



Published in final edited form as:

Cancer. 2010 April 1; 116(7): 1691–1698. doi:10.1002/cncr.24913.

## Model-based prediction of defective DNA mismatch repair using clinicopathological variables in sporadic colon cancer patients

Frank Sinicrope, MD<sup>1,2</sup>, N.R. Foster, MS<sup>4</sup>, D.J. Sargent, PhD<sup>4</sup>, S.N. Thibodeau, PhD<sup>3</sup>, T.C. Smyrk, MD<sup>3</sup>, M.J. O'Connell, MD<sup>1</sup>, and the North Central Cancer Treatment Group

<sup>1</sup>Divisions of Oncology and Gastroenterology, Mayo Clinic, Rochester, MN

<sup>2</sup>Fiterman Center for Digestive Diseases, Mayo Clinic, Rochester, MN

<sup>3</sup>Division of Laboratory Medicine & Pathology, Mayo Clinic, Rochester, MN

<sup>4</sup>Biostatistics and Informatics, Mayo Clinic, Rochester, MN

### Abstract

**Background**—Colon cancers with defective DNA mismatch repair (MMR) have a favorable prognosis and may lack benefit from 5-fluorouracil-based adjuvant chemotherapy. We developed models to predict MMR deficiency in sporadic colon cancer patients using routine clinical and pathological data.

**Methods**—TNM stage II and III colon carcinomas (n= 982) from six 5-fluorouracil-based adjuvant therapy trials were analyzed for microsatellite instability and/or MMR protein expression. Tumor infiltrating lymphocytes (TILs) were quantitated (n= 326). Logistic regression and a recursive partitioning and amalgamation (RPA) analysis were used to identify predictive factors for MMR status.

**Results**—Defective MMR was detected in 147 (15%) cancers. Tumor site and histologic grade were the most important predictors of MMR status. Distal tumors had a low likelihood of defective MMR (3%; 13/468); proximal tumors had a greater likelihood (26%; 130/506). Using tumor site, grade, and gender, the logistic regression model showed excellent discrimination (c-statistic = 0.81). Proximal site, female gender, and poor differentiation showed a positive predictive value (PPV) of 51% for defective MMR. In a patient subset (n= 326), a model including proximal site, TILs (> 2 / HPF), and female gender showed even better discrimination (c-statistic =0.86) with a PPV of 81%.

**Conclusions**—Defective MMR is rare in distal, sporadic colon cancers that should generally not undergo MMR testing. Proximal site, poor differentiation, and female gender detect 51% of tumors with defective MMR; substituting TILs for grade increases the PPV to 81%. These data can increase the efficiency of MMR testing to assist in clinical decision-making.

### Keywords

microsatellite instability; DNA mismatch repair; colon cancer; sporadic; tumor infiltrating lymphocytes

## Introduction

While the majority of sporadic colorectal cancers (CRCs) show chromosomal instability, an alternative pathway of tumorigenesis is characterized by defective DNA mismatch repair (MMR) 1. Microsatellite instability (MSI) is a molecular marker of defective function of the MMR system 1, 2. Most CRCs with high frequency MSI (MSI-H) are sporadic wherein the MMR defect develops due to inactivation of the *hMLH1* gene by DNA methylation 1, 3, 4. Of the approximately 150,000 incident colon cancer cases that were diagnosed in the United States in 2008 5, at least 20,000 patients are expected to be sporadic MMR deficient tumors. The identification of patients with sporadic MMR deficient tumors has potentially important clinical implications. Multiple retrospective studies 6–9, including a population-based study 10 and a meta-analysis 11, have demonstrated that patients with MMR deficient colon cancers have a more favorable stage-adjusted prognosis compared to patients whose tumors have intact MMR. While the explanation for the better outcome of MMR deficient colon cancers is poorly understood, we previously reported that a higher density of TILs, most of which are CD3+ T lymphocytes 12, was associated with better disease-free survival in these cases with defective versus intact MMR 13. Evidence also indicates that patients with colon cancers with defective MMR do not benefit from 5-fluorouracil (5-FU)-based adjuvant therapy in contrast patients whose tumors exhibit intact MMR 11, 14–18. More recently, a retrospective study and pooled analysis has validated the prognostic and predictive impact of MMR status in patients treated in 5-FU-based adjuvant therapy trials in North America and Europe 19. In these studies, MMR status has been determined by MSI testing or by immunohistochemical (IHC) evaluation of MMR protein expression. MSI testing requires a molecular laboratory and is analyzed by detecting instability at selected microsatellite loci in tumor tissue 2, 20. Analysis of tumors for MMR proteins by IHC is an alternative or complimentary test, performed in a routine pathology laboratory, and has been shown to be highly concordant with MSI testing 21.

