Published in final edited form as:

Eur J Cancer. 2017 May ; 76: 197–204. doi:10.1016/j.ejca.2017.02.008.

Improving Survival Prognostication of Gastroenteropancreatic Neuroendocrine Neoplasms: Revised Staging Criteria

Jacob A. Martin^a, Richard R.P. Warner, MD^a, Juan P. Wisnivesky, MD, DrPH^b, and Michelle Kang Kim, MD, PhD^a

^aDivision of Gastroenterology, Department of Medicine, Icahn School of Medicine at Mount Sinai, One Gustave L Levy Place, New York, NY 10029, United States of America

^bDivision of General Internal Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, One Gustave L Levy Place, New York, NY 10029, United States of America

Abstract

Purpose—Current staging criteria for gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs), while useful, have limitations. In this study, we used a population-based registry to evaluate the prognostic utility of the current staging systems and assess whether evidence-based modifications can improve survival predictions.

Methods—We identified patients with confirmed GEP-NENs from the Surveillance, Epidemiology, and End Results registry. We assigned TNM status according to American Joint Committee on Cancer and European Neuroendocrine Tumor Society criteria. We derived a revised staging classification using Kaplan-Meier methods and Cox regression to assess disease-specific survival and compared the accuracy of potential models based on the Akaike Information Criterion (AIC) and Harrell's C-index. The revised classification was validated in an independent set.

Results—We identified 10,268 patients with GEP-NENs. We found that multiple stages, as determined by current criteria, misclassified patients' prognosis. In particular, stage IIIB (T1-4N1) had overlapping survival with stage IIIA (T4N0). A revised system which reclassifies N1 disease into different stages based on T status (T1-2N1, T3N1, and T4N1) had an improved AIC (difference = 38) and C-index (0.86) compared to current staging. These revisions improved predictions in patients with both low and high grade tumors from all primary sites. Results also were confirmed across all primary sites in the validation set.

Conclusion—Current staging guidelines misclassify the prognosis of N1 patients. Our results suggest that a revised system could lead to better prognostication for GEP-NEN patients. Further validation followed by implementation of these revisions may improve treatment selection and design of clinical trials.

Corresponding Author: Michelle K. Kim, MD, PhD, The Mount Sinai Medical Center, One Gustave L. Levy Place, Box 1069, New York, NY, 10029, 212-241-9645, michelle.kim@mountsinai.org.

Conflicts of interest: none declared

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

Gastroenteropancreatic neuroendocrine tumors; Neuroendocrine tumor; Carcinoid; Prognosis; Staging; Survival Analysis

---

Introduction

Gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs) are relatively rare, currently affecting approximately 65,000 patients in the United States, but their prevalence has risen as abdominal imaging and endoscopy have become increasingly utilized.[1–4] GEP-NENs arise from primary sites such as the stomach, small intestine, pancreas, appendix, colon, and rectum. GEP-NENs encompass both well differentiated neuroendocrine tumors (NETs) and poorly differentiated neuroendocrine carcinomas (NECs), and they result in diverse outcomes that can be difficult to predict.

Prognostic staging of GEP-NENs is critical in recommending appropriate treatments, stratifying research participants, and counseling patients on the severity of their disease. In 2006, the European Neuroendocrine Tumor Society (ENETS) published the first consensus recommendations for foregut GEP-NENs based on single-center experiences.[5,6] In 2009, the 7th edition of the American Joint Committee on Cancer (AJCC) staging manual included neuroendocrine tumors for the first time.[7,8] The ENETS and AJCC stages are identical for tumors of the stomach, small intestine, colon, and rectum, but differ for the pancreas and appendix.[9]

Several groups have evaluated the prognostic accuracy of current GEP-NEN staging and found that both the European and American TNM classifications have stages that overlap and contain considerable variability in survival.[10–15] Given these limitations, recent work by our group has proposed revised staging criteria for small intestinal NENs which classifies node-positive patients differently according to tumor size and extent.[16,17] However, the staging for primary sites other than the small intestine has not been evaluated or revised.

In this study, we used population-based data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) registry to comprehensively evaluate the prognostic ability of the current American and European staging systems for all GEP-NENs and assessed whether refinement of these classifications can improve discrimination of patients' long-term survival.

