

**Title****Incorporating apical lymph node status into the seventh edition of TNM classification improves prognosis prediction in stage III colon cancer: A multicenter cohort study in Japan**

*Hironori Kawada (HK), MD<sup>1\*</sup>; Noriaki Kurita (NK), MD, FJSIM<sup>2\*</sup>; Fumiaki Nakamura (FN), MD, PhD<sup>3</sup>; Junichiro Kawamura (JK), MD, PhD<sup>1</sup>; Suguru Hasegawa (SH), MD, PhD<sup>1</sup>; Kenjiro Kotake (KK), MD, PhD<sup>4</sup>; Kenichi Sugihara (KS), MD, PhD<sup>5</sup>; Shunichi Fukuhara (SF), MD, FACP, DMSc<sup>2,6</sup>; and Yoshiharu Sakai (YS), MD, PhD<sup>1</sup>*

Department of Surgery, Graduate School of Medicine, Kyoto University, Kyoto, Japan<sup>1</sup>

Department of Healthcare Epidemiology, School of Public Health in the Graduate School of Medicine, Kyoto University, Kyoto, Japan<sup>2</sup>

Department of Public Health, The University of Tokyo Graduate School of Medicine, Tokyo, Japan<sup>3</sup>

Department of Surgery, Tochigi Cancer Center, Utsunomiya, Tochigi, Japan<sup>4</sup>

Department of Surgical Oncology, Graduate School, Tokyo Medical and Dental University, Tokyo, Japan<sup>5</sup>

Center for Innovative Research for Communities and Clinical Excellence (CIRC<sup>2</sup>LE), Fukushima Medical University, Fukushima, Japan<sup>6</sup>

\*H.K. and N.K. contributed equally to this work.

**Address correspondence to:** Shunichi Fukuhara, MD, FACP, DMSc

Department of Healthcare Epidemiology

School of Public Health in the Graduate School of Medicine

Kyoto University

Yoshida Konoe-cho, Sakyo-ku

Kyoto 606-8501, Japan

**TEL:** 81-75-753-4646

**FAX:** 81-75-753-4644

**E-mail:** fukuhara.shunichi.6m@kyoto-u.ac.jp

**Running title:** apical lymph node metastasis in colon cancer

**Sources of funding for research and/or publication:** none.

**Article type:** Original article.

**Whether the paper is based on a previous communication to a society or meeting? :** No.

## **Abstract**

**Background:** The node classification outlined in the seventh edition of the tumor-node-metastasis (TNM) classification is based solely on the number of metastasized lymph nodes. We examined the prognostic value of apical lymph node metastasis and the additive value of incorporating apical lymph node status into a risk model based on the seventh edition.

**Methods:** Our cohort study involved 1355 patients with Stage III colon cancer who underwent tumor resection with dissection of regional (including apical) lymph node at 71 hospitals across Japan between 2000 and 2002. The main exposure was pathologically confirmed apical lymph node metastasis, and the primary endpoint was cancer-specific death.

**Results:** Apical lymph node metastasis was present in 113 (8.3%) of the patients. During 5,356 patient-years (median 5.0 year) of follow-up, 221 (16.3%) instances of cancer-specific death were observed. After adjustment for tumor and node classification (as described in the seventh edition) and other prognostic factors, apical lymph node metastasis was found to be independently associated with cancer-specific death (hazard ratio, 2.29; 95% confidence interval [CI], 1.49-3.52). Incorporating apical lymph node metastasis into the prognostic model based on the seventh TNM edition significantly improved discriminative performance for cancer-specific death (difference in concordance index, 0.0146; 95% CI, 0.003-0.026) and risk reclassification for cancer-specific death at 5 years (category free net reclassification improvement, 19.4%; 95% CI, 5.0%-33.4%), respectively.

**Conclusions:** Assessment of apical lymph node metastasis provided independent prognostic information beyond that achievable with the seventh TNM edition in patients with Stage III colon cancer.

## **Introduction**

Accurate prognostic estimation of colorectal cancer is important for facilitating accurate decision-making in oncological management. The tumor-node-metastasis (TNM) classification developed by the American Joint Committee on Cancer (AJCC) and the Union for International Cancer Control (UICC) is the most commonly used staging system worldwide. Under this system, lymph node staging is defined based only on the number of metastatic lymph nodes,<sup>1</sup> representing an extremely simple way of classifying patients and saving surgeons or pathologists the trouble of classifying retrieved lymph nodes by anatomical location, such as paracolic area, intermediate area, or apical area. However, anatomical location of lymph nodes remains an important point from a prognostic perspective, especially given that apical lymph nodes (ALNs) are considered the entry point for lymph node metastases further proximal to the aorta (such as the para-aortic lymph node) or for hematogenous metastasis. As such, although ALN metastases have high potential utility as a marker for prediction of future systemic metastasis and subsequent mortality, whether or not ALN metastasis has such prognostic value remains unclear.

Several previous studies support the notion that the location of the metastasized lymph node has prognostic value. For example, presence of inferior mesenteric artery lymph node metastasis has been shown to predict para-aortic nodal recurrence in sigmoid colon and rectal cancer patients.<sup>2</sup> Other studies have shown distribution of lymph node metastasis to be an independent predictor of overall survival for sigmoid colon and rectal cancer.<sup>3,4</sup> However, these previous studies were limited by the disease being restricted to a partial area of the colon and by being single-center studies. Further, these studies failed to show the value of incorporating the location of the metastasized lymph node into existing TNM classification-based prediction models.

Here, using Japanese multicenter registry data, we examined the association between ALN metastasis with cancer-specific death among colon cancer patients as well as assessed whether or not the addition of ALN metastasis to the prognostic model based on the seventh edition of the UICC TNM staging system (UICC7) improves discrimination of cancer-specific death.

