



Published in final edited form as:

Cancer. 2011 September 15; 117(18): 4155–4165. doi:10.1002/cncr.26003.

Prognostic Web-based Models for Stage II and III Colon Cancer: a Population and Clinical Trials-based Validation of Numeracy and Adjuvant! Online

Sharlene Gill¹, Charles Loprinzi², Hagen Kennecke¹, Axel Grothey², Garth Nelson², Ryan Woods¹, Caroline Speers¹, Steven R. Alberts², Aditya Bardia³, Michael J. O'Connell⁴, and Daniel J. Sargent²

¹ British Columbia Cancer Agency, Vancouver, BC

² Mayo Clinic, Rochester, MN

³ Sidney Kimmel Comprehensive Cancer Center, Baltimore, MD

⁴ NSABP Operations Center, Pittsburgh, PA

Abstract

Purpose—Numeracy (www.mayoclinic.com/calcs) and Adjuvant! (www.adjuvantonline.com) are two web-based calculators widely used to estimate the prognosis and potential benefit of adjuvant 5FU-based therapy for patients with Stage II and III colon cancer. This study compares the predicted survival estimates from these models with the actual observed estimates in independent datasets derived from a population cohort and from clinical trials

Methods—The population cohort was derived from the British Columbia Colorectal Cancer Outcomes Unit database which identified referred patients with stage II and III colon cancer from 1995–1996 and 1999–2003. Patients enrolled in NCCTG trials 94651 and 914653 were included in the trials dataset. Patient and disease data were used to determine the predictions for 5-year relapse free and overall survival for both tools.

Results—In the population-based dataset (N= 2033), Adjuvant! offered more reliable predictions of prognosis for patients treated with surgery alone, but similar reliability as Numeracy for patients treated with adjuvant 5FU. Both models tended to overestimate survival in 5FU treated patients with stage II disease. In the trials dataset of patients treated with surgery and 5FU (N= 1729), Numeracy and Adjuvant! demonstrated similar performance and improve correctness.

Conclusions—This independent validation analysis demonstrates that both Numeracy and Adjuvant! have similar predictive performance and acceptable reliability for patients with stage III disease. Survival outcomes of patients with stage II colon cancer treated with adjuvant 5-FU were slightly lower than estimated by either model.

Introduction

In 2009, an estimated 130,000 new cases of colon cancer were diagnosed in the US and Canada, representing the third leading cause of cancer-related incidence and mortality in men and women.^{1,2} The use of adjuvant 5-fluorouracil (5FU) based chemotherapy for stage III and high-risk stage II colon cancer represents the most notable advance for improving

Correspondence: Sharlene Gill, MD, MPH, BC Cancer Agency, 600 West 10th Ave, Vancouver, BC CANADA V5Z4E6, Tel: 604-877-6000, Fax: 604-877-0585, sgill@bccancer.bc.ca.

Previously presented in part at the 2009 American Society of Clinical Oncology Annual Session

survival in the modern day chemotherapeutic management of colon cancer. The benefit of adjuvant therapy has been established through several well-conducted multinational randomized controlled trials of 5FU-based chemotherapy, demonstrating a 35% reduction in risk of death for node-positive disease³. More recently a benefit of the addition of oxaliplatin to 5FU-based chemotherapy has been demonstrated, establishing 5FU and oxaliplatin (FOLFOX or FLOX) as the current standard of adjuvant care for stage III colon cancer.^{4,5}

While the disease-free and overall survival (OS) benefits are well enumerated in clinical trials, predicting the prognosis for individual patients and expected absolute benefit of adjuvant therapy for an individual patient can be complex. Multiple factors are relevant including disease-specific factors, efficacy of the treatment intervention, and competing morbidity and mortality risks. To aid in individualizing decisions regarding adjuvant therapy, prognostic decision calculators have been developed with the intent of providing clinicians with a tailored estimate of a patient's baseline prognosis and predicted treatment benefit for stage II and III colon cancer. Currently, two such web-based programs are available for on-line use: 1) Numeracy, available at www.mayoclinic.com/calcs, a model developed at the Mayo Clinic derived from a pooled individual patient-data meta-analysis of seven randomized clinical trials including 3341 trial subjects³, and 2) Adjuvant!, available at www.adjuvantonline.com, a model developed using prognosis estimates derived from US SEER tumor registry-reported outcomes for colon cancer patients in the general population, and estimates of adjuvant therapy efficacy derived from proportional risk reductions published in the literature.

