Published in final edited form as:

Semin Oncol. 2010 February ; 37(1): 39–46. doi:10.1053/j.seminoncol.2009.12.005.

Adjuvant Chemotherapy for Resected Stage II and III Colon Cancer: Comparison of Two Widely Used Prognostic Calculators

Aditya Bardia^a, Charles Loprinzi^b, Axel Grothey^b, Garth Nelson^c, Steven Alberts^b, Smitha Menon^b, Stephan Thome^d, Sharlene Gill^e, and Dan Sargent^c

^aDepartment of Oncology, Johns Hopkins University, Baltimore, MD.

^bDepartment of Oncology, Mayo Clinic, Rochester, MN.

^cDepartment of Biostatistics, Mayo Clinic, Rochester, MN.

^dDepartment of Preventive Medicine, Creighton University, Omaha, NE.

^eUniversity of British Columbia, Vancouver, BC, Canada.

Abstract

Two Web-based prognostic calculators (Adjuvant! and Numeracy) are widely used to individualize decisions regarding adjuvant therapy among patients with resected stage II and III colon cancer. However, these tools have not been directly compared. Hypothetical scenarios were formulated for the Numeracy calculator based on all potential combinations of age, lymph nodes status, tumor stage, and grade of tumor. These were then applied to three postsurgical therapy choices: observation, 5-fluorouracil (5-FU), or FOLFOX (5-FU, leucovorin, and oxaliplatin chemotherapy) to obtain the predicted 5-year disease-free survival (DFS) and overall survival (OS). Wilcoxon signed rank tests were used to compare the numerical predictions between the Adjuvant! and Numeracy calculators for each combination. A total of 192 hypothetical patient scenarios were obtained. For these patients, DFS and OS predictions from Adjuvant! were statistically significantly different than Numeracy ($P < .05$), except for four of 144 categories. While the estimated benefit in DFS and OS for 5-FU compared to surgery obtained from Adjuvant! and Numeracy were similar, the benefit in DFS and OS for FOLFOX over 5-FU, obtained from the Adjuvant! tool was slightly lower than that estimated from Numeracy. Among patients with resected stage II and III colon cancer, the DFS and OS estimates obtained from Numeracy and Adjuvant!, regarding the benefit of 5-FU over surgery, are similar, but the benefits of FOLFOX over 5-FU differ. Validation studies are needed to clarify the discrepancy and to assess the accuracy of these tools for predicting actual patient outcomes.

Colorectal cancer is the third most common cancer in the United States and Canada, with approximately 148,610 new cases of large bowel cancers diagnosed in the United States each year.¹ Surgical resection with curative intent is the standard initial treatment for those with stage I to III colon cancer. The use of 5-FU–based adjuvant therapy with oxaliplatin has been demonstrated to improve disease-free survival (DFS) and overall survival (OS) among patients with stage III disease.^{2,3} The use of adjuvant therapy for patients with stage II disease continues to be debated.⁴

To help individualize decisions regarding adjuvant therapy, prognostic decision models have been developed for colon cancer, similar to those used for breast cancer.^{5–7} These tools aim

to provide an estimate of the prognosis for an individual patient based on predefined patient and tumor characteristics in the absence of postsurgical therapy, as well as an estimate of the net benefit of adjuvant therapy for colon cancer, in terms of DFS and OS.

Two Web-based calculators are currently available to estimate the prognosis of resected colon cancer, with and without adjuvant chemotherapy. One of the tools, called *Numeracy*, was developed at the Mayo Clinic and is available online at www.mayoclinic.com/calcs.⁸ Using a pooled data set of 3,302 patients with stage II and III colon cancer from seven randomized trials comparing 5-fluorouracil (5-FU) + leucovorin or 5-FU + levamisole to surgery alone, the authors found that age, nodal status, T stage, and tumor grade were important prognostic factors. In addition, model-derived estimates of adjuvant 5-FU– based treatment benefit were also derived from this pooled analysis. Data regarding the benefit of FOLFOX (5-FU, leucovorin, and oxaliplatin chemotherapy) versus 5-FU– based treatment for this tool were derived from the hazard ratios observed in the MOSAIC trial.² The other prognostic Web-based tool, called *Adjuvant!*, is available online at <https://www.adjuvantonline.com>. The Adjuvant! tool utilizes prognostic estimates derived from the US Surveillance Epidemiology and End Results (SEER) tumor registry, and therapeutic efficacy estimates derived from the literature.⁵ While both tools aim to provide risk estimates regarding the prognosis of individual patients with resected colon cancer, the comparability of these calculators has not been published.

