar trend is observed for WBC. At the time of diagnosis, 32% of those with poor prognosis have abnormal WBC while only 18% of patients with good prognosis have abnormal WBC. This difference increases after two PC cycles leading to 37% and 12% of patients with poor and good prognoses having abnormal WBC, respectively.

**Prediction accuracy**

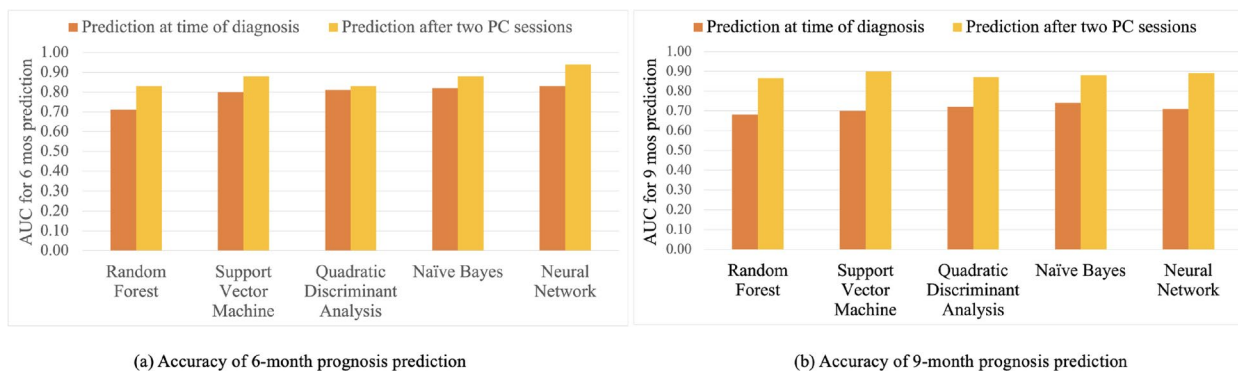
Looking at Figs. 4 and 5, it is clear that there is not a single blood marker that determines the prognosis outcome. We used machine learning approaches to investigate the possibility of predicting prognosis based on the collective information provided by different blood markers. Specifically, the machine learning models predict whether a patient would live beyond 6 or 9 months. Figure 6 shows the accuracy of different classifiers for prognosis prediction. For 6 month prognosis prediction, it can be seen that, even at the time of diagnosis, most machine learning models provide more than 80% accuracy. Once the values of blood markers after 2 PC cycles are included, the accuracy increases to more than 85%. Specifically, it is observed that, regardless of the type of machine learning algorithm that is used, prediction accuracy is improved with the inclusion of the values of blood markers after 2 PC cycles. Such improvement is even greater for the 9 month prognosis predictions. Overall, the results show that it is possible to predict, with accuracy above 85%, whether a patient would live beyond 6 or 9 months at very early stages of the treatment. This paves the way for more personalized treatments of end-stage cancer

patients, where no patients would suffer the demon of PC when it comes with no benefit.

**Discussion**

Gastric and esophageal cancers are the fifth and eighth most common cancers worldwide [34, 35]. The prognosis of these cancers are favorable only in the early stages. However, more than a third of patients are diagnosed with stage-IV disease [5] as these cancers are mostly asymptomatic during the early stages. Consequently, the overall 5-year observed survival rate for patients diagnosed with gastric and esophageal cancers are among the lowest for all cancers. Specifically, the 5-year survival rate for metastatic gastric and esophageal cancer is less than 5% [5, 6].

Considering the short life expectancy of stage-IV patients with gastric and esophageal cancers, PC is established as the standard first-line treatment. PC is associated with modest survival benefits and patient quality-of-life improvement [36]. However, the response to PC varies significantly among the patients. This means that not only are there patients that do not necessarily benefit from PC but also their quality of life deteriorates significantly because of the substantial and cumulative side effects of chemotherapy [37]. It is estimated that between 20% to 50% of terminally ill cancer patients undergo chemotherapy in the last thirty days of their lives without any clear benefits, and in many cases experience significant toxicities, financial costs, and decreased quality of life [38]. Moreover, receiving PC has been shown to be associated with higher rates of cardiopulmonary resuscitation and mechanical ventilation in the last week of life. Also, patients receiving PC were more likely to die in an intensive care unit rather than in their preferred place [39]. End-stage patients with gastric and esophageal cancers can, therefore, tremendously benefit from early predictions of their prognosis. Such a prediction can help the patients and their physicians in the timely



**Fig. 6** Accuracy of 6/9-month survival predictions at the time of diagnosis and after 2 PC cycles

termination of PC when it is not beneficial, thereby improving the patients' end-of-life quality. Patients and their care team may also leverage early prediction of response to PC to switch to a possibly more effective second-line treatment.