While both tools are widely used by health professionals to assist in providing estimates of benefit of adjuvant 5FU-based therapy, the validity of the assumptions inherent in the projections provided by these on-line prognostic tools have not yet been independently validated. This study examines the accuracy and validity of Numeracy and Adjuvant! by comparing the predicted recurrence free survival (RFS) and OS estimates with the observed outcomes from two independent data sources: a population-based dataset of patients referred to a Canadian provincial cancer agency and a clinical trial dataset including patients enrolled in two adjuvant therapy trials, NCCTG 894651 and NCCTG 914653.^{6,7}

Methods

Population-Based Dataset Validation

Patients with resected stage II and III colon cancer referred to the British Columbia Cancer Agency (BCCA) between 1995 to 1996, and 1999 to 2003, were identified from the Colorectal Cancer Outcomes Unit (CRCOU) database. The BCCA provides a population-based cancer control program for residents of the province of British Columbia, Canada and consists of regional cancer centers in partnership with an extensive community oncology network (CON). For these time periods, over 60% of BC patients with resected colon cancer were referred to the BCCA, with non-referred patients representing patients either deemed to be unsuitable for an adjuvant therapy referral on the basis of age, comorbidities or indication, or patients who were primarily managed in the CON. In accordance with the BCCA systemic therapy guidelines during the study period, adjuvant chemotherapy was recommended for patients with stage III disease but was not routinely recommended for stage II disease. The endorsed chemotherapy regimen at that time was the Mayo Clinic monthly regimen of bolus 5FU and low-dose leucovorin.

Data collected from the CRCOU database included patient demographics (age, gender, year of diagnosis), disease status (histologic grade of well, moderate or poor; tumor depth of invasion; number of positive nodes and number of examined nodes) and treatment received (surgery alone or surgery plus 5FU chemotherapy). Observed patient outcomes were

recorded as measured from date of diagnosis to observed first relapse and, date and cause of death where applicable.

Clinical Trial Dataset Validation

NCCTG 894651 was a trial of 915 patients with stage II or III colon cancer with double random assignment to receive standard 5FU and levamisole versus 5FU plus leucovorin and levamisole, and to receive 12 months versus 6 months of chemotherapy.⁶ No difference in survival between 6 months and 12 months of chemotherapy or the choice of chemotherapy was demonstrated. NCCTG 914653 was a trial of 878 patients with stage II or III colon cancer randomly assigned to receive standard dose levamisole combined with 5FU plus leucovorin versus high dose levamisole combined with the same chemotherapy.⁷ No difference in survival was demonstrated with high dose levamisole. The individual patient data from both trials was pooled for this analysis. Of note, these trials were not included in the pooled analysis for the development of the Numeracy tool³ and represent an independent trials dataset. Data collected from the trials database included patient demographics (age, gender) and disease status (histologic grade of low, classified as well or moderate, and high, classified as poor or anaplastic; tumor depth of invasion; and number of positive nodes). As the trials database did not include a variable for number of examined nodes, this variable was derived using an algorithm assignment of 4–10 examined nodes if the number of positive nodes was less than or equal to 3 and, greater than 10 examined nodes if the number of positive nodes was greater than 3. Observed patient outcomes were recorded as measured from date on study to date of first recurrence, and date of death where applicable.

Analysis

The observed RFS was measured from the date of diagnosis for the population-based dataset and date on study for the clinical trial-based dataset, to the date of recurrence or death if no recurrence. The observed OS was measured from the date of diagnosis or date on-study to date of death. Predicted RFS and OS estimates were derived for each patient in both validation datasets using the Numeracy (v2003) and Adjuvant!Online (v2005) calculators. The input options for Numeracy include lymph nodes (none, 1–4, 5+), tumor stage (T1/T2, T3, T4), grade (low, high) and age (49 years or younger, 50–59, 60–69, 70 years or older). The Adjuvant! tool has three additional inputs for gender (male, female), number of examined lymph nodes (0, 1–3, 4–10, >10) and comorbidity (perfect health, minor problems, average for age, or major problems). As comorbidity measures were not available for either dataset, a default comorbidity assumption of ‘Minor Problems’ was applied. A sensitivity analysis with “average for age” comorbidity was also performed. In addition, the nodal and grade categories for the two tools differed slightly. Hence in this analysis, the ‘1–3’ and 4+ positive node categories in Adjuvant! were considered interchangeable with the ‘1–4’ and 5+ positive node categories respectively in Numeracy. Likewise, the grade 1 and 2 categories in Adjuvant! were considered interchangeable with the low grade category in Numeracy.

Observed 5 yr RFS and OS by the method of Kaplan-Meier (KM) were compared to predicted estimates from Numeracy and Adjuvant! by univariate analysis by relevant prognostic factors and by combined groupings of T stage, N stage, and grade. Groups with fewer than 10 observations were removed. The comparative analyses are presented in a descriptive manner using the absolute difference in percentiles between observed and predicted 5-year outcomes by univariate analysis, the percent closer rate across combined prognostic subgroups by multivariate analysis, and the percent correct predictions of 5 year status for each model. Correctness was deemed accurate if a patient was alive and predicted to be alive with probability greater than or equal to 50%, or dead with a predicted (probability of being alive less than 50%. For RFS and OS, the Adjuvant! and Numeracy

predictions were divided into 5% intervals. Intervals were grouped so that each interval contained at least 50 observations. The observed KM estimations for each interval subset were plotted against the average prediction for both tools.

Results

Population-Based Dataset

A total of 2033 patients are included in the BCCA cohort with a median follow-up period of 5.6 years. 924 patients (45%) received surgery alone and 1109 patients (55%) were treated with surgery plus 5FU. Table 1 summarizes the demographic and staging characteristics of this overall cohort and by treatment status, with 5FU treated patients noted to be younger and more commonly with node-positive disease.