## Method

We obtained approval to use the registry data of Japanese Society for Cancer of the Colon and Rectum (JSCCR) in our cohort study. The JSCCR registry project was approved by the institutional review board of Tochigi Cancer Center. Patients diagnosed with colorectal cancer who underwent surgery at 89 member hospitals across different areas in Japan between January 2000 and December 2002 were voluntarily registered in the JSCCR registry. This nationwide registry is hospital-based (not population-based) and covers approximately 6.5% of all patients with colorectal cancer in Japan between 2000 and 2002. Eligibility criteria for this study were (1) patients with colon cancer who underwent curative resection with dissection of regional lymph node including ALN, (2) pathologically diagnosed Stage III with colon cancer, and (3) no distant metastasis. To maintain comparability between patients with or without ALN metastasis, neither Stage 0, I, nor II was included in the present study, as no patients in any of those Stages have ALN metastasis. Whether or not distant metastasis is present is usually examined using a combination of chest X-ray, abdominal ultrasound, and abdominal (and chest) computed tomography.<sup>5</sup> Exclusion criteria were (1) patients who underwent preoperative treatment or (2) patients with multiple primary cancers. We also excluded patients with rectal cancer, which was defined as cancer of the bowel below the lower margin of the second sacral vertebra,<sup>6</sup> as these patients have two main lymphatic drainage routes. Specifically, both routes continuing to the root of the inferior mesenteric arteries and the internal iliac arteries are involved in lymphatic spread. To correctly determine ALN metastasis, bilateral lymph node dissection surrounding the root of the internal iliac arteries is required but is not always performed in actual practice, as it is often complicated by voiding or sexual dysfunction. As informed consent to use existing registry data is not mandatory for observational studies in Japan, such consent was not obtained for this study.

### *Lymph node staging and apical lymph node metastasis*

For the JSCCR registry, data on number and anatomical location of positive lymph node were collected, while only the anatomical location of positive lymph nodes had been used for staging according to the Japanese Classification of Colorectal Carcinoma between 2000 and 2002. For the present study, using data on the number of positive lymph nodes, lymph node staging was defined according to the UICC version 7 (UICC7) and classified into four categories: N1a, N1b, N2a, and N2b.<sup>1</sup> The category of N1c was excluded because tumor deposit was not registered in the Japanese Classification system at that time. The definition of ALNs corresponds to the N3 station in the Japanese Classification of Colorectal Carcinoma, which locates the roots of major arteries, such as the iliocecal artery, right colic artery, middle colic artery, or inferior mesenteric artery.<sup>6</sup> These lymph node metastases were considered to be present if the retrieved lymph nodes were found to be metastasized pathologically. In Japan, well-trained surgeons usually perform regional lymph node dissection involving dissection of ALN, and surgeons in the theatre assign lymph node stations during surgery or postoperatively, in accordance with rules of Japanese classification of colorectal cancer.<sup>6</sup> Fat clearance technique is not generally performed.

### *Cancer-specific death*

Cause of death was categorized as defined by the JSCCR as death related to surgery, death specifically related to colon cancer, death related to other primary cancer, death not related to cancer, death of unknown cause. This outcome ascertainment is routine with the JSCCR registry and was performed independent of the study. Death specifically related to colon cancer was defined as the primary outcome and referred to as “cancer-specific death”. Follow-up started from post-operative day 1 and ended on the date of death or the date of the patient’s final visit to the hospital. In Japan, given that patients receiving curative resection

generally continue regular outpatient visits for at least five years<sup>5</sup> and patients are hospitalized before dying, most deaths were able to be collected through dedicated follow-up by surgeons.

#### *Measurement of other covariates*

Other clinical variables potentially useful in predicting cancer-specific death used in the present study were age, gender, histological type, and pathological T category. Histological type was dichotomized as “well or moderately differentiated” and “poorly differentiated, signet ring cell, or mucinous”. Pathological T category was classified into four categories according to the seventh TNM classification: T1/T2, T3, T4a, and T4b.

#### *Statistical analysis*

All statistical analyses were conducted using R version 2.13.1 ([www.r-project.org](http://www.r-project.org)) and Stata version 11.0 (Stata Corp., College Station, TX, USA). Demographic characteristics, histological type, pathological T category, UICC7-based lymph node staging, presence of ALN metastasis, and number of retrieved lymph nodes were described. Continuous variables and categorical variables were summarized as median (with 10th to 90th percentiles) and percent, respectively.

Cancer-specific death during the follow-up period was described as number, proportion, and incidence rate. Cancer-specific survival curves for those with and without ALN metastasis were estimated using Kaplan-Meier method, and difference between the two was assessed via the log-rank test. These analyses were stratified by each category of UICC7-based lymph node staging.

Effect measures in the present study were hazard ratios (HRs) for cancer-specific death using Cox models. Two models were constructed: a risk model of UICC7 based lymph node staging (Model 1) and a risk model of UICC7-based lymph node staging plus ALN status

(Model 2). To estimate adjusted HRs, all of the covariates described above were forced into multivariate analyses. Facility clustering effects of these analyses were addressed using a robust variance estimator,<sup>7</sup> and the proportional hazards assumption was tested and verified for each covariate.