Patients and Methods

Study Population and Variables

We used data from SEER, a population-based registry that collects cancer information covering approximately 30% of the United States population. SEER is subjected to extensive quality control measures including data audits and centralized monitoring.[18] Patients were included if they were over age 18 and diagnosed between 1988 and 2011 with pathologically confirmed GEP-NENs of the stomach, pancreas, appendix, colon, or rectum. We excluded

patients diagnosed by autopsy or using death certificate data as well as those with missing Tumor-Node-Metastasis (TNM) data.

We coded T, N, and M status according to both ENETS and AJCC specifications using SEER data that is based on a combination of clinical, operative, and pathological assessment. Tumor grade was assigned according to degree of differentiation, as defined in SEER using the following ICD-O-3 specifications: grade I, well differentiated; grade II, moderately differentiated; grade III, poorly differentiated; and grade IV, undifferentiated (anaplastic). SEER site-specific surgery codes were also used to define the extent of surgical resection (none, local, partial/simple, total or radical, or resection NOS).

Survival time was calculated as the period from the date of diagnosis until death. Those alive as of the last follow-up date reported in SEER (December 31, 2012) were treated as censored observations. GEP-NEN-specific survival was used as the primary endpoint for assessing cancer prognosis. Cause of death was determined by SEER using data abstracted from death certificates. Individuals who died from non-NEN causes or were alive at last contact were treated as censored observations. Overall survival was used as a secondary endpoint.

Evaluation of Current Staging Criteria

We divided the cohort into a training set containing 70% of the sample and a validation set including the remaining 30% of patients. Group assignment was conducted randomly with stratification by T, N, and M status, as well as primary site. The training set was used to evaluate the candidate staging models. The validation set was used to assess the validity of the best candidate model in an independent group of GEP-NEN patients.

Kaplan-Meier estimates of disease-specific survival and overall survival were plotted for the AJCC staging system. Luminal NENs (gastric, colon, and rectal NENs) were plotted as a single group because they share the same AJCC staging. Pancreatic and appendiceal NENs are staged according to different criteria and, thus, were plotted separately. Ten-year disease-specific survival was calculated for each stage, and log-rank tests were used to compare differences in survival. These analyses were repeated using ENETS TNM classifications. We evaluated our previously revised SI-NEN staging criteria against all models at every primary site.[16] We also plotted a Kaplan-Meier estimate of all node-positive (N1) patients stratified by T status to assess for potential interactions.

Assessment of Potential Candidate Models

We created a set of candidate staging models by reassigning T and N status into new stages. We first plotted Kaplan-Meier estimates for all combinations of T and N status (e.g., T1N0, T1N1, T2N0, etc.). Then, we used two rules in order to create rational stages for the candidate models: (1) only T and N groups with overlapping survival from the Kaplan-Meier could be combined into the same stage and (2) only adjacent curves could be combined into a single stage. We also constructed a saturated TNM model which included every combination of T and N status in a separate stage resulting in 9 different stages.

Derivation and Comparison of Potential Models

We formulated three criteria to select the staging system with the greatest prognostic accuracy. First, the best model would have a low Akaike Information Criterion (AIC) relative to the other models. The AIC is a measure of the predictive ability of a statistical model that allows models to be directly compared on the same data set. A lower AIC indicates a more explanatory model. The AIC difference (Δ) is the change in AIC between a candidate model and the model with the lowest AIC ($AIC_i - AIC_{min}$). A Δ greater than 10 indicates that the candidate model has considerably lower fit and should be discarded. The second criterion was that the best model would have an adequate Harrell's C-index, a measure of discrimination.[19] The third criterion was that Kaplan-Meier survival curves would show no overlap between different stages.

Using the above criteria, we evaluated all candidate models and selected the model with the lowest AIC as the "best" model. Then, we compared the best model against the AJCC and ENETS staging system separately for each primary site. We performed similar analyses on the subset of patients with available tumor grade, comparing the best model to the AJCC staging system separately in grade I–II and grade III–IV tumors. Additionally, the best candidate model, a six-stage model, was also compared to models with five, seven, and eight stages. A six-stage model, for example, means a model with six possible TNM classifications.

Validation in an Independent Data Set

Once the final model was selected, we validated its performance in an independent validation set. To assess the validity of the model, we compared the best model to other potential candidate models as well as the AJCC and ENETS systems. We compared the AIC, C-index, and performance of the Kaplan-Meier survival curves.