Table 2 presents the overall and subgroup univariate mean predictions for Numeracy and Adjuvant!, and the observed KM rates with 95% confidence intervals (CI) for 5-year RFS and OS. For surgery alone overall, the 5-year OS and RFS estimates from Numeracy and Adjuvant! differed from the mean observed rate by 6% and 3%, and 3% and 2% respectively. For the node-negative, stage II subgroup (n=621), the Numeracy and Adjuvant! estimates for OS differed from observed by 4% and 2%, and for RFS differed by 2% and 2% respectively. Across all 14 subgroups by age, gender, grade, T-stage and nodal status, the Adjuvant! estimates were more commonly within the observed 95% CI for OS (14/11 versus 9/14) and RFS (13/14 versus 10/14) as compared to Numeracy. The differences between the predicted and observed estimates were within 5% more frequently for Adjuvant! (12/14 for OS, 13/14 for RFS) versus Numeracy (8/14 for OS, 9/14RFS). In untreated patients, both calculators systematically under-estimated risk of recurrence/death compared to observed rates for young patients (age <50) and over-estimated risk of recurrence/death in older patients and patients with 4–10 lymph nodes.

In table 2, for surgery plus 5FU, the difference between estimates in the overall group by Numeracy and Adjuvant! from the observed 5-year OS rates were 2% in both calculators, and 4% and 6% for 5-year RFS. In the pooled treated group, the predicted estimates from both calculators were beyond the upper limit of the 95% CI for the observed RFS estimate. For stage II treated patients (n=193), the predictions for OS and RFS were more notably overestimated versus observed, differing by 7% for OS and 6% for RFS with Numeracy versus 8% for OS and 11% for RFS for Adjuvant!. Among the 15 subgroups by age, gender, grade, T-stage and nodal status, the Adjuvant! estimates were within the observed 95% CI for OS with the same frequency as Numeracy (12/15) while, for RFS, Numeracy estimates were more commonly within the observed 95% CI as compared to Adjuvant! (11/15 versus 8/15). The frequency of a difference between the predicted and observed estimate that was within 5% was comparable (Adjuvant!: 12/15 for OS and 8/15 RFS; Numeracy: 11/15 for OS and 8/15 for RFS).

The percent of predictions closer to the observed data across prognostic groups by multivariate analysis are presented in table 3 being notably higher for Numeracy for RFS in the surgery plus 5FU cohort versus higher rates for Adjuvant for OS in the surgery alone cohort. Table 3 also illustrates the percent correct rates for RFS and OS which are similar for both prognostic models within both cohorts. The 5 year Numeracy and Adjuvant! average predicted estimates for OS and RFS divided into 5% intervals are shown in Figures 1A–1D against the observed KM estimates for surgery alone and surgery plus 5FU.

Clinical Trials Based Dataset

A total of 1729 patients eligible for this analysis and enrolled in 894651 (n=887) and 914653 (n=842) were included with a median follow-up of 8.2 years. Table 1 presents the

characteristics for this pooled cohort, showing a similar distribution of characteristics for both studies. All patients received surgery plus 5FU chemotherapy and 21% were node-negative. In Table 4, the overall OS and RFS predictions differed from the observed rates respectively by -2% and -1% for Numeracy, versus 0% and 4% for Adjuvant!. The pooled RFS estimate for Adjuvant! was beyond the 95% CI. For patients with node-negative disease (n=363), the predicted estimates differed from observed OS and RFS by -5% and -5%, versus -4% and -1% for Numeracy versus Adjuvant!. Among the 15 subgroups defined by age, gender, grade and T-stage, nodes, the Numeracy estimates were less commonly within the observed 95% CI for OS than Adjuvant! (11/15 versus 13/15), but were more commonly within the 95% CI for RFS (15/15 versus 10/15). Both calculators overestimated prognosis in patients with 10 or more positive nodes, although the sample size in this group was small (n=18). The absolute differences between mean predicted and observed estimates were comparably within 5% for Numeracy (11/15 for OS and 12/15 for RFS) versus Adjuvant! (11/15 for OS and 11/15 for RFS).

As shown in Table 5, the percent closer rate by multivariate analysis across prognostic groups was higher with Adjuvant! for RFS and OS. The percent correct predictions are very similar for both prognostic models. Figures 1E and 1F illustrate the 5 year Numeracy and Adjuvant! average predicted estimates for OS divided into 5% intervals compared to the observed KM estimates.