The purpose of the current study was to assess the whether the Numeracy and Adjuvant! colon cancer prognostic tools produced similar results for patients with resected stage II and III colon cancer, based on a set of hypothetical patients (using all possible combinations of comparable data).

METHODS

The Numeracy calculator has the following input options:

- three choices for lymph nodes (none, 1–4, 5+ lymph nodes),
- three choices for tumor stage (T1/T2, T3, T4),
- two choices for grade (low, high),
- four choices for age (49 years or younger, 50–59, 60–69, 70 or older).

Thus, there are 72 hypothetical patient scenarios ($3 \times 3 \times 2 \times 4$) that account for all potential combinations. Since chemotherapy is not indicated for T1 and T2 tumors in the absence of lymph node involvement (N0), subtracting these variations ($N = 12$), leaves a total of 64 hypothetical patient scenarios. Adding the three postsurgical therapy choices (observation, 5-FU and leucovorin, FOLFOX), 192 (64×3) patient scenarios cover all possible results generated from the Numeracy tool for DFS and OS.

These 192 hypothetical scenarios were entered into the Numeracy program, using an electronic automated system to obtain the relevant outputs. The same hypothetical patients were manually entered into the Adjuvant! tool to obtain comparative outputs. Of note, the Adjuvant! tool has three additional options that are not included in the Numeracy tool: gender (male, female), comorbidity (perfect health, minor problems, average for age, or major problems), and the number of examined lymph nodes (0, 1–3, 4–10, >10). Results were calculated separately for males and females. The default setting for Adjuvant! was “minor problems” as the comorbidity, and thus this was used for the current analysis. However, a sensitivity analysis with “average for age” comorbidity was also performed. The number of examined lymph nodes used for these calculations was fixed at “>10” due to current practice guidelines that specify a minimum of 12 lymph nodes be examined to confirm nodal disease status. The nodal categories for

Adjuvant! and Numeracy differ slightly: for Adjuvant! the categorization “1–3” positive nodes is used, whereas for Numeracy it is “1–4,” these were treated as interchangeable in this analysis. Correspondingly, 4+ nodes in Adjuvant! were treated as identical to 5+ for Numeracy.

Comparative analyses between the two tools are primarily descriptive. The Wilcoxon signed rank test was used to compare the numerical predictions between the calculators for each treatment and outcome (comparing the 192 predictions between the two tools), and for age, gender, T stage, and the number of nodes subgroups.

RESULTS

A total of 192 hypothetical patient scenarios were obtained from Numeracy, and 384 were obtained from Adjuvant! (192 for males and 192 for females). The results are displayed in Figures 1 and 2.

Figures 1A–F and 2A–F display the output (DFS and OS respectively) for different adjuvant interventions from both prognostic tools, for various potential combinations of age, grade of tumor, stage of tumor, or lymph node status of tumor. The figures demonstrate that the vast majority of DFS and OS estimates for prognosis obtained from Adjuvant! were higher than those obtained from Numeracy. The difference was more apparent for females (as compared to males), DFS (as compared to OS), and patients receiving no chemotherapy (as compared to those receiving FOLFOX).