Discrimination for the two risk models was conducted using Harrell's c index, which is similar to the area-under-the-curve statistic for receiver-operating characteristic plots but allows calculation of concordance in time to event data. The c-indices of the two risk models and their difference were calculated with 2,000-replication bootstrapping, and bias-corrected 95% confidence intervals (95% CIs) were reported. Both the bootstrapping technique used in the present paper and the data-splitting technique estimate the likely performance of the prediction model on a new sample of patients from the present patient population.<sup>8</sup> However, the bootstrapping technique is more efficient than the data-splitting technique in that it preserves the sample size, improving precision and power.<sup>9, 10</sup> Thus, 95% CIs for the c-indices of the two risk models and their difference can be viewed as validated discrimination measures applied to a new sample from the same population.<sup>9, 10</sup>

The predicted risk of cancer-specific death at five years was estimated from a Cox model containing UICC7-based variables alone or UICC7-based variables plus ALN status. The predicted risk was determined by first running a Cox model to obtain the baseline survival function at five years, expressed as  $S_0(5)$ .<sup>11</sup> A risk score for each value of UICC7-based variables, with or without ALN status, was then calculated by multiplying the observed value for the model parameter by its corresponding coefficient from the Cox model. The estimated probability of observing cancer-specific death at five years was then calculated using the formula:  $P(5) = 1 - S_0(5) \times \exp(\text{risk score})$ .

The potential for ALN status to improve risk prediction beyond a UICC7-based risk model (Model 1) for cancer-specific death was further assessed based on category-free net

reclassification improvement (NRI) and integrated discrimination improvement (IDI) as, to date, clinically relevant cut-off values for the risk of cancer-specific death are absent among patients with colon cancer, and both category-free NRI and IDI are independent of arbitrarily defined risk thresholds.<sup>12</sup> Category-free NRI is a measure for estimating overall improvement in reclassification of patients on incorporating ALN status into the UICC7-based risk model and is calculated as the net proportion of patients whose predicted risk improved among those with (case) and without (control) cancer-specific death. For each patient, the predicted risk of cancer-specific death at five years was determined using the UICC7-based risk model, and the relative improvement in patient reclassification associated with ALN metastasis was then assessed. In contrast, IDI is a measure for examining the ability of ALN status added to the UICC7-based risk model to increase average sensitivity without reducing average specificity<sup>12</sup> and is calculated as improvement in the difference in the predicted risk between patients with (case) and without (control) cancer-specific death. Any events outside the five-year time frame were censored. Patients with censored data were treated as controls.

We also conducted sensitivity analyses. (1) Cancer-specific death was evaluated using patients with complete data, considering deaths other than cancer-specific death as competing risk according to the competing-risk regression model developed by Fine and Grey.<sup>13</sup> (2) Cancer-specific death was analyzed by imputing missing covariate and lymph node staging values. Multiple imputation using a chained equations method was used for estimating adjusted HRs,<sup>14, 15</sup> with results showing that 162 participants (10.7%) had missingness. Multiple imputation with five imputed datasets using `aregImpute` function in the `Hmisc` package was used for estimating difference in c-indices of the two risk models, category-free NRI, and IDI.<sup>16</sup> (3) All-cause death was analyzed using patients with complete data.  $P < 0.05$  was considered statistically significant.

## Results

Of the 17,739 patients in the registry, 9,378 with primary colon cancer underwent curative resection (Figure 1). Of those, 2,840 had Stage III disease, with at least 1,593 undergoing regional lymph node dissection involving ALN (277 underwent lymph node dissection of unknown extent). After excluding 76 (for missing data for survival time) and 162 patients (for either or both missing data for number of lymph node metastases [159] or having at least one confounding variable missing [11]), the remaining 1,355 with Stage III primary colon cancer who underwent regional lymph node dissection from 71 facilities were ultimately entered into primary analyses (85.1% of Stage III disease patients with regional lymph node dissection). Baseline characteristics of patients are presented in Table 1. Median age was 65 years. Well or moderately differentiated adenocarcinoma was the dominant histological type (91.0%). The most commonly assigned categories for pathological T and lymph node staging were T3 (55.4%) and N1a (39.4%), respectively. The median number of retrieved lymph nodes was 19.

Baseline characteristics were similar between participants with and without missing confounding variables or lymph node staging data (Table S1) except for lymph node staging and apical lymph node metastasis. As a group, their age, distribution of gender, distribution of histological type, and distribution of pathological T category were similar.

Overall, ALN metastasis was present in 8.3% (N=113) of patients. Respective proportions of apical lymph node metastasis were 2.1% (N=11), 7.1% (N=33), 10.2% (N=23), and 35.7% (N=46) among those with N1a, N1b, N2a, and N2b, respectively. As a group, patients with ALN metastasis had poorer histological type and more advanced pathological T and lymph node staging than those without metastases, whereas median number of retrieved lymph nodes were similar, regardless of ALN metastasis (Table 1).

Over the course of this study, a total of 5,356 patient-years (mean, 4.0 years; median, 5.0 years) and 221 (16.3%) instances of cancer-specific death were observed, giving an incidence rate of cancer-specific death of 4.13 per 100 patient-years. Overall, cancer-specific survival at 5 years was 81.0%. When stratified based on UICC7-based lymph node staging, cancer-specific survival at 5 years was worse among patients with ALN metastasis than among those without it (Figure 2), with respective survival rates of 67.5% and 87.5% for N1a (log-rank test  $P=0.068$ ), 62.5% and 83.3% for N1b ( $P=0.003$ ), 49.8% and 83.0% for N2a ( $P<0.001$ ), and 37.2% and 61.1% for N2b ( $P=0.008$ ).