MMR deficient colon cancers have distinct clinical and pathological features that include proximal colon predominance, poor differentiation or medullary-type pattern with or without mucinous histology, and intratumoral lymphocytic infiltration 6–8, 15, 22–25. Prior studies have examined histopathological features in an effort to prioritize sporadic colon cancers for MSI testing, yet morphological prediction of MSI-H had low sensitivity 22. Sporadic MMR deficient tumors are distinct epidemiologically from Lynch Syndrome cases in that they are associated with older age at diagnosis, female gender, and with cigarette smoking 4, 25, 26. However, established criteria to identify and select sporadic colon cancers for MSI testing are lacking. This limitation represents an impediment to the use of MMR status for prognostication and patient management in clinical practice. In contrast, Amsterdam and Bethesda criteria were developed to identify patients with Lynch Syndrome 27, 28, as germline MMR mutations are detected in only 2.8% of CRC cases 29. The revised Bethesda guidelines for Lynch syndrome recommend MSI testing for all CRCs in patients diagnosed before age 50 years and for those cancers diagnosed between ages 50 and 59 years with particular pathology features 28, 30. In a recent study, an algorithm was developed for identifying Lynch Syndrome patients using morphological tumor features that was restricted to patients under the age of 60 31. However, there is no such algorithm or predictive model to identify sporadic MMR deficient colon cancers.

The potential clinical utility of MMR status as a prognostic and predictive marker provides a strong impetus for developing models to predict the presence of MMR. Therefore, the objective of our study was to develop a predictive model for identifying MMR deficiency in sporadic colon cancer patients. Since only a subset of sporadic colon cancers show defective MMR, universal screening would be neither specific nor cost effective. Furthermore, such a

model should utilize routine clinical and pathological data to have the most utility in clinical practice.

## Materials and Methods

### Study Population

The study population consisted of 982 patients with surgically resected TNM stage II and III colonic adenocarcinomas enrolled in six randomized 5-fluorouracil-based phase III adjuvant studies conducted by Mayo Clinic and the North Central Cancer Treatment Group (NCCTG). Details of these completed, clinical trials have been previously reported 32, 33. Clinical and pathological data, collected in accordance with study protocol requirements, was obtained from the NCCTG study database. Available paraffin-embedded tumor blocks from a nonrandom subset (n= 982) of all study participants that contained sufficient tumor tissue for analysis were utilized. Primary tumor site was defined relative to the splenic flexure, and tumors located at the splenic flexure were included in the distal category. The current analysis was in accordance with the original informed consent document. Of the 982 patients, 821 were randomized to treatment arms and 161 received observation alone.

### Analysis of DNA Mismatch Repair

Primary colon carcinomas were analyzed for defective DNA mismatch repair (MMR) by molecular analysis of MSI, as described below, and/or by determination of MMR protein expression by immunohistochemistry (IHC).

**MSI Testing**—Tumors (n= 423) had been analyzed for MSI by detecting instability at selected microsatellite loci as previously described 6, 8, 25, 33. In 200 of these cases, IHC was also performed as described below. The tumors analyzed were composed of at least 60% tumor cells and had a minimum of five markers that successfully amplified for both normal DNA and tumor DNA. Given the preponderance of evidence indicating that tumors with MSI-L are not biologically distinct from MSS, we grouped them together in all analyses 7, 8, 25, 34–36. For patients with MSI data, tumors were classified as MSI-H, MSI-L or MSS as previously described 8, 25. In one adjuvant study (NCCTG 91-46-53) 33, MMR status was determined by analysis of instability at *BAT26* and MMR protein expression (hMLH1, hMSH2, hMSH6). In this study (n=396), defective MMR was defined by the presence of MSI at the marker *BAT 26* coupled with absence of protein expression for hMLH1, hMSH2, or hMSH6. For patients that were stable at *BAT 26* and showed normal protein expression for all three proteins, 16 additional microsatellite markers (*myc-L*, *D8S262*, *D8S1742*, *D8S261*, *D8S254*, *D8S133*, *D8S136*, *D8S560*, *D8S1055*, *D8S1820*, *D8S255*, *ANK1*, *D8S1760*, *D8S1720*, *D8S1842*, and *D8S1925*). If >30% of the markers showed MSI, then that tumor was classified as MSI-H and additional IHC testing was done with antibody for PMS2.

**MMR Protein Expression**—Paraffin-embedded tumors (n= 758) were analyzed for hMLH1 and MSH2 proteins, and hMSH6 (n= 387) and PMS2 were analyzed in a tumor subset. The IHC methods and antibodies utilized were as previously described 37. Slides were scored by a pathologist as showing either positive or negative based upon the presence (+) or absence (–) of nuclear staining for each MMR protein in the tumor cells. Each slide contained a unique number that enabled blinding with respect to patient identity and clinical characteristics.

### Histopathology

All colon cancers were evaluated by pathologists at Mayo Clinic or at NCCTG participating institutions. Tumor histologic grade was categorized as grade 1: well differentiated, grade 2:

moderately differentiated, grade 3: poorly differentiated, grade 4: undifferentiated 38. As required by the study protocols, representative blocks of normal and tumor tissue were sent to Mayo Clinic, Rochester for confirmation of the diagnosis of invasive adenocarcinoma of the colon. Histologic parameters were evaluated on hematoxylin and eosin (H&E)-stained slides without knowledge of MSI status.