Unfortunately, despite the significance of implications of a reliable and easy-to-use response prediction model for patients with metastatic disease, few studies address this topic. First, the majority of available work concerning survival analysis investigates only the correlation between survival and potential biomarkers [8–10, 12–15, 40], without providing rigorous survival predictions. Second, studies that do attempt to predict survival do not focus on patients with metastatic disease [24–27]. The accuracy of a prediction model developed using data from patients in different stages are likely not sufficient for end-stage patients with shorter survival time. Moreover, none of the existing works has investigated the advantages of incorporating the changes in blood markers after a few cycles of PC in predicting the response to PC.

This work marks the first study predicting the response to PC for patients with metastatic gastric and esophageal cancer. Unique to this study, we use data from a comprehensive metabolic panel and complete blood count tests after two cycles of chemotherapy. We find that blood markers, including WBC, Albumin, Creatinine, and total Bilirubin after two cycles of chemotherapy are among predictors of prognosis. Different machine learning tools are able to provide highly accurate predictions (with AUC above 85% shown in Fig. 6) given the same predictors. The consistency of results among different machine learning models adds confidence in the true predictive ability of the input variables.

One primary advantage of the developed prediction model is its dependence solely on the metastasis sites and blood markers. Blood laboratory tests are performed regularly for patients undergoing chemotherapy (e.g., normally before each session of chemotherapy). The blood markers are, therefore, easily accessible and can be readily used in prediction models. Recently, estimating survival based on radiomic features from CT scans or tumor genomic profiles has garnered more attention for more common types of cancers (e.g., lung and breast cancer) [41–45]. This is despite the fact that both radiomic features and tumor genomics are not readily available. Specifically, extracting radiomic features (which contain distinct attributes associated with attenuation, shape, size, location, intensity, and texture of tumors) include two non-trivial steps: 1) segmenting the region of interest in CT images using manually delineated contours [46] or (semi) automatic packages such as 3DSlicer [47], 2) passing region of interest into software programs (e.g., TexRAD [48] and MaZda [49]) that extract the desired features. As for tumor genomic profile, the cost

of tumor genomic profiling can be up to \$100,000 per patient [50]. Consequently, specialized expertise and high monetary cost prohibit the wide utilization of prediction models based on radiomic or genomic features. This further highlights the importance of developing prediction models that are practical, given the availability and cost of existing technologies. Nevertheless, future research could explore the possibility of improving the prediction accuracy using radiomic or genomic features, in hope that extracting these features becomes less cumbersome and costly with advancements in technology.

We also want to call attention to the fact that it is very hard and time-consuming to obtain cancer patient data. There is limited availability of public sources for detailed cancer patient data (i.e., data with information beyond gender, sex, and age of patients). In fact, to author's best knowledge there is no public dataset that includes the gastric or esophageal patients' blood biomarker during their course of treatment. For this reason, the majority of studies in the literature that are focused on survival analysis or prediction use data from a specific hospital, which limits the sample size. The small sample size excludes the potential use of some machine learning algorithms (such as deep learning) that require extensive training with large datasets. Moreover, the private status of the data makes the validation or replication of the results impossible for other researchers. As a step forward to address this issue, we share the anonymous data used in this work on GitHub (see data availability statement) and encourage other researchers in the field to do the same in order to maximize the utility of the conducted research. Integration of datasets from various hospitals would provide innumerable opportunities to substantially advance the state-of-art in the field of cancer prognosis prediction.