Results

We identified 25,952 patients with a GEP-NEN of the stomach, pancreas, colon, appendix, and rectum (Figure 1). A total of 15,684 patients were excluded due to incomplete T, N, and M data; lack of pathologic confirmation; or otherwise not fulfilling the inclusion criteria. Our final cohort consisted of 10,268 patients with GEP-NENs originating from the stomach (12% of the total), pancreas (19%), appendix (12%), colon (17%), and rectum (40%). The derivation and validation cohorts were equivalent with respect to sociodemographic and treatment characteristics (Table 1).

Kaplan-Meier survival curves, including all T, N, and M combinations as well as those for the AJCC and ENETS classifications, showed multiple overlaps across all sites. In luminal NENs, patients with AJCC stage IIIA (T4N0) and IIIB (T1-4N1) had comparable disease-specific survival (10-year DSS: 46%, 95% confidence interval [CI]: 34–58% vs. 53%, 95% CI: 48–58%, respectively; $p = 0.25$; Figure 2A). Additionally, patients with stage IIA (T3N0) and IIB (T1-3N1) pancreatic NENs had comparable DSS (10-year rates of 43%, 95% CI: 32–54% vs. 48%, 95% CI: 37–59%, respectively; $p = 0.09$; Figure 2B). In contrast to all other GEP-NEN staging systems, appendiceal NENs are classified into four AJCC stages

and showed three clustered groups with high survival and the fourth with low survival without an intermediate stage (Figure 2C). We found similar overlaps and limitations in ENETS staging.

In node-positive and node-negative patients, survival was worse as T status increased ($p < 0.001$) (Figure 3A, B). In patients with node-positive GEP-NENs from all sites, T3 had worse survival than T2 (hazard ratio [HR]: 1.9, 95% CI: 1.3–2.8) and T4 had worse survival than T3 (HR: 1.8, 95% CI: 1.4–2.3). However, T1 and T2 cases had similar survival (HR: 1.1, 95% CI: 0.5–2.5).

We constructed disease-specific survival curves for all combinations of T and N status. According to our pre-specified criteria, we then created sixteen candidate staging models. In luminal GEP-NENs, a single stage grouping, hereafter referred to as the ‘best’ model, had the lowest AIC in unadjusted and adjusted Cox models (Figure 4; Table 2). The original AJCC staging had a χ^2 of 38, indicating inferior fit compared to the best model. The C-index of the best model was 0.90, confirming its highly accurate predictive ability. The best model, containing six stages, had a comparable AIC to the fully saturated TNM model and models with seven or eight stages, despite being more parsimonious. Additionally, the best 5-stage model had a χ^2 of 20, indicating worse fit. Overall, these results supported the selection of a 6-stage classification.

In secondary analyses, we assessed the best model in the subset of luminal NENs with tumor grade data. The best model retained its prognostic accuracy for all tumor grades; that is, the best model was valid for well and intermediately differentiated tumors (C-index: 0.87, χ^2 6) and for poorly differentiated and undifferentiated tumors (C-index: 0.70, χ^2 3).

In site-specific analyses, the best model was superior to the AJCC criteria with χ^2 s of 5, 23, 28, and 16 in stomach, appendix, colon, and rectum, respectively. The best model also performed better than the ENETS criteria. In pancreatic NENs, the best model had a χ^2 of 1. The best model had a high predictive ability at all sites with a C-index of 0.86 in the stomach, 0.73 in the pancreas, 0.82 in the appendix, 0.78 in the colon, and 0.92 in the rectum. We also evaluated our previously described revised staging classification for small intestinal NENs.[16] In most primary sites, the best model had a better AIC than the small intestinal classification. However, in the small bowel, the revised small intestinal staging system outperformed the the AJCC staging with a χ^2 of 20 and the best model from this study with a χ^2 of 17. This indicates that the best model is a superior fit for the stomach, pancreas, appendix, colon, and rectum, while our previously described small intestinal NEN staging criteria perform well specifically in the small intestine.

Kaplan-Meier plots of the best model for all GEP-NENs showed adequate separation between different stages and lack of overlap between groups ($p < 0.0001$; Figure 5A). Ten-year DSS was 96% (95% CI: 95–99%) for stage IA, 93% (95% CI: 91–95%) for stage IB, 66% (95% CI: 58–72%) for stage II, 55% (95% CI: 47–62%) for stage IIIA, 39% (95% CI: 28–50%) for stage IIIB and 23% (95% CI: 15–32%) for stage IV. Similar results were obtained in analyses assessing overall survival.