Tumor Infiltrating Lymphocytes (TILs) were analyzed in a patient subset (n= 326) by morphology in (H&E)-stained tissue sections at light microscopy at Mayo Clinic. The number of TILs was determined by counting 5 consecutive high power (40X) fields (HPFs). The mean number of TILs was expressed as an average per HPF and categorized as follows: 1+ (0–1), 2+ (2–4), 3+ (5–8), and 4+ ( $\geq$ 9) 13.

### Statistical Analysis

The Chi-square ( $\chi^2$ ) test was used to test for an association between clinicopathological features, and the Wilcoxon Rank Sum tests were used to correlate 2-level categorical variables with continuous data. Logistic regression and a recursive partitioning and amalgamation (RPA) analysis was used to identify important predictive factors of MMR status. Factors explored included age, gender, histologic grade, tumor site, stage, lymph node metastases, T-stage and TILs. Model discrimination (i.e., ability to discriminate patients with different MMR outcomes) was evaluated using the c-statistic. The c-statistic is analogous to the area under the receiver-operating characteristic curve and it ranges from .5 to 1 where a c-statistic of 1 provides perfect discrimination and a c-statistic of 0.50 provides predictions that are no better than chance alone. A c-statistic value that ranges from 0.70 to 0.80 shows acceptable discrimination, while a c-statistic value that ranges from 0.80 to 0.90 show excellent discrimination 39. Using MMR status as the gold standard diagnosis, we also assessed the sensitivity, specificity, positive predictive value (PPV), and negative predictive values (NPV) for specific model categories of interest. Statistical tests were two-sided, with  $p \leq 0.05$  considered significant. P-values were not adjusted for multiple comparisons. Statistical analyses were performed using SAS (SAS Institute, Cary, NC).

### Results

A total of 982 stage II and III colon cancer patients were included in this study from six completed adjuvant therapy trials. Defective MMR, defined as either MSI-H and/or loss of MMR protein expression was detected in 147 (15%) patients. Both MSI and MMR protein data were obtained on 200 tumors and the results were concordant in 98% of cases. MSI-H was detected in 75 of 423 (17.7%) of tumors tested and loss of a MMR protein(s) was found in 72 of 559 (12.9 %). MLH1 loss was observed in 60 of these 72 (83.3 %) cases, MSH2 and/or MSH6 loss was detected in nine cases, and two cases showed loss of PMS2.

A comparison of clinicopathological features in tumors with defective versus intact MMR is shown in Table 1. Defective MMR status was significantly associated with proximal tumor site, poor/undifferentiated histology, female gender, older age at diagnosis, and fewer lymph node metastases compared to tumors with intact MMR (Table 1). We found that 91% of cancers with defective MMR were located in the proximal colon. Older age was found to be significantly associated with defective MMR. Patients whose tumors showed defective MMR were more likely to be of older age compared to tumors with intact MMR, using cut points of age 60 ( $p=0.0591$ ) or age 70 ( $p=0.0053$ ) years.

We utilized clinical and pathological variables to construct a model to predict defective MMR, with the objective of increasing the efficiency of MMR testing. Logistic regression and recursive partition amalgamation (RPA) analyses, identified tumor site as the most important predictor of MMR status followed by histologic grade (Table 2). Distal tumor site

was associated with a low likelihood of defective MMR (3% rate overall; 13/468), whereas proximal tumors had a greater likelihood of defective MMR (26%; 130/506). Regardless of the combination of covariates, the likelihood of a distal colon cancer showing defective MMR did not exceed 5%, thus all distal tumors were combined into a single category for prediction purposes (Table 3). Patients with proximal colon cancers could be further categorized based on histologic grade and gender [see Table 3]. Using the three factors of tumor site, histologic grade, and gender, the logistic regression model showed excellent discrimination or predictive ability (c-statistic = 0.81) for MMR status across the five subgroups (Table 3). RPA analysis gave similar results, where tumor site and grade were the most important variables for predicting MMR status (data not shown).

The positive predictive value (PPV) for defective MMR using the combination of proximal tumor site, poor/undifferentiated histopathology, and female gender was 51%, 95% confidence interval (40 – 61%). Sensitivity for this same category was 32%, specificity was 95%, and negative predictive value (NPV) was 89%. In an effort to restrict this model to sporadic colon cancer patients, we excluded patients (n= 26) whose tumors showed loss of MSH2, MSH6 or PMS2 expression, along with patient tumors that were negative for MLH1 and < 60 years of age. This subset of 941 patients yielded similar results to those observed in the full study cohort with a c-statistic of 0.84 and a PPV of 48% for the combination of proximal site, poor/undifferentiated histology, and female gender. As an alternative approach, we limited our analysis to all patients at least 60 years of age and found that the PPV for this model was 56% (c-statistic was 0.84).

