Development and Validation of a Colon Cancer Risk Assessment Tool for Patients Undergoing Colonoscopy

Fay Kastrinos, M.D., M.P.H.[1,5], **John I. Allen, MD, MBA**[6], **David H. Stockwell, M.D., M.P.H.**[1,2,5], **Elena M. Stoffel, M.D., M.P.H.**[1,3,5], **Earl F. Cook, ScD**[4,5], **Muthoka L. Mutinga, M.D.**[2,5], **Judith Balmaña, M.D.**[7], and **Sapna Syngal, M.D., M.P.H.**[1,3,5]

[1]Division of Gastroenterology, Brigham and Women's Hospital, Boston, MA

[2]Division of Gastroenterology, Faulkner Hospital, Boston, MA

[3]Population Sciences Division, Dana-Farber Cancer Institute, Boston, MA

[4]Division of Epidemiology, Harvard School of Public Health, Boston, MA

[5]Harvard Medical School, Boston, MA

[6]Minnesota Gastroenterology, Minneapolis, MN

[7]Medical Oncology Department, Hospital Vall d'Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain

Abstract

BACKGROUND & AIM—Diagnostic criteria for hereditary colorectal cancer (CRC) are complex. "Open-access" colonoscopy makes it challenging to identify who needs genetic evaluation, intensive surveillance, and screening for extracolonic tumors. Our aim was to develop a simple, pre-procedural risk assessment tool to identify who may be at highest risk for CRC.

METHODS—631 outpatients undergoing colonoscopy at two academic practices completed a questionnaire assessing personal and family history of CRC, polyps, and Lynch Syndrome (LS)-associated malignancies. Subjects were considered high-risk if one of nine prespecified characteristics of hereditary CRC syndromes were met. Through recursive partitioning analysis, an algorithm of fewest questions needed to capture the most high-risk individuals was developed. Results were validated in 5335 individuals undergoing colonoscopy at five private endoscopy centers and tested in 285 carriers of mismatch repair (MMR) mutations associated with LS.

RESULTS—17.7% and 20.0% of individuals were classified as high-risk in the development and validation cohorts, respectively. Recursive partitioning revealed three questions most informative for identifying high-risk patients:1."Do you have a first-degree relative (FDR) with CRC or LS-related cancer diagnosed before age 50?" 2."Have you had CRC or polyps diagnosed before age

---

Correspondence to: Sapna Syngal, MD, MPH Dana-Farber Cancer Institute 1 Jimmy Fund Way, SM 20 Boston, MA 02115 ssyngal@partners.org Phone: 617-632-5022, Fax: 617-632-4088.

50?" 3."Do you have ≥ 3 relatives with CRC?" When asked successively, these questions identified 77% of high-risk individuals in both cohorts and 271/285 (95%) of mutation carriers.

CONCLUSIONS—Approximately one in five individuals undergoing colonoscopy would benefit from further risk assessment. We developed a simple, three-question CRC Risk Assessment Tool to identify the majority of patients who require additional assessment and possible genetic evaluation.

INTRODUCTION

Appropriate risk stratification improves the overall efficiency and effectiveness of endoscopic screening and surveillance for colorectal cancer (CRC).[1] Use of colonoscopy for CRC screening has increased in the United States[2,3] and commonly targets individuals at average risk due to age, or those at moderate risk with a history of colon polyps or a family history of colonic neoplasia. However, as many as 10% of CRC cases may be associated with an inherited syndrome[4] and intensive endoscopic surveillance along with genetic evaluation and predictive testing have been shown to reduce CRC related morbidity and mortality.[5-7] While many physicians routinely ask patients about their family history of colon cancer, identifying the subset of individuals who may be at risk for a hereditary cancer syndrome is difficult because of multiple defining criteria and the requirement of a detailed personal and family history assessment.