Associations between two lymph node staging models and cancer-specific death, described in Table 2, were slightly attenuated after adjustment for covariates. For Model 1, N2b was associated with higher cancer-specific death rate than N1a (adjusted HR, 3.86; 95% CI, 2.28-6.53). For Model 2, N2b was still associated with higher cancer-specific death rate than N1a, but slightly more weakly than seen in Model 1 (adjusted HR, 2.98; 95% CI, 1.73-5.12). In the same model, the presence of ALN metastasis was associated with an increased cancer-specific death rate (adjusted HR, 2.29; 95% CI, 1.49-3.52). Sensitivity analyses in which non-cancer-specific deaths were considered competing risks showed that adjusted HRs of lymph node staging and ALN metastasis were similar to those presented in Table 2 (Table S2). We therefore chose the Cox models for primary analysis of cancer-specific death and based further analyses on these models.

Concordance indices and differences in them between these two risk models are presented in Table 3. Harrell's c statistics were 0.694 and 0.708 for the risk model based on UICC7 and the risk model based on UICC7 and ALN status, respectively, and the latter was better than the former (difference, 0.014; 95% CI, 0.005-0.030), suggesting that the risk model based on UICC7 and ALN has better performance of discrimination than one based on UICC7 alone.

Reclassification of risk was assessed for cancer-specific death after addition of ALN status to the risk model based on UICC7 (Table 4). Both of the risks predicted using the UICC7-based model with or without addition of ALN status at the individual level were visualized as plots (Figure S1). Addition of ALN status to the UICC7-based model resulted in 28.5% correct (up) reclassifications and 71.5% incorrect (down) reclassifications in the 221 patients with cancer-specific death. In contrast, 81.2% correct (down) reclassifications and 18.8% incorrect (up) reclassifications were noted in the 1134 patients without cancer-specific death. Overall, 19.4% of patients were correctly reclassified on addition of ALN status (category-free NRI, 19.4% [95% CI, 5.0%-33.4%; P=0.008]). Further, addition of ALN status to the UICC7-based model significantly improved stratification between the patients with cancer-specific death and those without it (IDI, 0.0146 [95% CI, 0.003-0.026; P=0.014]), indicating that the difference in average predicted risk between patients with cancer-specific death and those without increased by 0.0146 on incorporation of ALN status into the UICC7-based risk model.

In sensitivity analyses using imputed data for 1,517 participants, adjusted HR of ALN status for cancer-specific death (Table S3), the difference in the c-indices of the two risk models (Table S4), category-free NRI, and IDI (Table S5) were similar. When all-cause death was analyzed among the 1,355 patients, 290 (21.4%) of all-cause death was observed (overall survival at 5 years: 76.0%). The results were similar to those for cancer-specific death (Figure S2; Table S6; Table S7; and Table S8).

## Discussion

In this multicenter cohort study of stage III colon cancer patients, pathologically-confirmed ALN metastasis was independently associated with cancer-specific death. In addition, incorporating ALN metastasis into the seventh TNM classification-based prognostic model

significantly improved risk reclassification for cancer-specific death. Results were similar when all-cause death was examined as an outcome. Our findings suggest that identification of ALN metastasis by anatomical classification may provide additional prognostic value for risk stratification of cancer-specific death for patients with colon cancer.

Findings from several previous studies support the prognostic value of anatomical classification of the metastasized lymph nodes. Indeed, one study showed that the inferior mesenteric artery (IMA) lymph node metastasis (i.e. an apical lymph node metastasis) was associated with para-aortic nodal recurrence.<sup>2</sup> However, participants in this previous study were restricted to patients with sigmoid colon and rectal cancer, and the association between IMA lymph node metastasis and mortality was not examined. Another study showed that anatomical classification of metastasized lymph node was independently associated with overall survival among patients with sigmoid colon and rectal cancer.<sup>3</sup> The importance of anatomical classification of metastasized lymph nodes in predicting overall survival was also shown among patients with colon or rectal cancer.<sup>4, 17</sup> However, these previous studies involve several shortcomings, such as their single-center status, the fact that most had relatively small sample size, their therapeutic heterogeneity (especially adjuvant chemotherapy), and most importantly, their failure to examine the additive value of incorporating the anatomical classification of the metastasized lymph nodes into the existing TNM classification-based risk model. Our multicenter study involved 1355 patients with any portion of the Stage III colon cancer patients from more than 71 facilities. In addition, we focused on the prognostic importance of ALN metastasis and showed that incorporating the ALN into the seventh TNM classification based risk model significantly improved predictive ability for both cancer-specific and all-cause death.

We feel that our findings here will influence the activities of surgeons and cancer researchers for several reasons. First, a more accurate staging system enables surgeons to

provide detailed information to postoperative patients. Further, identification of those subgroups with a relatively poor prognosis may indicate necessity of more extensive adjuvant therapy to such patients. Second, our findings may represent a basis for conducting a future study examining the effectiveness of dissecting ALNs for colorectal cancer. The ALNs are considered the entry point for lymph node metastases further proximal to the aorta (such as the para-aortic lymph node) or for hematogenous metastasis, and therefore their retrieval may offer prognostic benefit. Indeed, complete mesenteric excision (CME) including ALN has been proposed as a method of improving oncological outcomes.<sup>18</sup> However, CME is often associated with a greater technical requirement, involving high ligation of a major artery, and might therefore increase incidence of postoperative complications. Further study is therefore warranted to clarify the role of CME in apical lymph node dissection. Third, the present study demonstrated that the TNM classification can be restructured based on the evidence to predict patient prognosis better by incorporating ALN metastasis into the risk model based on the present TNM classification.<sup>19</sup> Of note, ALN was an important feature in the lymph node staging in the fourth TNM classification but has since been dropped from the fifth TNM classification due to presumed lack of importance.<sup>20</sup> The seventh TNM classification has come under criticism due to lack of an evidence base for changing the definitions of classifications (e.g. tumor deposits) from previous editions [check – is this what you mean?<-- How is this? Tumor deposits is one example of “definition of classification” and is not main theme in the present study.].<sup>21</sup> Further studies regarding the additive impact of pathological findings on prognosis prediction will be required to revise the TNM classification scientifically.