We then evaluated the ability of TILs, a feature of MMR deficient colon cancers 23-40, to increase the utility of the predictive model. TIL density within colon cancers was determined in a patient subset from which 58 of 326 (18%) patients analyzed showed defective MMR. Using logistic regression and RPA analyses, we found that both TILs and tumor site were important predictors of MMR status, followed by histologic grade and gender (Table 4). We then integrated TILs into the predictive model, at which point tumor grade did not contribute to prediction and was therefore removed. This simplified model included tumor site, TILs, and gender and provided excellent discrimination (c-statistic 0.86) with only four categories. The observed rate of defective MMR varied from a low of 3% in the distal patient group to a high of 81% in the proximal, female gender, and increased TILs group (Table 5). Specifically, the combination of proximal site, female gender and TILs ( $\geq 2$ /HPF) yielded a PPV of 81% for detecting defective MMR (Table 5). This same group yielded a sensitivity of 38%, a specificity of 98% and a NPV of 88%. This model showed the highest efficiency for selecting patients for MMR testing as compared to the model developed without TILs on the full study cohort. Male gender in this model with proximal site and increased TILs ( $> 2$ +/HPF) resulted in a PPV of 50%. When excluding patients whose tumors showed loss of MSH2, MSH6, or PMS2 and those that were negative for MLH1 and < 60 years of age, modeling of this subset of 312 patients yielded similar results as did the full study cohort (c-statistic = 0.86, PPV of 81%). Alternatively, when including only those patients  $\geq$  age 60, the PPV was still 81% (c-statistic = 0.84). We emphasize that TILs were determined in routine histological sections and can be analyzed at the time of diagnostic pathology assessment. Therefore, we believe that this model is practical and clinically relevant.

## Discussion

Currently, there are no guidelines used to detect patients whose sporadic colon cancers exhibit MMR deficiency, the majority of which occur in patients over the age of 60 years as shown here. The importance of recognizing patients with sporadic colon cancers with defective MMR is gaining acceptance due to the recognition that MMR status can be used



for prognostication and may influence treatment decisions. A recent retrospective pooled analysis has validated the prognostic and predictive impact of the MMR status in colon cancer patients from completed adjuvant therapy trials conducted the U.S. and Europe 19. Based upon these data and prior studies, we anticipate that the clinical demand for MSI testing in sporadic colon cancer patients is likely to increase in the near future. Consistent with population-based studies, the prevalence of defective MMR in stage II and III colon cancers was 15% 10. We found that the MMR phenotype was significantly associated with lower tumor stage, proximal site, poor/undifferentiated histology, female gender and older age as has been consistently reported. In addition, 91% of tumors with defective MMR were located in the proximal colon in contrast to those with intact MMR. Of note, prior data suggest that sporadic tumors with defective MMR are more frequently proximal compared to those in Lynch Syndrome 41· 42. Sporadic MMR deficient cancers are distinct from Lynch Syndrome cases in that they show widespread DNA hypermethylation and appear to originate within serrated polyps with *BRAF* mutations 41· 43.

The screening of every colorectal cancer patient for MSI and/or MMR protein expression has been advocated, but would be a nonspecific and expensive approach. Therefore, the goal of this study was to develop a model to accurately identify patients whose sporadic colon cancer would be likely to demonstrate MMR deficiency, thereby increasing the efficiency of MSI testing to enable such information to be used in clinical decision-making. Among stage II and III colon cancer patients who participated in adjuvant chemotherapy trials, we found a 15% prevalence of defective MMR that is consistent with population-based studies 10· 31. We used logistic regression and RPA analyses of clinical and routine pathological information from the adjuvant studies in an attempt to predict a patient's MMR tumor status. We found that tumor site was the most important predictor of defective MMR status followed by histological grade in the full study cohort. Defective MMR was rare in distal and increased in proximal tumors such that distal colon cancers do not warrant screening for defective MMR based on our models. The combination of proximal site, poor differentiation, and female gender resulted in a 51% likelihood of defective MMR. In a study that evaluated several morphological features of MMR deficient colon cancers, interobserver agreement among five pathologists was good only for evaluation of poor differentiation (kappa, 0.69) 22. These data support the reproducibility of our model and we emphasize that our study utilized routine diagnostic pathology reports generated by multiple pathologists from institutions affiliated with our cooperative group clinical trials mechanism. We emphasize that our predictive models apply only to sporadic colon cancer patients, and are not intended to have utility in Lynch Syndrome cases.