Clinical criteria to identify patients at risk for Lynch Syndrome (also known as hereditary nonpolyposis colorectal cancer (HNPCC)), the most common hereditary CRC syndrome, are widely accepted[8-11] but both the Amsterdam Criteria and Bethesda Guidelines are cumbersome and impractical to routinely use in clinical practice, particularly in asymptomatic screening populations where the vast majority of individuals are not at risk for the disease. Recent studies have shown that gastroenterologists have difficulty recognizing Lynch syndrome and recommend appropriate screening at very low rates.[12,13] Patients affected with Familial Adenomatous Polyposis (FAP) classically develop hundreds to thousands of colorectal adenomas at a young age and can be easily identified, but a subset of patients have a less obvious phenotype. Patients with 10 or more (but less than 100) cumulative colorectal adenomas, may have a variant of FAP called attenuated FAP (AFAP). In addition, a recessive form of adenomatous polyposis has been described which is due to mutations in the *MYH* gene ("*MYH* associated polyposis" or MAP). While the true incidence of atypical forms of polyposis is unknown, gene mutations in both *APC* and *MYH* are associated with a high CRC risk and should be considered in individuals with multiple adenomas. Under-recognition of clinical criteria used to identify individuals with Lynch syndrome, AFAP, and other polyposis syndromes leads to incomplete risk assessment with many potential repercussions for patients who may be at the highest risk of developing cancer.

National guidelines recommend genetic evaluation in individuals at high risk for CRC based on personal or family medical history.[14] Identifying gene carriers through predictive genetic testing improves the efficiency of cancer surveillance and helps identify which family members may require intense endoscopic evaluation versus those who can undergo screening at average risk intervals.[15,16] However, referral rates for genetic evaluation are low[17] despite the estimated 20% of all CRC patients diagnosed each year who meet revised Bethesda criteria.[18-20]

Time constraints placed on today's busy healthcare providers increase the difficulty of identifying patients at risk for hereditary CRC syndromes. In the setting of "open-access" colonoscopy, gastroenterologists meet patients just minutes before performing colonoscopy with little time to obtain the complete personal and family history needed to properly assess

CRC risk. Incomplete risk assessment inevitably leads to inappropriate recommendations for surveillance colonoscopy and a missed opportunity for genetic evaluation referral. The goal of our study was to develop a simple, practical and efficient tool to improve pre-procedural CRC risk assessment by identifying those individuals who may be at high risk for hereditary CRC syndromes.

METHODS

The study proceeded in five phases: (1) pilot testing of the data collection instrument; (2) data collection phase, where patients undergoing lower endoscopy completed a questionnaire and were risk-stratified based on personal or family history characteristics; individuals were classified as being "high-risk" if they fulfilled one of 9 pre-defined criteria, (3) development of the CRC Risk Assessment Tool, where recursive partitioning analysis was used to derive an algorithm that would identify the most high-risk individuals with the fewest number of questions, (4) validation testing of the CRC Risk Assessment Tool in an external population, and (5) determining the Tool's ability to identify known carriers of mismatch repair (MMR) gene mutations associated with Lynch Syndrome in a second external validation cohort. This study received approval from the Dana-Farber Cancer Institute's (DFCI) Institutional Review Board (IRB) and the Partners Human Research Committee.

From February 2004 to June 2004, data were collected from consecutively referred patients for lower endoscopy, including colonoscopy and flexible sigmoidoscopy. Subjects were recruited from Brigham and Women's Hospital, a tertiary-care institution, and Faulkner Hospital, a community-affiliated hospital, both in Boston, MA.

Eligible subjects included those who were scheduled for colonoscopy or flexible sigmoidoscopy performed for screening, surveillance or diagnostic purposes. All subjects were age 18 years and older and were able to give informed consent. Exclusion criteria included the inability to read and write in English, since all subjects were required to complete a self-administered written questionnaire.

Potential participants were identified from the endoscopy center schedules and were invited to participate by mail prior to their procedure. Because this type of mailing required identification of potential subjects in advance, a waiver of consent from the IRB was obtained for the initial part of recruitment. Potential study subjects received a cover letter describing the study, a study questionnaire and consent form. Those patients interested in participating completed the questionnaire prior to their procedure date and returned the completed study materials on the day of their exam.

Phases One and Two: Instrument Development and Data Collection

All recruited subjects completed the "CRC Risk Assessment Questionnaire," a 34-item self-administered survey developed to assess CRC risk based on personal and family history of CRC and polyps, as well as Lynch syndrome-related malignancies. The questionnaire also elicited information about subjects' prior history of lower endoscopies, including number of procedures and the indication for the procedures. Standard demographic information such as age, gender, race/ethnicity, marital status, income, level of education and type of health insurance was also obtained.