Several strengths to the present study warrant mention. First, this cohort study was conducted using a nationwide registry of major teaching hospitals across different areas of Japan, and our survival analyses accounted for variability in survival across the facilities

using a robust variance estimator.<sup>7</sup> Our study design and analysis plan ensures the generalizability of our findings. Second, the relatively short, uniform enrollment period prevented our findings from being confounded by advances in perioperative management or development of new anti-cancer agents such as oxaliplatin, which were not marketed in Japan between 2000 and 2002. Third, the association between ALN metastasis and cancer-specific death might be less likely to be affected by competing risks because Japanese patients tend to experience fewer cardiovascular disease risks than patients in Western countries.<sup>22</sup> Indeed, results for sensitivity analyses with all-cause death as the outcome were similar to those in primary analyses. Fourth, by calculating the difference in the c-indices, category-free NRI, and IDI, we were able to demonstrate that incorporating ALN status into the seventh TNM classification-based risk model not only improved the discriminative performance but also improved accuracy in risk reclassification for cancer-specific death and all-cause death.

However, several limitations to the present study warrant mention. We were unable to incorporate other potential confounders such as tumor markers or degree of lymphatic vessel invasion or vascular invasion, which might have resulted in overestimation of the strength in the association between ALN metastasis and cancer-specific death. In addition, we were unable to incorporate the use of adjuvant chemotherapy. However, lack of adjustment for adjuvant chemotherapy tends to result in underestimation of the strength of the association between ALN metastasis and cancer-specific death, as adjuvant chemotherapy is more likely to be provided to those with more advanced lymph node staging involving ALN. Second, we were unable to incorporate N1c category of the seventh TNM classification in the present study, as tumor deposit was not used for staging between 2000 and 2002. Further study is warranted to examine the prognostic value of ALN metastasis. Third, among the patients experiencing cancer-specific death, the risk model based on UICC-7 based variables plus ALN status proved less accurate in its predictability than the model based on UICC-7 alone,

although ALN metastasis itself was an independent prognostic factor. Further research is needed to construct a more precise risk model.

In conclusion, ALN metastasis is a significant prognostic factor in stage III colon cancer, and integration of ALN into the seventh TNM classification-based risk model improves accuracy of prediction of prognosis. ALN should therefore be incorporated into the next TNM classification to better stratify patient prognosis.

## References

1. Sobin LH, Gospodarowicz MK, Wittekind C. *TNM Classification of Malignant Tumours (ed 7)*. 7 ed. West Sussex, United Kingdom: Wiley-Blackwell; 2009.
2. Kang J, Hur H, Min B, Kim N, Lee K. Prognostic Impact of Inferior Mesenteric Artery Lymph Node Metastasis in Colorectal Cancer. *Ann Surg Oncol*. 2011/03/01 2011;18(3):704-710.
3. Huh JW, Kim YJ, Kim HR. Distribution of Lymph Node Metastases Is an Independent Predictor of Survival for Sigmoid Colon and Rectal Cancer. *Annals of Surgery*. 2012;255(1):70-78 10.1097/SLA.1090b1013e31823785f31823786.
4. Kobayashi H, Ueno H, Hashiguchi Y, Mochizuki H. Distribution of lymph node metastasis is a prognostic index in patients with stage III colon cancer. *Surgery*. 2006;139(4):516-522.
5. Kobayashi H, Mochizuki H, Sugihara K, et al. Characteristics of recurrence and surveillance tools after curative resection for colorectal cancer: A multicenter study. *Surgery*. 2007;141(1):67-75.
6. . *Japanese Society for Cancer of the Colon and Rectum. Japanese classification of colorectal carcinoma. Second English Edition*. . Tokyo: Kanehara & Co., Ltd; 2009.
7. Klein J, Moeschberger M. *Survival Analysis Techniques for Censored and Truncated Data*. New York: Springer; 1997:416 - 418.
8. McGinn TG, Guyatt GH, Wyer PC, et al. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. *JAMA*. 2000;284(1):79-84.
9. 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(4):361-387.

10. Harrell Jr FE. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer; 2001:1-568.
11. Gulati A JAIF, et al. Association of fibrosis with mortality and sudden cardiac death in patients with nonischemic dilated cardiomyopathy. *JAMA*. 2013;309(9):896-908.
12. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat. Med.* 2011;30(1):11-21.
13. Fine JP, Gray RJ. A Proportional Hazards Model for the Subdistribution of a Competing Risk. *Journal of the American Statistical Association*. 1999/06/01 1999;94(446):496-509.
14. Rubin D. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons; 1987.
15. van Buuren S, Boshuizen HC, Knook DL. Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in medicine*. 1999;18(6):681-694.
16. Harrel FEJ. Hmisc: Harrell Miscellaneous. R package version 3.12-2. 2013; <http://cran.r-project.org/web/packages/Hmisc/index.html>. Accessed Accessed Jul 24, 2013.
17. Kobayashi H, Enomoto M, Higuchi T, et al. Clinical Significance of Lymph Node Ratio and Location of Nodal Involvement in Patients with Right Colon Cancer. *Digestive Surgery*. 2011;28(3):190-197.
18. Hohenberger W, Weber K, Matzel K, Papadopoulos T, Merkel S. Standardized surgery for colonic cancer: complete mesocolic excision and central ligation – technical notes and outcome. *Colorectal Disease*. 2009;11(4):354-364.
19. Quirke P, Cuvelier C, Ensari A, et al. Evidence-based medicine: the time has come to set standards for staging. *The Journal of Pathology*. 2010;221(4):357-360.