Discussion

The present study evaluated the performance of two commonly used web-based prognostic models for the adjuvant treatment of stage II and III colon cancer: Mayo Clinic's 'Numeracy' and Adjuvant!, using descriptive analyses of the predicted estimates of prognosis and adjuvant 5FU treatment benefit against observed estimates from two independent datasets: a population-based experience of 2033 patients referred to the BCCA and a pooled set of 1729 patients enrolled in two NCCTG adjuvant 5FU trials. A recent comparative analysis based upon hypothetical patient scenarios using all possible combinations of comparable data yielded similar outputs for Numeracy and Adjuvant! in predicting the absolute benefit of 5FU chemotherapy over surgery alone although some differences were noted, particularly in estimates of baseline prognosis for untreated patients.⁸ While the Adjuvant! tool has previously been validated for use in breast cancer⁹, neither tool has been validated for use in the adjuvant colon cancer setting. Using a number of descriptive comparative metrics, this current analysis has demonstrated that while both Numeracy and Adjuvant! exhibit overall acceptable and similar reliability, some important areas of disagreement are present.

For baseline prognosis estimates for surgery alone using the population cohort, Adjuvant! offered more reliable predictions than Numeracy. This is not entirely unexpected as Adjuvant! derives its estimates from the US SEER registry and US mortality data which is a population-based representation. By the comparison measures utilized in this analysis, the estimates for patients treated with 5FU were similar in the population and trials datasets by the measures of comparison utilized in this analysis. With respect to the robustness of these tools, both models had acceptable performance for the population and trials validation of treated patients despite the comparatively younger age of the trials cohort.

More of a difference was noted for patients with stage II disease. Compared to the population-based dataset, both Numeracy and Adjuvant! frequently overestimated the RFS and OS. This may reflect the reality that treated stage II patients in the population cohort were likely to have other features suggestive of high-risk disease which would have compelled treatment and were consequently associated with inferior 5-year observed

outcomes. Evaluating the models in terms of the difference between observed and predicted estimates, both Numeracy and Adjuvant! demonstrated better performance in the trials dataset with closer albeit slight under-estimations in predicting RFS and OS in treated patients. This may be partly explained as both models employ estimates of 5FU treatment efficacy from trials-derived proportional reductions of recurrence risk and death. Nonetheless, as these tools are clinically applied in the real-world, the tendency for modestly over-optimistic estimates of treatment benefit for stage II disease should be considered.

At the individual patient level, both tools do leave room for improvement. Each tool in each case accurately predicted outcome at the patient level (judged as correct if the patients predicted probability of 5-year survival was > 50% and the patient was alive, and similar for death probability) just two-thirds of the time. Additional markers of prognosis and treatment benefit in the adjuvant setting are required. While not yet routine, molecular prognosticators of interest such as microsatellite instability and, more recently, multigene expression profiles for recurrence risk assessment, hold the potential to influence adjuvant decision-making.^{11, 12} It would be of future interest to evaluate the real-world impact of these biomarkers as compared to, or in conjunction with, the Numeracy and Adjuvant! clinical prediction tools.

A major strength of this study is that it employs the use of individual patient data from two independent sources which were not used to derive the prognostic models for Numeracy or Adjuvant!. The primary limitation of this analysis is that it validates the predictive performance of these models for 5FU therapy while the current standard for patients with stage III disease is 5FU plus oxaliplatin (FOLFOX). Both the Numeracy and Adjuvant! programs presently offer survival estimates for FOLFOX treatment based by application of the proportional benefits observed in MOSAIC and NSABP C-06.^{4,5} While certain observations with respect to predicting baseline prognosis and risk of over-estimating treatment benefit may be inferred, a subsequent validation for FOLFOX is also desirable.

So, if the treatment benefit estimates are comparable, is one prognostic model preferred? Both tools are web-accessible and available free-of-charge. Numeracy at www.mayoclinic.com/calcs does not require registration or login, and only involves four data inputs of age, T stage, N stage and grade. Adjuvant! at www.adjuvantonline.com requires registration and login, and involves additional data points for input including comorbidity, gender and number of examined lymph nodes. Its interface is more user-friendly than Numeracy with presentation graphics and illustrations which may be useful when discussing adjuvant therapy with patients¹⁰. The choice of which decision aid to use will therefore be driven by individual preferences and user familiarity.

The decision to offer adjuvant chemotherapy needs to be tailored to include an assessment of treatment-related risks and benefits for the individual patient. Prognostic estimates provided by clinicians tend to be divergent and inconsistent.¹³ The availability of easily accessible and more reliable predictions of outcomes, with or without chemotherapy, provides patients with better prognostic and predictive information, and facilitates the decision-making for the physician, patient and family. While neither tool is 100% accurate, this independent validation analysis demonstrates that both Numeracy and Adjuvant! provide similar predictive reliability which, in concert with clinical judgment and patient discussion, are worthwhile decision aids for 5FU-based adjuvant chemotherapy in stage II and III colon cancer.