This study instrument was developed in collaboration with the hereditary colon cancer group at DFCI and within the Division of Gastroenterology at Brigham and Women's Hospital. The CRC Risk Assessment Questionnaire was pilot tested to determine if participants found any questions difficult to understand. Seventy-one of 284 subjects were part of the pilot

testing and were recruited from Brigham and Women's Hospital. Data from subjects enrolled in the pilot study were not included in the final analysis and in generating the CRC Risk Assessment Tool.

Our goal was to identify a subset of individuals who may benefit from further risk assessment for hereditary CRC. We developed a category of "high-risk" that was defined by criteria adapted from the modified and revised Bethesda guidelines[8-11] as well as the American Gastroenterological Association's (AGA) review on hereditary colorectal cancer and genetic testing.[14] Subjects reporting any of the following criteria were classified as high-risk: Personal history of CRC diagnosed at age < 50 years, polyps at age < 50 years, endometrial cancer at age < 50 years, any Lynch syndrome-associated cancer diagnosed at age <50 years, 10 cumulative polyps; Family history of first-degree relative (FDR) with CRC at age <50 years, FDR with polyps at age < 50 years, 3 relatives with CRC, FDR with Lynch syndrome-associated cancer at age <50 years. Lynch syndrome-associated cancers include endometrial, ovarian, urinary tract (including bladder, ureter, kidney), gastric, small bowel, biliary, pancreatic or brain cancers.

The proportion of individuals fulfilling any high-risk criteria was calculated. To verify the accuracy of self-reported personal and family histories, those subjects who were classified as high-risk and provided consent for future participation in the study, were contacted by telephone by the study's research coordinator. The information provided on the "CRC Risk Assessment Questionnaire" was reviewed for each individual high-risk subject in order to assess the accuracy of the histories provided. Simple descriptive statistics were used for the analysis of patient demographics and the proportion of subjects fulfilling each high-risk criterion.

Phase Three: Derivation of the CRC Risk Assessment Tool

We next analyzed data from the subject-completed questionnaires to construct a risk assessment tool that would identify the greatest number of high-risk individuals with the fewest possible number of questions. Through recursive partitioning analysis (using Classification And Regression Tree (CART) software),[21] a simple decision-making tool to identify high-risk individuals was derived. In this analysis, subgroups of patients fulfilling high-risk criteria were identified in a sequential manner based on component criteria used to define high-risk. The initial high-risk subgroup was based on one of the nine clinical factors deemed to be the single, most informative high-risk criterion. The presence of this first criterion defined the subgroup with the largest percentage of high-risk subjects. Once this criterion was identified, the remaining subject sample was divided again using the next best criterion. This process was repeated using subsequently chosen criterion until no further partitioning was useful.

Through recursive partitioning analysis we developed an algorithm of the fewest number of questions that would identify the largest proportion of individuals fulfilling high-risk criteria. We analyzed the ability of each potential criterion to identify the greatest proportion of high-risk individuals and determined which criteria would provide the maximum sensitivity in correctly stratifying subjects into the high-risk group.

Phase Four: Validation Testing of the CRC Risk Assessment Tool

The goal of the validation phase was to prospectively evaluate the proportion of high-risk individuals and performance of CRC Risk Assessment Tool in a second cohort from a different geographic location as well as practice setting. The validation cohort was comprised of consecutive individuals presenting for colonoscopy to any one of five ambulatory centers affiliated with Minnesota Gastroenterology, a large, privately-held

practice. The nine questions used to identify high-risk individuals were incorporated into each patient's pre-procedural risk assessment evaluation (and electronic medical record (EMR). The necessary study data was obtained in a two-step manner. Patients were first asked to complete a printed "History Form" which is routinely sent by mail to all individuals scheduled for an endosocpic procedure at Minnesota Gastroenterology. The CRC Risk Assessment Tool questions were incorporated into the History Form. When patients presented for their scheduled colonoscopy, their responses to the History Form/CRC Risk Assessment Tool were then confirmed by either a nurse or by an admission specialist during the admission (triage) process. No further attempts were made to verify the histories reported by eligible participants in the validation cohort. Information on the frequency of incorrect histories provided by study participants and elicited during the triage process was not recorded.

Data on patient age, gender, and responses to the nine questions was prospectively collected from May to September 2007 in order to validate the Risk Assessment Tool. The proportion of individuals fulfilling high-risk criteria was calculated. Similar to the methods in the first phase of the study, simple descriptive statistics were used to analyze patient demographics and the proportion of subjects fulfilling each high-risk criterion. The sensitivity of each sequential question of the CRC Risk Assessment Tool was determined and compared to the findings from the derivation cohort.