Based on studies that have demonstrated that TILs are a marker for colon cancers with defective MMR 23 and in an attempt to further improve upon the PPV, we analyzed and dichotomized TIL counts in routine histological tumor sections from a patient subset (n= 326). Substituting TILs ( $\geq 2$ /HPF) for tumor grade in our three variable model increased the PPV, in the cohort with the greatest likelihood of a defective MMR tumor (proximal, TILs ( $\geq 2$ /HPF), female), to 81%. Substituting male for female gender reduced the PPV to 50%, which remains a high rate sufficient to justify MMR testing. PPV is the most important measure of a diagnostic method as it indicates the probability that a positive test reflects the underlying condition being tested for. Therefore, our model with the category of proximal tumor site, high TILs ( $\geq 2$ /HPF), and female gender allows the accurate prediction of the presence of defective MMR in about eight of ten cases with those characteristics, and in males in one of two cases. If validated, this model suggests that TIL counts should become a component of the diagnostic pathology review of all proximal colon cancers in clinical practice.

Our findings suggest that anatomic tumor site and TILs are the most important discriminating variables for MMR status 40–44. In prior studies that evaluated several morphological features of MMR deficient colon cancers and involved expert gastrointestinal (GI) pathologists, poor differentiation, medullary carcinoma, and intraepithelial lymphocytosis were the best discriminators between MSI-H and microsatellite-stable cancers with high specificity but very low sensitivity and poor interobserver agreement except for evaluation of poor differentiation 22. These findings underscore the need for a simplified strategy, as proposed here, for identifying patients whose sporadic colon cancers may exhibit defective MMR. While an independent review of the histopathology by gastrointestinal pathologists and the evaluation of a greater number of morphological tumor features has the potential to further improve the PPV of our model, such an approach cannot easily be implemented into routine clinical practice and a more simplified approach is therefore more practical. Based upon aforementioned factors, we believe that our proposed model is likely able to be generalized and reproducible for selecting sporadic colon cancer cases for MSI testing. Nonetheless, independent validation of our model is warranted. Our models are intended to be used as an adjunct to clinical judgment and are not meant to be the exclusive indicator of which patients should receive testing for MMR.

In summary, anatomic subsite and TILs were the most predictive markers of defective MMR status in sporadic cases. Based upon the data presented here, we recommend that only proximal colon cancers be tested for MMR deficiency. If proximal tumors with poor differentiation from females are tested, approximately half will show defective MMR. If we substitute TIL counts for histologic grade, we can accurately predict MMR deficient cancer in 80% of the female patients and 50% of the male patients whose tumors are proximal with increased TILs. The use of the proposed predictive models can greatly increase the efficiency of MMR testing in sporadics which has obvious implications for reducing costs. We emphasize that the proposed models include data that can be obtained at the time of routine diagnostic pathology review. The identification of patients whose colon cancers exhibit defective MMR for use in clinical decision-making would represent an important step toward individualized cancer medicine.

## Acknowledgments

Funded by: National Cancer Institute grant CA104683-02 (FAS) and Mayo Clinic Cancer Center core grant CA15083.

## References

1. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology*. 2008; 135(4):1079–1099. [PubMed: 18773902]
2. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. *Science*. 1993; 260(5109):816–819. [PubMed: 8484122]
3. Cunningham JM, Christensen ER, Tester DJ, Kim CY, Roche PC, Burgart LJ, et al. Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. *Cancer Res*. 1998; 58(15):3455–3460. [PubMed: 9699680]
4. Poynter JN, Siegmund KD, Weisenberger DJ, Long TI, Thibodeau SN, Lindor N, et al. Molecular characterization of MSI-H colorectal cancer by MLH1 promoter methylation, immunohistochemistry, and mismatch repair germline mutation screening. *Cancer Epidemiol Biomarkers Prev*. 2008; 17(11):3208–3215. [PubMed: 18990764]
5. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, et al. Cancer statistics, 2008. *CA Cancer J Clin*. 2008; 58(2):71–96. [PubMed: 18287387]
6. Sinicrope FA, Rego RL, Halling KC, Foster N, Sargent DJ, La Plant B, et al. Prognostic impact of microsatellite instability and DNA ploidy in human colon carcinoma patients. *Gastroenterology*. 2006; 131(3):729–737. [PubMed: 16952542]