Acknowledgments

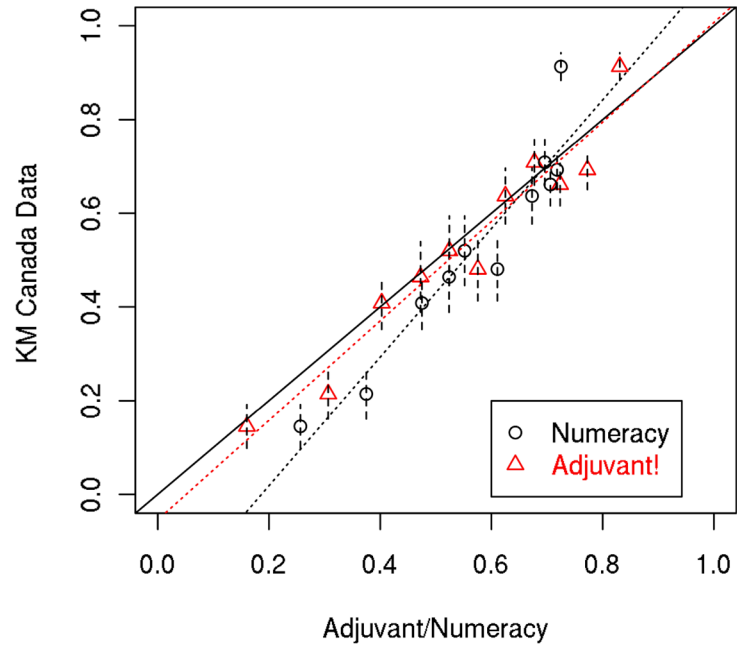
This work was supported, in part, by the following United States National Institutes of Health Grants-CA 25224 and CA 124477

References

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin*. 2009; 59:225–249. [PubMed: 19474385]
2. Committee CCSS. Canadian Cancer Statistics 2009. 2009.
3. Gill S, Loprinzi CL, Sargent DJ, et al. Pooled analysis of fluorouracil-based adjuvant therapy for stage II and III colon cancer: who benefits and by how much? *J Clin Oncol*. 2004; 22:1797–1806. [PubMed: 15067028]
4. Andre T, Boni C, Navarro M, et al. Improved overall survival with oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment in stage II or III colon cancer in the MOSAIC trial. *J Clin Oncol*. 2009; 27:3109–3116. [PubMed: 19451431]
5. Kuebler JP, Wieand HS, O'Connell MJ, et al. Oxaliplatin combined with weekly bolus fluorouracil and leucovorin as surgical adjuvant chemotherapy for stage II and III colon cancer: results from NSABP C-07. *J Clin Oncol*. 2007; 25:2198–2204. [PubMed: 17470851]
6. O'Connell MJ, Laurie JA, Kahn M, et al. Prospectively randomized trial of postoperative adjuvant chemotherapy in patients with high-risk colon cancer. *J Clin Oncol*. 1998; 16:295–300. [PubMed: 9440756]
7. O'Connell MJ, Sargent DJ, Windschitl HE, et al. Randomized clinical trial of high-dose levamisole combined with 5-fluorouracil and leucovorin as surgical adjuvant therapy for high-risk colon cancer. *Clin Colorectal Cancer*. 2006; 6:133–139. [PubMed: 16945169]
8. Bardia A, Loprinzi C, Grothey A, et al. Adjuvant chemotherapy for resected stage II and III colon cancer: comparison of two widely used prognostic calculators. *Semin Oncol*. 2010; 37:39–46. [PubMed: 20172363]
9. Olivotto IA, Bajdik CD, Ravdin PM, et al. Population-based validation of the prognostic model ADJUVANT! for early breast cancer. *J Clin Oncol*. 2005; 23:2716–2725. [PubMed: 15837986]
10. Hochster HS. Web Alert: Adjuvant therapy for colon cancer. *Curr Colorectal Cancer Reports*. 2007:167–168.
11. Locker GY, Hamilton S, Harris J, et al. ASCO 2006 update of recommendations for the use of tumor markers in gastrointestinal cancer. *J Clin Oncol*. 2006; 24:5313–5327. [PubMed: 17060676]
12. Kerr D, Gray R, Quirke P, et al. A quantitative multigene RT-PCR assay for prediction of recurrence in stage II colon cancer: Selection of the genes in four large studies and results of the independent, prospectively designed QUASAR validation study. *J Clin Oncol*. 2009; 27 suppl abstr 4000.
13. Loprinzi CL, Ravdin PM, de Laurentiis M, Novotny P. Do American oncologists know how to use prognostic variables for patients with newly diagnosed primary breast cancer? *J Clin Oncol*. 1994; 12:1422–1426. [PubMed: 8021733]

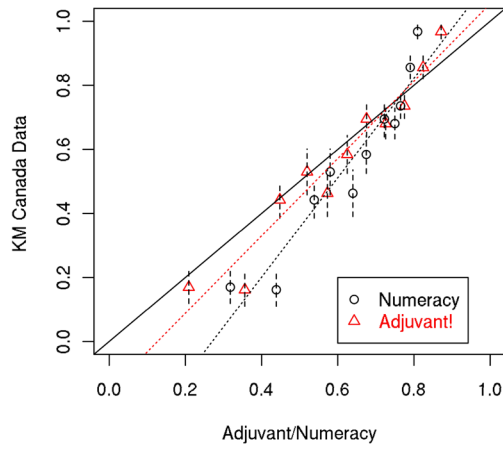
1A

Recurrence Free Survival 5% Prediction Plot (Surgery Only)