The Wilcoxon signed rank test was used to compare the difference of Adjuvant! versus Numeracy within each group (DFS for surgery only, DFS for 5-FU, DFS for FOLFOX, OS for surgery only, OS for 5-FU, OS for FOLFOX) as outlined in Table 1. Overall, Adjuvant! predicted a higher DFS and OS than Numeracy ($P < .05$ for all comparisons), consistent with the differences observed in Figure 1. Within subgroups (age, T stage, positive nodes, grade), the majority of estimates obtained from the Adjuvant! tool were higher than those obtained from Numeracy tool, with the exception of only four of 144 categories, as indicated in Table 1.

Given that the main purpose of using these tools in clinical practice is to estimate the benefit that may be obtained for different available treatment options, the estimated benefit of 5-FU over surgery, and FOLFOX over 5-FU, was compared between the two calculators. Figures 3A–H display the outputs of differences in DFS and OS for 5-FU versus surgery, and FOLFOX versus 5-FU, from both the tools, for various potential combinations of age, grade of tumor, stage of tumor, or lymph node status of tumor. The figures suggest that the estimates of DFS and OS benefit for 5-FU compared to surgery, obtained from the Adjuvant! tool were generally similar to that obtained from the Numeracy tool for both males and females. However, the benefit for DFS and OS for FOLFOX compared to 5-FU, obtained from Adjuvant! tended to be lower than the estimate obtained from Numeracy. Similar results were obtained when Wilcoxon signed rank tests were used, as outlined in Table 2.

When sensitivity analyses were conducted using the “average for age” comorbidity in Adjuvant! (rather than “minor for age” comorbidity), the results were similar.

DISCUSSION

This study, based on a set of hypothetical patients (using all possible combinations of comparable data), suggests that the two currently available Web-based colon cancer prognostic tools, Numeracy and Adjuvant!, produce similar results for predicting the numerical benefit in DFS and OS of using 5-FU chemotherapy over surgery alone. The results for the advantage of FOLFOX over 5-FU are not as similar, with the estimated benefit obtained using the Adjuvant!

tool typically lower than the predicted estimate derived from Numeracy. It should be noted that while the results from the prognostic tools differ, they are each likely to be much more accurate for individual patient prognoses than would be expected from clinician estimates without them. This latter statement is based, in part, on data collected from clinicians attempting to estimate individual patient prognoses for resected breast cancer.⁹

Why Is There a Difference in the Estimates of the Two Calculators?

There are several potential explanations for the differences between the two calculators. Foremost, the differences in baseline prognosis estimates for patients not receiving adjuvant chemotherapy could be due to differences in the populations used to generate the calculators. The Numeracy tool is based on data from patients enrolled in randomized clinical trials (in the United States and Europe), while the Adjuvant! tool is based on information from US-based community-derived data (SEER tumor registry). While patients in clinical trials generally have a better performance status than the general population, this is unlikely to account for the observed results, as the comparison of the tools found that Adjuvant! had a higher estimated DFS and OS than Numeracy (rather than the opposite, if the performance status hypothesis were true).

The difference in baseline prognosis could be related to the fact that SEER-based data are unable to identify which patients received adjuvant chemotherapy. To correct for this, mathematical adjustments were made to the baseline prognosis (for stage II patients only) within the Adjuvant! program based on assumed proportions of patients treated and the associated benefit of adjuvant chemotherapy. In addition, relapse data are not available in the SEER tumor registry, therefore Adjuvant! indirectly estimates the DFS by making certain assumptions about the general survival of patients after relapse, and the observed mortality to indirectly calculate the relapse, which is always greater than the risk of colon cancer specific mortality.

Another possible reason for the differences is that Numeracy uses 10-year age ranges while Adjuvant! uses individual age in years. Similarly, a difference could be related to Numeracy having fewer choices (12) as compared to Adjuvant! (24 along with age being a continuous variable). This could, in particular, explain the slight difference in the estimates obtained for males and females. The data on which the Numeracy tool are based are from clinical trials conducted in the late 1980s, compared to more recent SEER data used for Adjuvant!. While the standard of care for the adjuvant treatment of colon cancer did not change in that period, other temporal changes in health care or other factors could influence the patient baseline prognosis.