7. Gafa R, Maestri I, Matteuzzi M, Santini A, Ferretti S, Cavazzini L, et al. Sporadic colorectal adenocarcinomas with high-frequency microsatellite instability. *Cancer*. 2000; 89(10):2025–2037. [PubMed: 11066042]
8. Halling KC, French AJ, McDonnell SK, Burgart LJ, Schaid DJ, Peterson BJ, et al. Microsatellite instability and 8p allelic imbalance in stage B2 and C colorectal cancers. *J Natl Cancer Inst*. 1999; 91(15):1295–1303. [PubMed: 10433618]
9. Lanza G, Gafa R, Santini A, Maestri I, Guerzoni L, Cavazzini L. Immunohistochemical test for MLH1 and MSH2 expression predicts clinical outcome in stage II and III colorectal cancer patients. *J Clin Oncol*. 2006; 24(15):2359–2367. [PubMed: 16710035]
10. Samowitz WS, Curtin K, Ma KN, Schaffer D, Coleman LW, Leppert M, et al. Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. *Cancer Epidemiol Biomarkers Prev*. 2001; 10(9):917–923. [PubMed: 11535541]
11. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol*. 2005; 23(3):609–618. [PubMed: 15659508]
12. Dolcetti R, Viel A, Doglioni C, Russo A, Guidoboni M, Capozzi E, et al. High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *Am J Pathol*. 1999; 154(6):1805–1813. [PubMed: 10362805]
13. Sinicrope FA, Rego RL, Garrity-Park MM, Foster NR, Sargent DJ, Goldberg RM, et al. Alterations in cell proliferation and apoptosis in colon cancers with microsatellite instability. *Int J Cancer*. 2007; 120(6):1232–1238. [PubMed: 17187355]
14. Ribic CM, Sargent DJ, Moore MJ, Thibodeau SN, French AJ, Goldberg RM, et al. Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med*. 2003; 349(3):247–257. [PubMed: 12867608]
15. Benatti P, Gafa R, Barana D, Marino M, Scarselli A, Pedroni M, et al. Microsatellite instability and colorectal cancer prognosis. *Clin Cancer Res*. 2005; 11(23):8332–8340. [PubMed: 16322293]
16. Jover R, Zapater P, Castells A, Llor X, Andreu M, Cubiella J, et al. Mismatch repair status in the prediction of benefit from adjuvant fluorouracil chemotherapy in colorectal cancer. *Gut*. 2006; 55(6):848–855. [PubMed: 16299036]
17. Carethers JM, Smith EJ, Behling CA, Nguyen L, Tajima A, Doctolero RT, et al. Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer. *Gastroenterology*. 2004; 126(2):394–401. [PubMed: 14762775]
18. Jover R, Zapater P, Castells A, Llor X, Andreu M, Cubiella J, et al. The efficacy of adjuvant chemotherapy with 5-fluorouracil in colorectal cancer depends on the mismatch repair status. *Eur J Cancer*. 2009; 45(3):365–373. [PubMed: 18722765]
19. Sargent DJ, Marsoni S, Thibodeau SN, et al. Confirmation of deficient mismatch repair (dMMR) as a predictive marker for lack of benefit from 5-FU based chemotherapy in stage II and III colon cancer (CC): A pooled molecular reanalysis of randomized chemotherapy trials. 2008 ASCO Annual Meeting. 2008 May 20.26 suppl abstr 4008.
20. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, et al. A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. *Cancer Res*. 1998; 58(22):5248–5257. [PubMed: 9823339]
21. Lindor NM, Burgart LJ, Leontovich O, Goldberg RM, Cunningham JM, Sargent DJ, et al. Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors. *J Clin Oncol*. 2002; 20(4):1043–1048. [PubMed: 11844828]
22. Alexander J, Watanabe T, Wu TT, Rashid A, Li S, Hamilton SR. Histopathological identification of colon cancer with microsatellite instability. *Am J Pathol*. 2001; 158(2):527–535. [PubMed: 11159189]
23. Smyrk TC, Watson P, Kaul K, Lynch HT. Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma. *Cancer*. 2001; 91(12):2417–2422. [PubMed: 11413533]