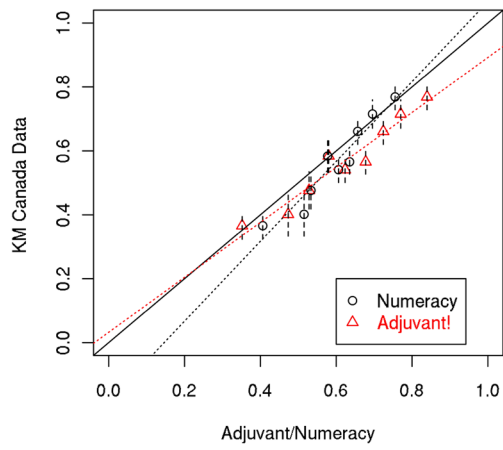
1B

**Overall Survival 5% Prediction Plot
(Surgery Only)**



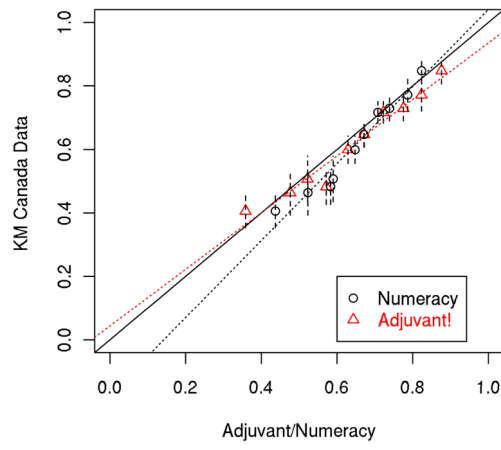
1C

Recurrence Free Survival 5% Prediction Plot



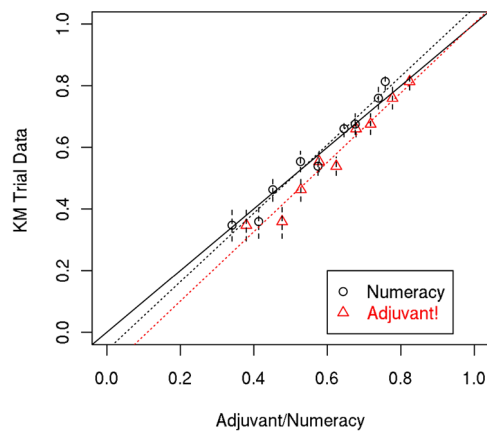
1D

Overall Survival 5% Prediction Plot

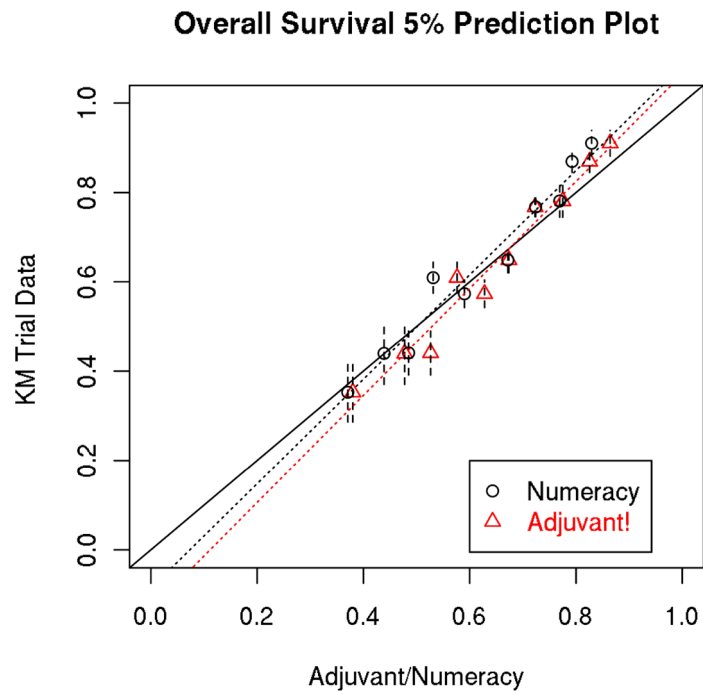


1E

Recurrence Free Survival 5% Prediction Plot



1F

**Figure 1.**

A, B, C and D: Population-Based Data: 5% Prediction Interval Plot for Surgery Alone 5-year RFS (A) and OS (B), and for Surgery plus 5FU 5-year RFS (C) and OS (D). E and F: Trials-Based Data: 5% Prediction Interval Plot for Surgery plus 5FU 5-year RFS (E) and OS (F)

Table 1

Demographics of Cohorts

Population cohort			
	Surgery Alone N=924	Surgery + 5FU N=1109	Total N=2033
Median Age	73y	65y	68y
Male Gender	51.3%	55%	53%
T1/2 stage	2.5%	8.9%	6%
T3	83.4%	73.3%	77.9%
T4	14.1%	17.8%	16.1%
Node negative	67.2%	17.4%	40%
1–4 positive nodes	27.3%	66.7%	48.8%
5+ positive nodes	5.5%	15.9%	11.2%
High Grade	15.3%	16.1%	15.7%
Trials Cohort			
	894651 N=887	914653 N=842	Total N=1729
5FU Chemo	100%	100%	100%
Median Age	65y	63y	64y
Male Gender	51%	53.7%	52.3%
T1/2 stage	10.4%	12.8%	11.6%
T3	75.4%	73.2%	74.3%
T4	14.2%	14%	14.1%
Node negative	17.1%	25.1%	21%
1–3 positive nodes	57.4%	49.8%	53.7%
4+ positive nodes	25.4%	25.2%	25.3%
High Grade	29.3%	26.1%	27.8%