A further reason for the difference in tools could be the Will Rogers effect of including 1–4 positive nodes in one group in Numeracy and 1–3 positive nodes in Adjuvant!. This choice would be expected to result in improved prognostic estimates for both node-positive groups (1–3 and 4+) in Adjuvant!. Brief reflection on Rogers' famous quip shows why: "When the Okies left Oklahoma and went to California, the average intelligence of both states went up."¹⁰ Still, this effect can not explain the observed differences in the node negative population prognostic estimates, nor any predictive differences noted.

Perhaps the most clinically relevant and important question to assess is which of these calculators is actually better at predicting the increase in DFS and OS when using chemotherapy, over surgery alone. The subject of tool accuracy has been addressed for the Adjuvant! tool for breast cancer. Olivotto et al¹¹ used the British Columbia Breast Cancer Outcomes Unit (BCOU) database of 4,083 women with breast cancer (T1-2, N0-1, M0) diagnosed between 1989 and 1993 in British Columbia, Canada. The relevant demographic, staging, and treatment data on these women were inserted into the Adjuvant! tool for breast

cancer to calculate the predicted 10-year OS and DFS. Overall, the estimated outputs from the Adjuvant! tool were within 1% of the actual OS and DFS, and were within 2% for most subgroups. Conducting a similar validation study for the two colon cancer prognostic tools would be very helpful.

This study has limitations. Not all possible patient-level combinations could be compared as Numeracy has fewer input options (12) as compared to Adjuvant! (24, along with age being a continuous variable). However, this should not alter the results substantially as the most important prognostic variables were present in both the calculators. Moreover, sensitivity analyses were conducted to explore gender and comorbidity differences. Finally, while this study suggests that the estimates obtained from Adjuvant! are slightly different from Numeracy for benefit of FOLFOX over surgery, it does not (and cannot) address which of the tools most accurately reflects clinical reality.

In conclusion, among hypothetical case scenarios of patients with resected stage II and III colon cancer, the DFS and OS estimates obtained from Numeracy and Adjuvant! regarding the benefit of 5-FU over surgery are similar, but the benefit estimates of FOLFOX over 5-FU do differ. The clinical significance of this difference in predicted benefit requires validation with a real-world cohort of patients. Such a population-based validation is underway.

Acknowledgments

This work was supported by the following United States National Institutes of Health Grant-CA 124477 (PI Charles Loprinzi, MD).

REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ. CA Cancer statistics, 2008. *Cancer J Clin* 2008;58:71–96.
2. André T, Boni C, Mounedji-Boudiaf L, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. Multicenter International Study of Oxaliplatin/5-Fluorouracil/Leucovorin in the Adjuvant Treatment of Colon Cancer (MOSAIC) Investigators. *N Engl J Med* 2004;350:2343–2351. [PubMed: 15175436]
3. de Gramont A, Tournigand C, André T, Larsen AK, Louvet C. Adjuvant therapy for stage II and III colorectal cancer. *Semin Oncol* 2007;34:S37–S40. [PubMed: 17449351]
4. Figueredo A, Coombes ME, Mukherjee S. Adjuvant therapy for completely resected stage II colon cancer. *Cochrane Database Syst Rev* 2008;CD005390. [PubMed: 18646127]
5. Ravdin PM, Siminoff LA, Davis GJ, et al. Computer program to assist in making decisions about adjuvant therapy for women with early breast cancer. *J Clin Oncol* 2001;19:980–991. [PubMed: 11181660]
6. Loprinzi CL, Thome SD. Understanding the utility of adjuvant systemic therapy for primary breast cancer. *J Clin Oncol* 2000;19:972–979. [PubMed: 11181659]
7. Loprinzi CL, Ravdin PM. Decision-making for patients with early breast cancer: individualized decisions for and by patients and their physicians. *JNCCN* 2003;1:189–196. [PubMed: 19768877]
8. 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]
9. 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.
10. Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med* 1985;312:1604–1608. [PubMed: 4000199]

11. 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]

Figure 1.