Lastly, it must be noted that the data set used in this study lacked information about details of chemotherapy regimens. It might be possible to further improve the results by considering this additional information. Moreover, relatively recent evidence supports standard first-line treatment for stage IV esophageal and gastric cancer as combined chemotherapy and immunotherapy in subsets of patients based upon PDL1 status. Currently, we do not have enough data to evaluate the performance of predictions models in predicting survival for patients receiving chemoimmunotherapy. Nevertheless, our results show that it is possible to predict how long a patient would survive at very early stages of treatment. As a future research path, we will study data from various institutions and will develop prediction models for different treatments (i.e., PC +/- immunotherapy). Such models built on large multi-intuitions data for different types of treatment can tremendously help patients and their care team to make an informed decision about their choice of treatment, and its continuation or termination.

### Acknowledgements

We thank the Engineering in Medicine Program at the University of Virginia for facilitating collaboration between the engineering and medicine teams.

### Authors' contributions

P.K., E.P., N.Alemazkour and N.Aviles conceived the idea. N.Alemazkour and P.K. supervised the study. X.M. conducted the data cleaning, processing, and prediction modeling. E.P. led the data collection. H.A. helped in sense-making and writing the manuscript. All authors reviewed and edited the manuscript.

### Funding

Funding for this project was provided by the Engineering in Medicine Program at the University of Virginia.

### Availability of data and materials

The raw data (including pathology and radiology reports and bi-weekly laboratory results) is not publicly available. However, the cleaned and processed dataset used in this study is available on Github: [github.com/Elsixiaoyuan/gastric\\_esophageal\\_cancer\\_data](https://github.com/Elsixiaoyuan/gastric_esophageal_cancer_data). The code is available upon request.

### Declarations

#### Ethics approval and consent to participate

Data used in this study was anonymized before its use. Approval for this study has been obtained by the institutional review board for health sciences research at the University of Virginia on July 8, 2021 (IRB-HRS #23298).

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 17 April 2023 Accepted: 20 September 2023