24. Kim H, Jen J, Vogelstein B, Hamilton SR. Clinical and pathological characteristics of sporadic colorectal carcinomas with DNA replication errors in microsatellite sequences. *Am J Pathol.* 1994; 145(1):148–156. [PubMed: 8030745]
25. Thibodeau SN, French AJ, Cunningham JM, Tester D, Burgart LJ, Roche PC, et al. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. *Cancer Res.* 1998; 58(8):1713–1718. [PubMed: 9563488]
26. Slattery ML, Curtin K, Anderson K, Ma KN, Ballard L, Edwards S, et al. Associations between cigarette smoking, lifestyle factors, and microsatellite instability in colon tumors. *J Natl Cancer Inst.* 2000; 92(22):1831–1836. [PubMed: 11078760]
27. Lynch HT, Boland CR, Rodriguez-Bigas MA, Amos C, Lynch JF, Lynch PM. Who should be sent for genetic testing in hereditary colorectal cancer syndromes? *J Clin Oncol.* 2007; 25(23):3534–3542. [PubMed: 17687158]
28. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Ruschoff J, et al. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst.* 2004; 96(4):261–268. [PubMed: 14970275]
29. Hampel H, Frankel WL, Martin E, Arnold M, Khanduja K, Kuebler P, et al. Feasibility of screening for Lynch syndrome among patients with colorectal cancer. *J Clin Oncol.* 2008; 26(35):5783–5788. [PubMed: 18809606]
30. Rodriguez-Moranta F, Castells A, Andreu M, Pinol V, Castellvi-Bel S, Alenda C, et al. Clinical performance of original and revised Bethesda guidelines for the identification of MSH2/MLH1 gene carriers in patients with newly diagnosed colorectal cancer: proposal of a new and simpler set of recommendations. *Am J Gastroenterol.* 2006; 101(5):1104–1111. [PubMed: 16696788]
31. Jenkins MA, Hayashi S, O'Shea AM, Burgart LJ, Smyrk TC, Shimizu D, et al. Pathology features in Bethesda guidelines predict colorectal cancer microsatellite instability: a population-based study. *Gastroenterology.* 2007; 133(1):48–56. [PubMed: 17631130]
32. Allegra C. Thymidylate synthase levels: prognostic, predictive, or both? *J Clin Oncol.* 2002; 20(7):1711–1713. [PubMed: 11919223]
33. O'Connell MJ, Sargent DJ, Windschitl HE, Shepherd L, Mahoney MR, Krook JE, 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(2):133–139. [PubMed: 16945169]
34. Sinicrope FA, Rego RL, Halling KC, Foster NR, Sargent DJ, La Plant B, et al. Thymidylate synthase expression in colon carcinomas with microsatellite instability. *Clin Cancer Res.* 2006; 12(9):2738–2744. [PubMed: 16675565]
35. Laiho P, Launonen V, Lahermo P, Esteller M, Guo M, Herman JG, et al. Low-level microsatellite instability in most colorectal carcinomas. *Cancer Res.* 2002; 62(4):1166–1170. [PubMed: 11861399]
36. Halford S, Sasieni P, Rowan A, Wasan H, Bodmer W, Talbot I, et al. Low-level microsatellite instability occurs in most colorectal cancers and is a nonrandomly distributed quantitative trait. *Cancer Res.* 2002; 62(1):53–57. [PubMed: 11782358]
37. Parc YR, Halling KC, Wang L, Christensen ER, Cunningham JM, French AJ, et al. HSMH6 alterations in patients with microsatellite instability-low colorectal cancer. *Cancer Res.* 2000; 60(8):2225–2231. [PubMed: 10786688]
38. Compton C, Fenoglio-Preiser CM, Pettigrew N, Fielding LP. American Joint Committee on Cancer Prognostic Factors Consensus Conference: Colorectal Working Group. *Cancer.* 2000; 88(7):1739–1757. [PubMed: 10738234]
39. Hosmer, D. *Applied Logistic Regression.* ed. 2nd.. New York: John Wiley & Sons; 2000.
40. Jass JR, Do KA, Simms LA, Iino H, Wynter C, Pillay SP, et al. Morphology of sporadic colorectal cancer with DNA replication errors. *Gut.* 1998; 42(5):673–679. [PubMed: 9659163]
41. Young J, Simms LA, Biden KG, Wynter C, Whitehall V, Karamatic R, et al. Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings: parallel pathways of tumorigenesis. *Am J Pathol.* 2001; 159(6):2107–2116. [PubMed: 11733361]
42. Aaltonen LA, Peltomaki P. Genes involved in hereditary nonpolyposis colorectal carcinoma. *Anticancer Res.* 1994; 14(4B):1657–1660. [PubMed: 7979203]

43. Kim YH, Kakar S, Cun L, Deng G, Kim YS. Distinct CpG island methylation profiles and BRAF mutation status in serrated and adenomatous colorectal polyps. *Int J Cancer*. 2008; 123(11):2587–2593. [PubMed: 18798261]
44. Halvarsson B, Anderson H, Domanska K, Lindmark G, Nilbert M. Clinicopathologic factors identify sporadic mismatch repair-defective colon cancers. *Am J Clin Pathol*. 2008; 129(2):238–244. [PubMed: 18208804]

**Table 1**

Clinicopathological Variables Stratified by DNA Mismatch Repair (MMR) Status in Human Colon Carcinomas

Variable	Intact MMR (N=835)	Defective MMR (N=147)	P-value <sup>1</sup>
<b>Histologic grade<sup>2a</sup></b>			<0.0001
Well/moderate	637 (76.5%)	70 (48.3%)	
Poor/undifferentiated	196 (23.5%)	75 (51.7%)	
<b>Gender</b>			0.0008
Female	380 (45.5%)	89 (60.5%)	
Male	455 (54.5%)	58 (39.5%)	
<b>Tumor site<sup>2b</sup></b>			<0.0001
Distal	455 (54.7%)	13 (9%)	
Proximal	377 (45.3%)	132 (91%)	
<b>TNM stage</b>			0.0027
III	690 (82.6%)	106 (72.1%)	
II	145 (17.4%)	41 (27.9%)	
<b>Age</b>			0.0107 <sup>3</sup>
Mean (SD)	62.1 (10.50)	64.1 (11.11)	
Median	63.0	66.5	
Range	(25.0–88.0)	(31.1–86.2)	
<b>Lymph node metastases<sup>3c</sup></b>			0.0081
0	145 (18.3%)	41 (29.1%)	
1–3	429 (54%)	71 (50.4%)	
>3	220 (27.7%)	29 (20.6%)	
<b>Age group</b>			0.0270
< 50	104 (83.9%)	20 (16.1%)	
50 – 59	196 (90.3%)	21 (9.7%)	
60 – 64	147 (85.5%)	25 (14.5%)	
65 – 69	179 (86.5%)	28 (13.5%)	
70+	209 (79.8%)	53 (20.2%)	
<b>T-stage<sup>2d</sup></b>			0.1053
T <sub>1,2</sub>	108 (13.1%)	12 (8.3%)	
T <sub>3,4</sub>	718 (86.9%)	133 (91.7%)	