(A–F) Correlation of DFS estimates for different therapeutic modalities, ie, surgery alone, surgery + 5-FU, and surgery + FOLFOX, between the Adjuvant! and Numeracy colon cancer prognostic tools for all potential combinations of age, grade of tumor, stage of tumor or lymph node status of tumor, among males and females. Curved lines represent a lowess smoothed estimate of association.

Figure 2.

(A–F) Correlation of OS estimates for different therapeutic modalities, ie, surgery alone, surgery + 5-FU, and surgery + FOLFOX, between the Adjuvant! and Numeracy colon cancer prognostic tools for all potential combinations of age, grade of tumor, stage of tumor or lymph node status of tumor, among males and females. Curved lines represent a lowess smoothed estimate of association.

Figure 3.

(A–H) Calculated absolute difference in DFS and OS estimates for 5-FU over surgery, and FOLFOX over 5-FU, from both Adjuvant! and Numeracy colon cancer prognostic tools for all potential combinations of age, grade of tumor, stage of tumor or lymph node status of tumor, among males and females. Curved lines represent a lowess smoothed estimate of association.

Table 1

Number of Patients Who Had Estimates Higher From Adjuvant! as Compared to Numeracy for DFS and OS by Therapeutic Modalities and Other Characteristics

	Age Group (yr)				Grade		Lymph Node-Positive			Stage of Tumor			
	50-59 (n = 16)	60-69 (n = 16)	≤49 (n = 16)	≥70 (n = 16)	High (n = 32)	Low (n = 32)	0 (n = 16)	1-4 (n = 24)	5+ (n = 24)	T1/2 (n = 16)	T3 (n = 24)	T4 (n = 24)	Total (n = 64)
DFS for surgery only													
Male	16 (100%)	16 (100%)	16 (100%)	14 (87.5%)	32 (100%)	30 (93.8%)	15 (93.8%)	23 (95.8%)	24 (100%)	15 (93.8%)	23 (95.8%)	24 (100%)	62 (96.9%)
Female	16 (100%)	16 (100%)	16 (100%)	16 (100%)	32 (100%)	32 (100%)	16 (100%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	24 (100%)	64 (100%)
DFS for 5-FU													
Male	16 (100%)	16 (100%)	16 (100%)	11 (68.8%)	31 (96.9%)	28 (87.5%)	14 (87.5%)	21 (87.5%)	24 (100%)	14 (87.5%)	22 (91.7%)	23 (95.8%)	59 (92.2%)
Female	16 (100%)	16 (100%)	16 (100%)	16 (100%)	32 (100%)	32 (100%)	16 (100%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	24 (100%)	64 (100%)
DFS for FOLFOX													
Male	16 (100%)	16 (100%)	16 (100%)	8* (50%)	29 (90.6%)	27 (84.4%)	13 (81.3%)	20 (83.3%)	23 (95.8%)	13 (81.3%)	20 (83.3%)	23 (95.8%)	56 (87.5%)
Female	16 (100%)	16 (100%)	16 (100%)	15 (93.8%)	32 (100%)	31 (96.9%)	16 (100%)	23 (95.8%)	24 (100%)	15 (93.8%)	24 (100%)	24 (100%)	63 (98.4%)
OS for surgery only													
Male	16 (100%)	16 (100%)	16 (100%)	13 (81.3%)	32 (100%)	29 (90.6%)	14 (87.5%)	23 (95.8%)	24 (100%)	15 (93.8%)	23 (95.8%)	23 (95.8%)	61 (95.3%)
Female	16 (100%)	16 (100%)	15 (93.8%)	16 (100%)	32 (100%)	31 (96.9%)	15 (93.8%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	23 (95.8%)	63 (98.4%)
OS for 5-FU													
Male	16 (100%)	16 (100%)	16 (100%)	11 (68.8%)	31 (96.9%)	28 (87.5%)	14 (87.5%)	21 (87.5%)	24 (100%)	14 (87.5%)	22 (91.7%)	23 (95.8%)	59 (92.2%)
Female	16 (100%)	16 (100%)	16 (100%)	16 (100%)	32 (100%)	32 (100%)	16 (100%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	24 (100%)	64 (100%)
OS for FOLFOX													
Male	16 (100%)	12 (75%)	16 (100%)	6* (37.5%)	28 (87.5%)	22 (68.8%)	10* (62.5%)	17* (70.8%)	23 (95.8%)	12 (75%)	18 (75%)	20 (83.3%)	50 (78.1%)
Female	16 (100%)	16 (100%)	16 (100%)	12 (75%)	32 (100%)	28 (87.5%)	14 (87.5%)	22 (91.7%)	24 (100%)	15 (93.8%)	22 (91.7%)	23 (95.8%)	60 (93.8%)