20. Sobin LH, Wittekind C. *TNM Classification of Malignant Tumours (ed 5)*. 7 ed. New York: Wiley-Liss; 1997.
21. Nagtegaal ID, Quirke P. Revised Staging: Is It Really Better, or Do We Not Know? *Journal of Clinical Oncology*. August 10, 2010 2010;28(23):e397-e398.
22. Henderson SO, Haiman CA, Wilkens LR, Kolonel LN, Wan P, Pike MC. Established Risk Factors Account for Most of the Racial Differences in Cardiovascular Disease Mortality. *PLoS ONE*. 2007;2(4):e377.

## **Figure legend**

### **Figure 1. Flow chart of the study population.**

JSCCR: Japanese Society for Cancer of the Colon and Rectum.

### **Figure 2. Description of cancer-specific survival by UICC7-based lymph node staging and apical lymph node status.**

Cancer-specific survival curves were estimated via the Kaplan-Meier method. Within each UICC7-based lymph node staging (i.e. N1a to N2b), survival curves are shown based on apical lymph node status. Solid and dashed lines indicate survival curves for those with and without apical lymph node metastasis, respectively. ALN: Apical lymph node metastasis, (-): Absent, (+): Present.

**Table 1. Baseline characteristics of the analysis population**

	Total (N = 1,355)	ALN metastasis absent (N=1,242)	ALN metastasis present (N=113)
Age, y			
<i>Median</i>	65	65	67
<i>10th to 90th percentile</i>	50 to 77	50 to 77	51 to 77
Gender, n (%)			
<i>Men</i>	704 (52.0)	655 (52.7)	49 (43.4)
<i>Women</i>	651 (48.0)	567 (47.3)	64 (56.6)
Histological type, n (%)			
<i>Well or moderately differentiated</i>	1,233 (91.0)	1,145 (92.2)	88 (77.9)
<i>Poorly differentiated/Signet ring cell/Mucinous</i>	122 (9.0)	97 (7.8)	25 (22.1)
Pathological T category, n (%)			
<i>T1/2</i>	100 (7.4)	96 (7.7)	4 (3.5)
<i>T3</i>	750 (55.4)	700 (56.4)	50 (44.3)
<i>T4a</i>	424 (31.3)	377 (30.4)	47 (41.6)
<i>T4b</i>	81 (6.0)	69 (5.6)	12 (10.6)
Lymph node staging according to UICC7, n (%)			
<i>N1a</i>	534 (39.4)	523 (42.1)	11 (9.7)
<i>N1b</i>	466 (34.4)	433 (34.9)	33 (29.2)
<i>N2a</i>	226 (16.7)	203 (16.3)	23 (20.4)
<i>N2b</i>	129 (9.5)	83 (6.7)	46 (40.7)
Apical lymph node metastasis (ALN), n (%)			
<i>Absent</i>	1,242 (91.7)		
<i>Present</i>	113 (8.3)		
Number of retrieved lymph nodes, n			
<i>Median</i>	19	19	18
<i>10th to 90th percentile</i>	8 to 40	8 to 40	7 to 39

UICC7: Union for International Cancer Control version 7

**Table 2. Association between lymph node staging and cancer-specific death (N=1,355)**

	Number of Participants	Cancer-specific death			
		Unadjusted HR (95% CI)	p-value	Adjusted <sup>a</sup> HR (95% CI)	p-value
<b><u>Model 1: UICC7</u></b>					
<i>Lymph node staging</i>					
<i>N1a</i>	534	Reference		Reference	
<i>N1b</i>	466	1.41 (0.98 - 2.01)	0.061	1.38 (0.94 - 2.02)	0.102
<i>N2a</i>	226	1.63 (1.06 - 2.52)	0.026	1.49 (0.97 - 2.29)	0.069
<i>N2b</i>	129	4.74 (2.91 - 7.71)	<0.001	3.86 (2.28 - 6.53)	<0.001
<b><u>Model 2: UICC7 + ALN status</u></b>					
<i>Lymph node staging</i>					
<i>N1a</i>	534	Reference		Reference	
<i>N1b</i>	466	1.32 (0.93 - 1.89)	0.123	1.31 (0.89 - 1.93)	0.166
<i>N2a</i>	226	1.49 (0.97 - 2.28)	0.066	1.39 (0.91 - 2.12)	0.128
<i>N2b</i>	129	3.47 (2.13 - 5.66)	<0.001	2.98 (1.73 - 5.12)	<0.001
<i>Apical lymph node metastasis</i>					
Absent	1,242	Reference		Reference	
Present	113	2.58 (1.76 - 3.78)	<0.001	2.29 (1.49 - 3.52)	<0.001

<sup>a</sup>Estimated from Cox models including age, gender, histological type, and pathological T stage accounting for facility clustering effects

HR: hazard ratio, CI: confidence interval, UICC7: Union for International Cancer Control version 7, ALN: apical lymph node

**Table 3. Concordance indices and their difference between the two prognostic models for cancer-specific death (N=1,355)<sup>a</sup>**

	<b>Cancer-specific death</b>
	<i>Harrell's C (95% CI)<sup>b</sup></i>
<b>Model 1: UICC7</b>	0.694 (0.651 - 0.724)
<b>Model 2: UICC7 + ALN status</b>	0.708 (0.666 - 0.737)
	<i>Difference in Harrell's C (95% CI)<sup>b</sup></i>
<b>Model 2 vs. Model 1</b>	0.014 (0.005 - 0.030)

<sup>a</sup>Derived from Cox models including age, gender, histological type, pathological T stage, and lymph node staging presented in Table 2, accounting for facility clustering effects.