Published online: 28 September 2023

### References

- Kuhl D, Stanbrook MB, Hébert PC. What people want at the end of life. *Can Med Assoc*; 2010.
- Kuhl D. What dying people want: Practical wisdom for the end of life. *PublicAffairs*; 2003.
- Nilsson J, Blomberg C, Holgersson G, Carlsson T, Bergqvist M, Bergström S. End-of-life care: where do cancer patients want to die? A systematic review. *Asia Pac J Clin Oncol*. 2017;13(6):356–64.
- Durães C, Almeida GM, Seruca R, Oliveira C, Carneiro F. Biomarkers for gastric cancer: prognostic, predictive or targets of therapy? *Virchows Arch*. 2014;464(3):367–78.
- Thrift AP, El-Serag HB. Burden of gastric cancer. *Clin Gastroenterol Hepatol*. 2020;18(3):534–42.
- Njei B, McCarty TR, Birk JW. Trends in esophageal cancer survival in United States adults from 1973 to 2009: a SEER database analysis. *J Gastroenterol Hepatol*. 2016;31(6):1141–6.
- Then EO, Lopez M, Saleem S, Gayam V, Sunkara T, Culliford A, et al. Esophageal cancer: an updated surveillance epidemiology and end results database analysis. *World J Oncol*. 2020;11(2):55.
- Saito H, Osaki T, Murakami D, Sakamoto T, Kanaji S, Tatebe S, et al. Effect of age on prognosis in patients with gastric cancer. *ANZ J Surg*. 2006;76(6):458–61.
- Gravalos C, Jimeno A. HER2 in gastric cancer: a new prognostic factor and a novel therapeutic target. *Ann Oncol*. 2008;19(9):1523–9.
- Tokunaga M, Ito S, Yoshikawa T, Nunobe S, Fukagawa T, Misawa K, et al. Prognostic factors for survival in patients with pT1 N+ or T2–3 N0 gastric cancer in Japan. *J Br Surg*. 2017;104(7):885–90.
- Adachi Y, Oshiro T, Mori M, Maehara Y, Sugimachi K. Tumor size as a simple prognostic indicator for gastric carcinoma. *Ann Surg Oncol*. 1997;4(2):137–40.
- Shiraishi N, Sato K, Yasuda K, Inomata M, Kitano S. Multivariate prognostic study on large gastric cancer. *J Surg Oncol*. 2007;96(1):14–8.
- Hosch SB, Stoecklein NH, Pichlmeier U, Rehders A, Scheunemann P, Nien-dorf A, et al. Esophageal cancer: the mode of lymphatic tumor cell spread and its prognostic significance. *J Clin Oncol*. 2001;19(7):1970–5.
- Greenstein AJ, Litle VR, Swanson SJ, Divino CM, Packer S, Wisnivesky JP. Prognostic significance of the number of lymph node metastases in esophageal cancer. *J Am Coll Surg*. 2008;206(2):239–46.
- Chau I, Norman AR, Cunningham D, Waters JS, Oates J, Ross PJ. Multi-variate prognostic factor analysis in locally advanced and metastatic esophago-gastric cancer—pooled analysis from three multicenter, randomized, controlled trials using individual patient data. *J Clin Oncol*. 2004;22(12):2395–403.
- Cruz JA, Wishart DS. Applications of machine learning in cancer prediction and prognosis. *Cancer Informat*. 2006;2:117693510600200030.
- Taninaga J, Nishiyama Y, Fujibayashi K, Gunji T, Sasabe N, Iijima K, et al. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Sci Rep*. 2019;9(1):1–9.
- Mccarthy JF, Marx KA, Hoffman PE, Gee AG, O'neil P, Ujwal ML, et al. Applications of machine learning and high-dimensional visualization in cancer detection, diagnosis, and management. *Ann N Y Acad Sci*. 2004;1020(1):239–62.
- Saba T. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *J Infect Public Health*. 2020;13(9):1274–89.
- Amrane M, Oukid S, Gagaoua I, Ensari T. Breast cancer classification using machine learning. In: 2018 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT). IEEE; 2018. p. 1–4.
- Tan AC, Gilbert D. Ensemble machine learning on gene expression data for cancer classification. *Appl Bioinformatics*. 2002;2(3):75–83.
- Mihaylov I, Nisheva M, Vassilev D. Application of machine learning models for survival prognosis in breast cancer studies. *Information*. 2019;10(3):93.
- Kourou K, Exarchos TP, Exarchos KP, Karamouzis MV, Fotiadis DI. Machine learning applications in cancer prognosis and prediction. *Comput Struct Biotechnol J*. 2015;13:8–17.
- Jiang Y, Xie J, Han Z, Liu W, Xi S, Huang L, et al. Immunomarker Support Vector Machine Classifier for Prediction of Gastric Cancer Survival and Adjuvant Chemotherapeutic Benefit Immunomarker SVM-Based Predictive Classifier. *Clin Cancer Res*. 2018;24(22):5574–84.
- Jiang Y, Liu W, Li T, Hu Y, Chen S, Xi S, et al. Prognostic and predictive value of p21-activated kinase 6 associated support vector machine classifier in gastric cancer treated by 5-fluorouracil/oxaliplatin chemotherapy. *EBioMedicine*. 2017;22:78–88.
- Que SJ, Chen QY, Liu ZY, Wang JB, Lin JX, Lu J, et al. Application of pre-operative artificial neural network based on blood biomarkers and clinicopathological parameters for predicting long-term survival of patients with gastric cancer. *World J Gastroenterol*. 2019;25(43):6451.
- Wang Y, Wang H, Li S, Wang L. Survival risk prediction of esophageal cancer based on the Kohonen network clustering algorithm and kernel extreme learning machine. *Mathematics*. 2022;10(9):1367.
- Wagner AD, Syn NL, Moehler M, Grothe W, Yong WP, Tai BC, et al. Chemotherapy for advanced gastric cancer. *Cochrane Database Syst Rev*. 2017;(8).
- Biglarian A, Hajizadeh E, Kazemnejad A, Zali M. Survival analysis of gastric cancer patients using Cox model: a five year study. *Tehran Univ Med J*. 2009;67.
- Moghimi Dehkordi B, Rajaeeferd A, Tabatabaee H, Zeighami B, Safaee A, Tabeie Z. Modeling survival analysis in gastric cancer patients using the proportional hazards model of Cox. *Iran J Epidemiol*. 2007;3(1):19–24.
- Abadi A, Yavari P, Dehghani-Arani M, Alavi-Majid H, Ghasemi E, Amanpour F, et al. Cox models survival analysis based on breast cancer treatments. *Iran J Cancer Prev*. 2014;7(3):124.
- Smoot BJ, Wong JF, Dodd MJ. Comparison of diagnostic accuracy of clinical measures of breast cancer-related lymphedema: area under the curve. *Arch Phys Med Rehabil*. 2011;92(4):603–10.



33. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med*. 2018;24(10):1559–67.
34. Herszenyi L, Tulassay Z. Epidemiology of gastrointestinal and liver tumors. *Eur Rev Med Pharmacol Sci*. 2010;14(4):249–58.
35. Wadhwa R, Song S, Lee JS, Yao Y, Wei Q, Ajani JA. Gastric cancer—molecular and clinical dimensions. *Nat Rev Clin Oncol*. 2013;10(11):643–55.
36. Hutt E, Da Silva A, Bogart E, Le Lay-Diemande S, Pannier D, Delaine-Clisant S, et al. Impact of early palliative care on overall survival of patients with metastatic upper gastrointestinal cancers treated with first-line chemotherapy: a randomised phase III trial. *BMJ Open*. 2018;8(1):015904.
37. Ooft SN, Weeber F, Dijkstra KK, McLean CM, Kaing S, van Werkhoven E, et al. Patient-derived organoids can predict response to chemotherapy in metastatic colorectal cancer patients. *Sci Transl Med*. 2019;11(513):2574.
38. Braga S. Why do our patients get chemotherapy until the end of life? *Ann Oncol*. 2011;22(11):2345–8.
39. Wright AA, Zhang B, Keating NL, Weeks JC, Prigerson HG. Associations between palliative chemotherapy and adult cancer patients' end of life care and place of death: prospective cohort study. *BMJ*. 2014;348:g1219.
40. Oyama K, Oba M, Oshima Y, Shimada H. Predicting short-term life expectancy of patients with end-stage gastric cancer using Onodera's prognostic nutritional index. *Int J Clin Oncol*. 2021;26(2):364–9.
41. van Timmeren JE, Leijenaar RT, van Elmpt W, Reymen B, Oberije C, Monshouwer R, et al. Survival prediction of non-small cell lung cancer patients using radiomics analyses of cone-beam CT images. *Radiother Oncol*. 2017;123(3):363–9.
42. Dietzel M, Schulz-Wendtland R, Ellmann S, Zoubi R, Wenkel E, Hammon M, et al. Automated volumetric radiomic analysis of breast cancer vascularization improves survival prediction in primary breast cancer. *Sci Rep*. 2020;10(1):1–11.
43. La Forgia D, Fanzini A, Campobasso F, Bellotti R, Didonna V, Lorusso V, et al. Radiomic analysis in contrast-enhanced spectral mammography for predicting breast cancer histological outcome. *Diagnostics*. 2020;10(9):708.
44. Sun D, Li A, Tang B, Wang M. Integrating genomic data and pathological images to effectively predict breast cancer clinical outcome. *Comput Methods Prog Biomed*. 2018;161:45–53.
45. Lopez-García G, Jerez JM, Franco L, Veredas FJ. Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data. *PLoS ONE*. 2020;15(3):0230536.
46. Coroller TP, Grossmann P, Hou Y, Velazquez ER, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol*. 2015;114(3):345–50.
47. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn Reson Imaging*. 2012;30(9):1323–41.
48. Chatwin C, Young R, Ganeshan B. TexRad-feedback plc - cancer management imaging software. Brighton (United Kingdom): University of Sussex, Industrial Informatics & Signal Processing Research Group; 2015.
49. Szczypinski PM, Strzelecki M, Materka A, Mazda-a software for texture analysis. In: 2007 international symposium on information technology convergence (ISITC 2007). IEEE; 2007. p. 245–9.
50. Dong OM, Poonnen PJ, Winski D, Reed SD, Vashistha V, Bates J, et al. Cost-Effectiveness of Tumor Genomic Profiling to Guide First-Line Targeted Therapy Selection in Patients With Metastatic Lung Adenocarcinoma. *Value Health*. 2022;25(4):582–94.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