“%” represents the percentage of patients that had estimates higher from Adjuvant!, as compared to Numeracy. N represents the number of hypothetical patients in each group.

* Values were not statistically significantly different between the two tools based on Wilcoxon signed rank test. The rest of the values were statistically significantly different, ie, the estimates were significantly higher from Adjuvant! as compared to Numeracy ($P < .05$).

Table 2

Number of Patients Who Had Estimates Lower From Adjuvant! as Compared to Numeracy for DFS and OS by Therapeutic Modalities and Other Characteristics

	Age Group (yr)			Grade		Lymph Node-Positive			Stage of Tumor			
	50-59 (n = 16)	60-69 (n = 16)	≥49 (n = 16)	High (n = 32)	Low (n = 32)	0 (n = 16)	1-4 (n = 24)	5+ (n = 24)	T1/2 (n = 16)	T3 (n = 24)	T4 (n = 24)	Total (n = 64)
DFS benefit of 5-FU over surgery only												
Male	11* (68.8%)	11* (68.8%)	10* (62.5%)	21* (65.6%)	24 (75%)	16 (100%)	20 (83.3%)	9* (37.5%)	16 (100%)	17 (70.8%)	12* (50%)	45 (70.3%)
Female	11* (68.8%)	10* (62.5%)	11* (68.8%)	19* (59.4%)	24 (75%)	15 (93.8%)	19 (79.2%)	9 (37.5%)	16 (100%)	17* (70.8%)	10* (41.7%)	43* (67.2%)
DFS benefit of FOLFOX over 5-FU												
Male	15 (93.8%)	16 (100%)	16 (100%)	32 (100%)	31 (96.9%)	15 (93.8%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	23 (95.8%)	63 (98.4%)
Female	16 (100%)	16 (100%)	16 (100%)	32 (100%)	32 (100%)	16 (100%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	24 (100%)	64 (100%)
OS benefit of 5-FU over surgery only												
Male	9* (56.3%)	12 (75%)	6* (37.5%)	16* (50%)	23 (71.9%)	11 (68.8%)	17 (70.8%)	11* (45.8%)	16 (100%)	16 (66.7%)	7* (29.2%)	39 (60.9%)
Female	10* (62.5%)	11* (68.8%)	6* (37.5%)	17* (53.1%)	23 (71.9%)	10 (62.5%)	18 (75%)	12* (50%)	16 (100%)	16 (66.7%)	8* (33.3%)	40 (62.5%)
OS benefit of FOLFOX over 5-FU												
Male	16 (100%)	16 (100%)	16 (100%)	32 (100%)	32 (100%)	16 (100%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	24 (100%)	64 (100%)
Female	16 (100%)	16 (100%)	15 (93.8%)	32 (100%)	31 (96.9%)	15 (93.8%)	24 (100%)	24 (100%)	16 (100%)	24 (100%)	23 (95.8%)	63 (98.4%)

% represents the percentage of patients who had estimates lower from Adjuvant! as compared to Numeracy. N represents the number of hypothetical patients in each group.

* Values were not statistically significant based on Wilcoxon signed rank test. The rest of the values were statistically significant, ie, the estimates were significantly lower from Adjuvant! as compared to Numeracy, with the exception of 5+ positive lymph nodes for DFS benefit of 5-FU over surgery alone for female where Adjuvant! was significantly higher as compared to Numeracy.