In the validation set, the best model was superior to AJCC and ENETS staging. The best model had a lower AIC than the AJCC staging in the stomach, appendix, colon, and rectum with Δ AICs of 3–8 in each site. In the pancreas, the best model had a Δ AIC of 4 compared to AJCC staging. The C-index of the best model in GEP-NENs from all primary sites was 0.86 in the validation set and ranged from 0.75–0.96 in the different primary sites. In all primary sites, Kaplan-Meier plots for the best model were similar between the training and validation sets and stages showed clear differences in survival (Figure 5B). The best model was also prognostic for well and intermediately differentiated GEP-NENs in the validation set with a C-index of 0.80.

Discussion

GEP-NENs have heterogeneous outcomes that can be difficult to predict in clinical practice. Validation studies have shown overlaps among different AJCC and ENETS stages as well as heterogeneity in the survival of patients within single groups.[11,14,16] In this study, we used objective criteria to derive and validate a robust staging system for GEP-NENs. Our findings suggest that the revised system performs well among tumors of multiple sites and may help improve prognostication.

Existing staging guidelines for most GEP-NENs group all N1 disease into the same stage despite substantial differences in survival among node positive patients. In this study, we showed that T status allows for improved classification of patients with N1 disease. Our best model separates N1 patients into three different stages leading to more prognostic accuracy than current staging. The revised staging classification resolved all overlaps in the AJCC and ENETS stages in both the derivation and validation sets. These findings strongly suggest the need for revising the grouping of patients with lymph node involvement.

Strengths of our study include the use of data from a population-based sample of over 10,000 GEP-NENs patients from multiple geographic regions in the United States, ensuring high generalizability. The duration of follow-up available in SEER was also advantageous to study long-term outcomes of GEP-NENs which can be generally indolent tumors. Another strength of our design was the use of objective criteria to develop and validate staging systems. This allowed us to systematically compare models in order to identify the most accurate prognostic classification for GEP-NENs.[20] Finally, our results were assessed in a independent cohort of patients enhancing their external validity to future patients. Finally, we developed uniform staging systems that can be applied to most GEP-NEN sites and facilitate adoption in clinical practice.

One study weakness was the exclusion of patients with missing TNM information in the SEER registry. Additionally, we were unable to specify the sizes and anatomic sites of lymph node metastases in order to analyze the heterogeneity of survival in N1 patients. We also had limited information on tumor grade. Current WHO and ENETS systems rely on the mitotic count or the Ki-67 index to differentiate low, intermediate, and high-grade GEP-NENs; SEER does not provide tumor grading according to this classification system.[5,6,21] However, SEER does provide degree of tumor differentiation, and we used this data as our measurement of grade to show that the proposed classification remained accurate after

controlling for this prognostic factor. Future studies would be needed to understand how the combination of tumor stage and grade can help predict outcomes in GEP-NEN patients.

In conclusion, our findings suggest that potential revisions to the current GEP-NEN staging guidelines, particularly in the grouping of N1 patients, may improve survival predictions. While additional studies should confirm these results in other patient cohorts, adoption of the revised guidelines may help in the clinical management of these patients as well as inform future clinical trials of emerging GEP-NENs therapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Sources of support: This work was supported by the American Cancer Society (MRS-14-014-01-CCE to M.K.K.) and the National Center for Advancing Translational Sciences of the National Institutes of Health (award number TL1TR001434 to J.M.).