* High Grade = poorly differentiated/anaplastic or Grade 3–4

Table 2
Population-Based Data: Observed and Predicted Outcomes with both calculators

	n	5y metric	Mean % of 5-year outcomes			% delta (pred-obs)		
			NUM	ADJ!	Observed	95% CI	NUM	ADJ!
Surgery alone								
Overall	924	OS	67%	64%	61%	58 - 65	6%	3%
		RFS	61%	60%	58%	54 - 61	3%	2%
Age								
<=49y	51	OS	79%	82%	87%	77 - 97	-8%	-5%
		RFS	67%	78%	81%	70 - 93	-14%	-3%
50-59y	100	OS	75%	78%	80%	72 - 89	-5%	-2%
		RFS	65%	73%	74%	65 - 83	-9%	-1%
60-69y	199	OS	71%	73%	69%	62 - 76	2%	4%
		RFS	64%	69%	64%	57 - 72	0%	5%
>=70y	574	OS	63%	58%	54%	49 - 58	9%	4%
		RFS	58%	53%	50%	46 - 55	8%	3%
Gender								
Male	474	OS	68%	64%	60%	56 - 65	8%	4%
		RFS	62%	60%	56%	52 - 61	6%	4%
Female	450	OS	66%	65%	63%	58 - 68	3%	2%
		RFS	60%	60%	59%	54 - 64	1%	1%
Grade								
High	141	OS	56%	56%	49%	41 - 59	7%	7%
		RFS	49%	51%	48%	40 - 58	1%	3%
Low	783	OS	69%	66%	64%	60 - 67	5%	2%
		RFS	63%	62%	59%	56 - 63	4%	3%
T-stage								
T1/2	23	OS	67%	70%	68%	47 - 89	-1%	2%
		RFS	61%	66%	68%	47 - 89	-6%	-2%
T3	771	OS	69%	67%	66%	63 - 69	3%	1%

	n	5y metric	Mean % of 5-year outcomes			% delta (pred-obs)	
			NUM	ADJ!	Observed	95% CI	NUM
T4	130	RFS	63%	63%	62%	58 – 66	1%
		OS	53%	48%	34%	26 – 42	14%
		RFS	44%	44%	31%	23 – 39	13%
Nodes							
N0	621	OS	76%	74%	72%	69 – 76	4%
		RFS	70%	70%	68%	64 – 72	2%
1-3+	228	OS	53%	50%	48%	42 – 56	2%
		RFS	46%	44%	44%	38 – 52	0%
4-10	67	OS	33%	30%	18%	10 – 30	12%
		RFS	28%	24%	18%	11 – 20	6%
Surgery + SFU							
Overall	1109	OS	66%	66%	64%	61 – 67	2%
		RFS	61%	63%	57%	54 – 60	6%
Age							
<=49y	137	OS	72%	72%	68%	60 – 77	4%
		RFS	60%	69%	63%	55 – 72	-3%
50-59y	243	OS	70%	71%	69%	63 – 76	1%
		RFS	62%	68%	58%	52 – 65	10%
60-69y	387	OS	66%	67%	66%	61 – 71	0%
		RFS	61%	64%	60%	55 – 65	4%
>=70y	342	OS	62%	60%	56%	51 – 62	4%
		RFS	60%	57%	52%	47 – 58	5%
Gender							
Male	610	OS	66%	65%	60%	56 – 64	5%
		RFS	61%	62%	54%	50 – 58	7%
Female	499	OS	67%	68%	68%	63 – 72	-1%
		RFS	60%	65%	61%	57 – 66	-1%
Grade							

	n	5y metric	Mean % of 5-year outcomes				% delta (pred-obs)	
			NUM	ADJ1	Observed	95% CI	NUM	ADJ1
High	178	OS	55%	57%	56%	49 – 64	-1%	1%
		RFS	49%	53%	55%	48 – 63	-6%	-2%
Low	931	OS	69%	68%	65%	62 – 68	4%	3%
		RFS	63%	65%	58%	55 – 61	5%	7%
T-stage								
T1/2	99	OS	75%	79%	84%	77 – 91	-9%	-5%
		RFS	72%	77%	82%	74 – 90	-10%	-5%
T3	813	OS	67%	67%	64%	61 – 67	3%	3%
		RFS	62%	64%	57%	54 – 60	5%	7%
T4	197	OS	59%	59%	53%	46 – 60	6%	6%
		RFS	50%	55%	47%	40 – 54	3%	8%
Nodes								
N0	193	OS	79%	80%	72%	65 – 79	7%	8%
		RFS	72%	77%	66%	59 – 74	6%	11%
1-3+	669	OS	69%	69%	67%	63 – 71	2%	2%
		RFS	63%	66%	60%	56 – 64	3%	6%
4-10	218	OS	51%	52%	49%	43 – 57	2%	3%
		RFS	46%	48%	43%	37 – 51	3%	5%
10+	29	OS	41%	30%	42%	27 – 66	-1%	-12%
		RFS	34%	26%	43%	28 – 67	-9%	-17%