<sup>b</sup>Confidence intervals were calculated from 2,000 replication bootstrapping with bias correction method. In one re-sampling with a new set of 1,355 patients (some were able to be re-sampled twice or more), we computed point estimates of c-indices (Harrell's c) from model 1 (C1) and from model 2 (C2) and determined difference in the two c-indices. The point estimate of difference in the c-indices is "C2 – C1" and indeed this was positive (over zero) in our analyses. By replicating this process 2,000 times, we obtained 2,000 estimates of C1, C2 and "C2 – C1". From the 2,000 estimates, we obtained the 95% CIs presented in Table 3. Therefore, even though 95% CIs of c-indices for model 1 and model 2 did overlap to some extent, 95% CIs of difference in the two indices were over zero.

CI: confidence interval, UICC7: Union for International Cancer Control version 7, ALN: apical lymph node

**Table 4. Category-free net reclassification improvement and integrated discrimination index for apical lymph node status added to the UICC7-based risk model for cancer-specific death at 5 years (N=1,355)**

<b><i>NRI</i></b>	
<i>Cases reclassified to higher risk</i>	28.5%
<i>Cases reclassified to lower risk</i>	71.5%
<i>Controls reclassified to lower risk</i>	81.2%
<i>Controls reclassified to higher risk</i>	18.8%
<i>Category-free net reclassification improvement</i>	19.4% (95% CI: 5.0%-33.4%); p=0.008
<b><i>IDI</i></b>	0.0146 (95% CI: 0.0030-0.0262); p=0.014

Cases denote patients who experienced cancer-specific death at 5 years whereas controls denote those who did not experience cancer-specific death at 5 years or were censored before 5 years. NRI: net reclassification improvement, IDI: integrated discrimination index, CI: confidence interval

Figure 1

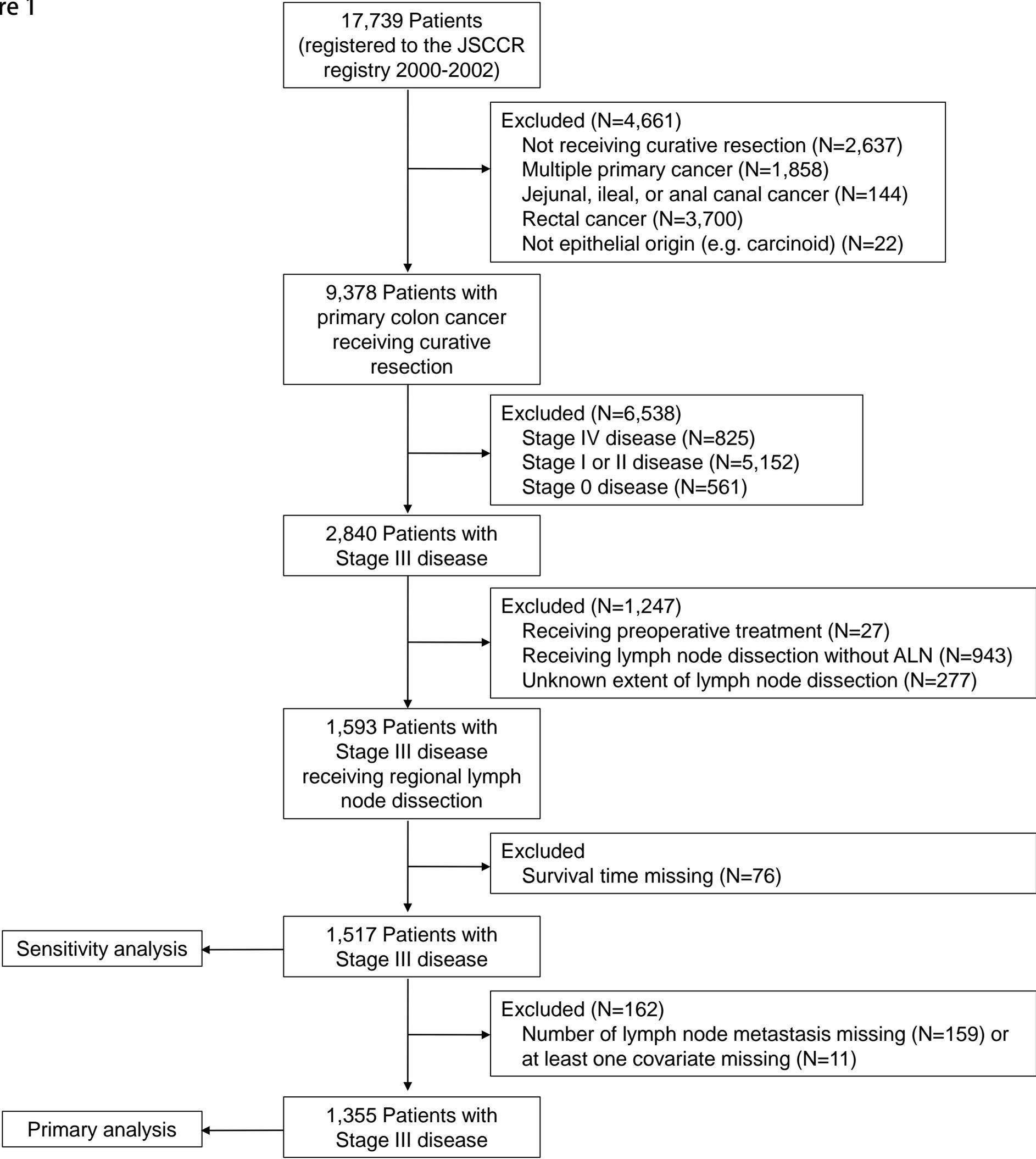
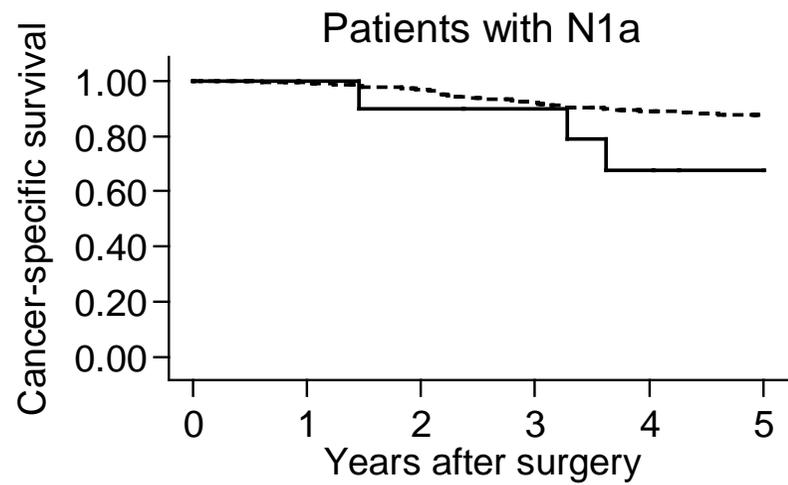
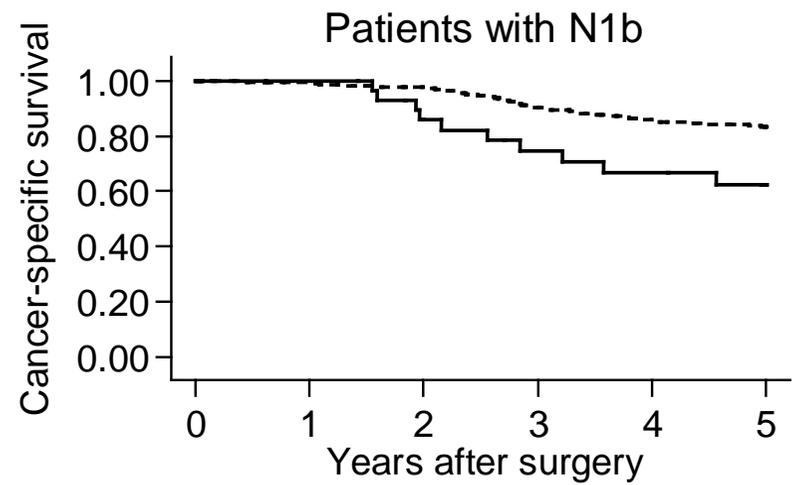