References

1. Lawrence B, Gustafsson BI, Chan A, Svejda B, Kidd M, Modlin IM. The epidemiology of gastroenteropancreatic neuroendocrine tumors. *Endocrinol Metab Clin North Am*. 2011; 40:1–18. vii. DOI: 10.1016/j.ecl.2010.12.005 [PubMed: 21349409]
2. Tsikitis VL, Wertheim BC, Guerrero MA. Trends of incidence and survival of gastrointestinal neuroendocrine tumors in the United States: a seer analysis. *J Cancer*. 2012; 3:292–302. [PubMed: 22773933]
3. Fraenkel M, Kim MK, Faggiano A, Valk GD. Epidemiology of gastroenteropancreatic neuroendocrine tumours. *Best Pract Res Clin Gastroenterol*. 2012; 26:691–703. DOI: 10.1016/j.bpg.2013.01.006 [PubMed: 23582913]
4. Yao JC, Hassan M, Phan A, Dagohoy C, Leary C, Mares JE, et al. One Hundred Years After “Carcinoid”: Epidemiology of and Prognostic Factors for Neuroendocrine Tumors in 35,825 Cases in the United States. *J Clin Oncol*. 2008; 26:3063–72. DOI: 10.1200/JCO.2007.15.4377 [PubMed: 18565894]
5. Rindi G, Klöppel G, Alhman H, Caplin M, Couvelard A, de Herder WW, et al. TNM staging of foregut (neuro)endocrine tumors: a consensus proposal including a grading system. *Virchows Arch*. 2006; 449:395–401. DOI: 10.1007/s00428-006-0250-1. [PubMed: 16967267]
6. Rindi G, Klöppel G, Couvelard A, Komminoth P, Körner M, Lopes JM, et al. TNM staging of midgut and hindgut (neuro) endocrine tumors: a consensus proposal including a grading system. *Virchows Arch*. 2007; 451:757–62. [PubMed: 17674042]
7. Bilimoria KY, Bentrem DJ, Merkow RP, Tomlinson JS, Stewart AK, Ko CY, et al. Application of the pancreatic adenocarcinoma staging system to pancreatic neuroendocrine tumors. *J Am Coll Surg*. 2007; 205:558–63. DOI: 10.1016/j.jamcollsurg.2007.05.009 [PubMed: 17903729]
8. Compton CC, Byrd DR, Garcia-Aguilar J, Kurtzman SH, Olawaiye A, Washington MK. *AJCC cancer staging atlas: a companion to the seventh editions of the AJCC cancer staging manual and handbook*. Springer Science & Business Media; 2012.
9. Klöppel G, Rindi G, Perren A, Komminoth P, Klimstra DS. The ENETS and AJCC/UICC TNM classifications of the neuroendocrine tumors of the gastrointestinal tract and the pancreas: a statement. *Virchows Arch*. 2010; 456:595–7. [PubMed: 20422210]
10. Fischer L, Kleeff J, Esposito I, Hinz U, Zimmermann A, Friess H, et al. Clinical outcome and long-term survival in 118 consecutive patients with neuroendocrine tumours of the pancreas. *Br J Surg*. 2008; 95:627–35. DOI: 10.1002/bjs.6051 [PubMed: 18306152]

11. Strosberg JR, Cheema A, Weber J, Han G, Coppola D, Kvols LK. Prognostic validity of a novel American Joint Committee on Cancer Staging Classification for pancreatic neuroendocrine tumors. *J Clin Oncol Off J Am Soc Clin Oncol*. 2011; 29:3044–9. DOI: 10.1200/JCO.2011.35.1817
12. Jann H, Roll S, Couvelard A, Hentic O, Pavel M, Müller-Nordhorn J, et al. Neuroendocrine tumors of midgut and hindgut origin: Tumor-node-metastasis classification determines clinical outcome. *Cancer*. 2011; 117:3332–41. DOI: 10.1002/cncr.25855 [PubMed: 21246527]
13. Strosberg JR, Cheema A, Weber JM, Ghayouri M, Han G, Hodul PJ, et al. Relapse-free survival in patients with nonmetastatic, surgically resected pancreatic neuroendocrine tumors: an analysis of the AJCC and ENETS staging classifications. *Ann Surg*. 2012; 256:321–5. DOI: 10.1097/SLA.0b013e31824e6108 [PubMed: 22415420]
14. Rindi G, Falconi M, Klersy C, Albarello L, Boninsegna L, Buchler MW, et al. TNM staging of neoplasms of the endocrine pancreas: results from a large international cohort study. *J Natl Cancer Inst*. 2012; 104:764–77. DOI: 10.1093/jnci/djs208 [PubMed: 22525418]
15. Strosberg JR, Weber JM, Feldman M, Coppola D, Meredith K, Kvols LK. Prognostic Validity of the American Joint Committee on Cancer Staging Classification for Midgut Neuroendocrine Tumors. *J Clin Oncol*. 2013; 31:420–5. DOI: 10.1200/JCO.2012.44.5924 [PubMed: 23248248]
16. Kim MK, Warner RRP, Roayaie S, Harpaz N, Ward SC, Itzkowitz S, et al. Revised Staging Classification Improves Outcome Prediction for Small Intestinal Neuroendocrine Tumors. *J Clin Oncol*. 2013; JCO.2013.51.1477. doi: 10.1200/JCO.2013.51.1477.
17. Kim MK, Warner RRP, Ward SC, Harpaz N, Roayaie S, Schwartz ME, et al. Prognostic Significance of Lymph Node Metastases in Small Intestinal Neuroendocrine Tumors. *Neuroendocrinology*. 2015; doi: 10.1159/000371807.
18. [accessed January 18, 2017] Quality Improvement - About SEER. n.d. <https://seer.cancer.gov/about/uses.html>
19. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996; 15:361–87. DOI: 10.1002/(SICI)1097-0258(19960229)15:4<361::AID-SIM168>3.0.CO;2-4 [PubMed: 8668867]
20. Anderson DR, Burnham KP. Avoiding pitfalls when using information-theoretic methods. *J Wildl Manag*. 2002:912–8.
21. Bosman FT, Carneiro F, Hruban RH, Theise ND, et al. WHO classification of tumours of the digestive system. World Health Organization; 2010.