Estimates are rounded to the nearest whole number

* Gray highlight indicates predicted estimate outside 95% CI of KM estimate

Table 3

Population-Based Data:

Percent Closer Rate of Observed to Predicted 5-year rates Across Combined Prognostic Subgroupings of age, T, N and grade *				
	5yr RFS		5yr OS	
	Numeracy	Adjuvant!	Numeracy	Adjuvant!
Surg Alone	44%	56%	12%	88%
Surg + 5FU	62%	38%	55%	45%
Percent Correct Predictions **				
	5yr RFS		5yr OS	
	Numeracy	Adjuvant!	Numeracy	Adjuvant!
Surg Alone	65%	66%	66%	69%
Surg + 5FU	58%	59%	64%	63%

* Cell level analysis with > 10 observations

** Correct if predicted probability of event ≥ 0.5 and patient alive with no recurrence (RFS) or alive (OS) at 5 years; or predicted probability of event recurred or dead (RFS) or death (OS) < 0.5 and patient death at 5 years

Table 4

Trials Data: Observed and Predicted Outcomes with both calculators (All patients received Surgery plus 5FU)

	n	5y metric	Mean % of 5-year outcomes			% delta (pred-obs)		
			NUM	ADJ!	Observed	95% CI	NUM	ADJ!
Overall	1729	OS	66%	68%	68%	66 - 70	-2%	0%
		RFS	60%	65%	61%	58 - 63	-1%	4%
Age								
<=49y	227	OS	69%	71%	68%	62 - 74	1%	3%
		RFS	57%	68%	58%	52 - 65	-1%	10%
50-59y	401	OS	68%	71%	72%	67 - 76	-4%	-1%
		RFS	61%	68%	64%	59 - 68	-3%	4%
60-69y	655	OS	65%	68%	71%	68 - 75	-6%	-3%
		RFS	60%	65%	63%	60 - 67	-3%	2%
>=70y	446	OS	62%	62%	60%	56 - 65	2%	2%
		RFS	59%	59%	55%	51 - 60	4%	4%
Gender								
Male	904	OS	66%	67%	67%	64 - 70	-1%	0%
		RFS	60%	64%	59%	56 - 63	1%	5%
Female	825	OS	65%	68%	69%	66 - 72	-4%	-1%
		RFS	59%	65%	62%	58 - 65	-3%	3%
Grade								
High	480	OS	57%	62%	60%	56 - 65	-3%	2%
		RFS	51%	59%	55%	51 - 60	-4%	4%
Low	1249	OS	69%	70%	71%	68 - 73	-2%	-1%
		RFS	63%	67%	63%	60 - 65	0%	4%
T-stage								
T1/2	200	OS	74%	79%	82%	77-87	-8%	-3%
		RFS	71%	77%	73%	67-80	-2%	4%
T3	1285	OS	66%	68%	67%	65-69	-1%	1%
		RFS	60%	65%	60%	57-62	0%	5%

	n	5y metric	Mean % of 5-year outcomes			% delta (pred-obs)	
			NUM	ADJ!	Observed	95% CI	NUM
T4	244	OS	58%	58%	61%	55-67	-3%
		RFS	49%	54%	54%	48-60	0%
Nodes							
N0	363	OS	78%	79%	83%	79 - 87	-4%
		RFS	72%	76%	77%	72 - 81	-1%
1-3+	928	OS	69%	69%	71%	68 - 74	-2%
		RFS	63%	66%	63%	60 - 66	3%
4-10	420	OS	48%	58%	51%	46 - 56	7%
		RFS	42%	54%	44%	39 - 49	10%
10+	18	OS	43%	34%	24%	10 - 56	10%
		RFS	38%	30%	23%	9 - 57	7%

Estimates are rounded to the nearest whole number

* Gray highlight indicates predicted estimate outside 95% CI of KM estimate

Table 5

Trials Data:

Percent Closer Rate of Observed to Predicted Across Combined Prognostic Subgroupings of age, T, N and grade*				
	5yr RFS		5yr OS	
	Numeracy	Adjuvant!	Numeracy	Adjuvant!
Surg + 5FU	45%	55%	42%	58%
Percent Correct Predictions**				
	5yr RFS		5yr OS	
	Numeracy	Adjuvant!	Numeracy	Adjuvant!
Surg + 5FU	64%	63%	70%	69%

* Cell level analysis with >10 observations

** Correct if predicted probability of event ≥ 0.5 and patient alive with no recurrence (RFS) or alive (OS) at 5 years; or predicted probability of event recurred or dead (RFS) or death (OS) < 0.5 and patient death at 5 years