<sup>1</sup> Chi-square p-value<sup>2</sup> Missing data: n= 4<sup>a</sup>, n= 5<sup>b</sup>, n= 47<sup>c</sup>, n= 11<sup>d</sup><sup>3</sup> Wilcoxon rank-sum p-value

**Table 2**

Univariate Association of Covariates with Defective versus Intact MMR Status (N= 982)

Variable or Model	Odds Ratio (95% CI) <sup>1</sup>	c-statistic <sup>2</sup>
Tumor Site (proximal vs. distal)	12.26 (6.82, 22.07)	0.764 (0.729)
Histologic Grade (poor/undifferentiated vs. well/moderate)	3.84 (2.65, 5.59)	0.682 (0.641)
Sex (male vs. female)	0.54 (0.38, 0.77)	0.618 (0.575)
Age ≥ 70	1.64 (1.13, 2.39)	0.602 (0.555)
LN metastasis (1–3 vs. 0) (> 3 vs. 0)	0.76 (0.46, 1.26) 0.61 (0.34, 1.11)	0.593 (0.568)
T-stage (3, 4 vs. 1, 2)	1.58 (0.84, 2.95)	0.585 (0.524)
Stage (II vs. III)	1.51 (0.96, 2.39)	0.582 (0.553)

<sup>1</sup> logistic models adjusted for adjuvant study; c-statistic was unadjusted.

<sup>2</sup> adjusted for adjuvant study (unadjusted in parenthesis).

**Table 3**

Positive Predictive Value of Covariate Combinations for MMR Status

Covariate Category <sup>1</sup>	Intact MMR	Defective MMR (PPV)	Risk of defective MMR vs, expected rate of 15%
Distal	455 (97%)	13 (3%)	Low Risk
Proximal, Well/moderate <sup>2</sup> , Male	151 (85%)	26 (15%)	Average Risk
Proximal, Well/moderate <sup>2</sup> , Female	123 (78%)	35 (22%)	Slightly Elevated Risk
Proximal, Poor/undifferentiated <sup>2</sup> , Male	57 (71%)	23 (29%)	Moderately Elevated Risk
Proximal, Poor/undifferentiated <sup>2</sup> , Female	45 (49%)	46 (51%)	High Risk

<sup>1</sup> c-statistic of 0.810, adjusted for adjuvant study

<sup>2</sup> differentiation of tumor



**Table 4**

Univariate Association of Covariates with Defective versus Intact MMR Status in Colon Cancers (N= 326)

Variable or Model	Odds Ratio (95% CI) <sup>1</sup>	c-statistic <sup>2</sup>
TILs ( $\geq 2$ /HPF) <sup>3</sup>	14.88 (7.56, 29.27)	0.791 (0.753)
Tumor Site (proximal vs. distal)	15.80 (5.56, 44.90)	0.755 (0.734)
Histologic Grade (poor/undifferentiated vs. well/moderate)	4.02 (2.21, 7.32)	0.680 (0.656)
Sex (female vs. male)	2.53 (1.39, 4.62)	0.632 (0.610)
Age $\geq 70$	2.21 (1.23, 3.98)	0.607 (0.589)
LN metastasis		0.566 (0.541)
(1–3 vs. 0)	0.74 (0.37, 1.49)	
(> 3 vs. 0)	0.70 (0.31, 1.56)	
Stage (II vs. III)	1.41 (0.74, 2.68)	0.549 (0.538)
T-stage ( <sub>3,4</sub> vs. <sub>1,2</sub> )	1.98 (0.58, 6.79)	0.543 (0.523)

<sup>1</sup> logistic models adjusted for adjuvant study<sup>2</sup> adjusted for adjuvant study (unadjusted in parenthesis).<sup>3</sup> TILs: tumor infiltrating lymphocytes

**Table 5**

## Covariate Models and MMR Status in Colon Cancers

Covariate Category <sup>1</sup>	Intact MMR	Defective MMR (PPV)	NPV	Risk of defective MMR <sup>2</sup>
Distal	144 (97%)	4 (3%)	70%	Low Risk
Proximal, TILs 1+	109 (83%)	22 (17%)	82%	Average Risk
Proximal, TILs 2-4+, Male	10 (50%)	10 (50%)	84%	High Risk
Proximal, TILs 2-4+, Female	5 (19%)	22 (81%)	88%	Very High Risk

<sup>1</sup> c-statistic of 0.86 adjusted for adjuvant study. Observed rates shown above are the same as the PPV values (unadjusted for study).

<sup>2</sup> Compared to expected MMR rate of 15%