Highlights

- Current staging for gastroenteropancreatic neuroendocrine neoplasms (GEP-NENs) misclassifies the prognosis of patients with lymph node metastasis.
- A revised staging system for GEP-NENs was developed systematically based on a large, population based cohort.
- This staging system was validated in an independent data set, and it is capable of improving survival predictions in GEP-NEN patients.

Figure 1.
Inclusion of patients in the study

Figure 2.

Disease-specific survival according to the current AJCC staging system. In luminal NENs (A), there was overlap of patients with stage IIIA (T4N0) and IIIB (T1-4N1) disease ($p=0.25$). In pancreatic NENs (B), patients with stage IIA (T3N0) and IIB (T1-3N1) overlapped ($p=0.09$). In appendiceal NENs (C), there are three groups are clustered with high survival and a fourth with low survival without an intermediate stage.

Figure 3.
Kaplan-Meier estimates of survival for N0 (A) and N1 (B) GEP-NENs from all sites according to T status. Survival is worse as T status increases ($P < 0.001$).

Figure 4.

AIC differences (Δ) among staging models in luminal NENs. One stage grouping, the best model, had the lowest AIC (Δ of 0, indicated by *). AJCC staging had a Δ of 38, and the fully saturated TNM model had a Δ of 45.

Figure 5.

Kaplan-Meier estimates of survival for the best model for all GEP-NENs in the derivation set (A) and validation set (B). Disease-specific survival was different between all stages ($p < 0.001$).

Table 1

Demographics and Clinical Characteristics

Baseline demographic characteristics

Characteristic	Training Set	Validation Set
Participants, n (%)	7,226 (70%)	3,042 (30%)
Age (years)		
Diagnosis, mean $\pm$ SD	57 $\pm$ 14	56 $\pm$ 14
Follow-up (years)		
Duration, median (IQR)	3.4 (6.8)	3.6 (6.8)
Female, n (%)	3,723 (52%)	1,625 (53%)
Marital status, n (%)		
Married	4,299 (59%)	1,792 (59%)
Other	2,927 (41%)	1,250 (41%)
Primary site, n (%)		
Stomach	884 (12%)	364 (12%)
Pancreas	1,367 (19%)	579 (19%)
Appendix	868 (12%)	362 (12%)
Colon	1,214 (17%)	509 (17%)
Rectum	2,893 (40%)	1,228 (40%)
T status, n (%)		
T1	3,076 (43%)	1,306 (43%)
T2	1,897 (26%)	799 (26%)
T3	1,485 (21%)	622 (20%)
T4	768 (11%)	315 (10%)
N status, n (%)		
N0	5,739 (79%)	2,431 (80%)
N1	1,487 (21%)	611 (20%)
M status, n (%)		
M0	6,399 (89%)	2,717 (89%)
M1	827 (11%)	325 (11%)
Surgery, n (%)		
Local	2,968 (41%)	1,269 (42%)
Partial/simple	2,468 (34%)	1,046 (34%)
Total or radical	732 (10%)	315 (10%)
Resection, NOS	85 (1%)	31 (1%)
None	973 (13%)	381 (13%)
Grade, n (%)		

Characteristic	Training Set	Validation Set
Well-differentiated	1,672 (23%)	732 (24%)
Moderately differentiated	494 (7%)	207 (7%)
Poorly differentiated	640 (9%)	237 (8%)
Undifferentiated	181 (2%)	63 (2%)
Not graded	4,239 (59%)	1,803 (59%)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2

Definition of stages in the best model. N1 patients have been separated into stage II (T1-2), IIIA (T3), and IIIB (T4).

Best Model	
IA	T1N0
IB	T2N0
II	T3N0 T1N1 T2N1
IIIA	T4N0 T3N1
IIIB	T4N1
IV	M1