Figure 2



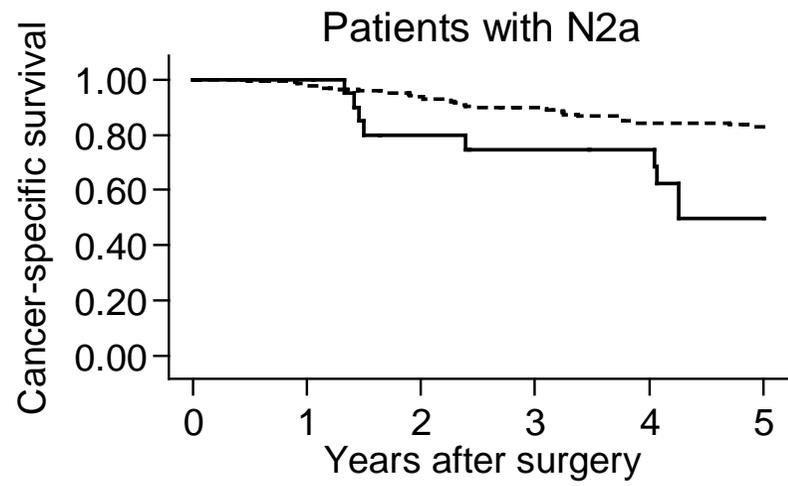
Number at risk

ALN = (-)	523	489	454	404	362	315
ALN = (+)	11	10	9	8	6	4



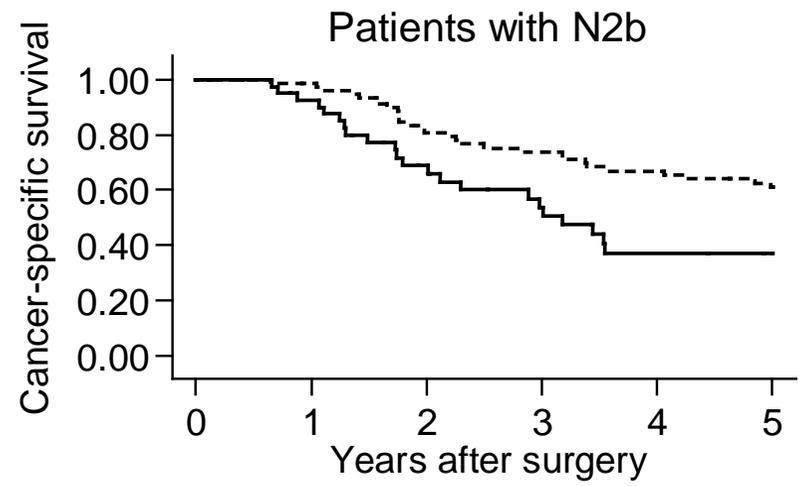
Number at risk

ALN = (-)	433	406	383	332	299	270
ALN = (+)	33	30	23	19	17	13



Number at risk

ALN = (-)	203	188	172	153	136	119
ALN = (+)	23	21	15	13	12	8



Number at risk

ALN = (-)	83	78	61	54	47	39
ALN = (+)	46	37	24	17	11	9