Phase Five: Testing the Ability of the CRC Risk Assessment Tool to Identify Individuals with Lynch Syndrome and Known MMR Gene Mutations

Subjects with a known mutation in the MMR genes most commonly associated with Lynch Syndrome, *MLH1* and *MSH2*, were identified from an existing anonymized database of 1914 unrelated probands who had undergone full gene sequencing and analysis of large rearrangements of these two genes by a commercial laboratory. This data has been previously used to develop a prediction model to identify gene mutation carriers.[22] The database contains information on the probands' demographic profile (including age, gender, ethnicity) and personal and family history of neoplasia (cancer and colonic adenomatous polyps) along with their ages of diagnoses. This information was originally provided on a test order form that was completed by the healthcare professional ordering the genetic testing. For phase five of this study, the personal and family histories of the identified mutation carriers were reviewed. The proportion of gene mutation carriers fulfilling each sequential criterion of the CRC Risk Assessment Tool was calculated in order to determine the instrument's overall ability to accurately identify these individuals with Lynch syndrome.

RESULTS

A total of 1,641 patients met entry criteria during the recruitment period from February 2004 to September 2004 and were offered enrollment in the study. Of these eligible subjects, 631 (38.5%) agreed to participate and returned completed questionnaires. A total of 16 patients were excluded: 11 had a known diagnosis of FAP or Lynch syndrome and five provided incomplete questionnaires. Overall, data from 615 subjects was analyzed. A total of 327 patients enrolled at Brigham and Women's Hospital and 288 enrolled at Faulkner Hospital, and patient demographics were comparable between the two groups. Table 1 depicts subject demographics for the entire cohort and those subjects that fulfilled high-risk criteria.

Of the 615 participants, 109 subjects (17.7%) met high-risk criteria. Detailed data regarding criteria by which they were defined as high-risk are presented in Table 2. 84/109 (77%) fulfilled at least one criterion related to family history and 33/109 (30.3%) fulfilled at least one criterion related to their personal history. There were eight patients who were defined as

being high-risk based on both personal and family history. 74% (81/109) of high-risk participants had been scheduled for their lower endoscopy via "open-access" and had not seen a gastroenterologist prior to their procedure.

Attempts were made to contact fifty-five subjects who were identified as high-risk had given consent to be contacted by telephone for future participation in this study in order to verify the accuracy of the personal and family histories provided. Twenty-three subjects were successfully contacted, personal and family histories were obtained and compared to the information these patients had provided on the CRC Risk Assessment Questionnaire. Eighty-seven percent (20/23) of these high-risk subjects reported the same personal and family history information.

Using recursive partitioning analysis, the nine defining criteria for high-risk were ranked from most informative to least informative in the form of a simple flowchart (Figure 1). The most informative question was that which identified the largest proportion of high-risk subjects. Each subsequent question was ranked based on its ability to capture the most high-risk individuals remaining and yield the highest possible cumulative sensitivity when the questions were asked sequentially. The recursive partitioning technique initially segregated patients with history of a "FDR with a Lynch syndrome-associated cancer at age < 50 years" into the first high-risk subgroup containing 31 of the 109 high-risk subjects (sensitivity=28.4%), thereby making it the most informative question to ask when determining a patient's risk for a hereditary CRC syndrome. Among the remaining 584 subjects, "having more than three relatives with CRC" was the next selected criterion, identifying another 25 of the high-risk subjects (cumulative sensitivity = 51.4%).

Certain criteria were then combined *a priori* to create single, two-part questions. For example, the criteria "FDR with CRC <50 years" and "FDR with a Lynch syndrome-associated cancer at age < 50 years" were merged to create a single question, "Do you have a FDR with a Lynch syndrome-related cancer or CRC diagnosed before age 50?", resulting in a sensitivity of 43.1% (Figure 2). This was done to create questions that would be easier to ask and appear less redundant to patients. As a result, a total of six questions were used in the final analysis and the abbreviated algorithm is depicted in Figure 2. The abbreviated algorithm was able to capture a higher percentage of high-risk individuals with each question when compared to the original algorithm employing all nine individual questions. The CRC Risk Assessment Tool was then extrapolated from the abbreviated algorithm and is comprised of the first three questions of the algorithm since these questions were deemed the most informative when asked in the given order. The decision to exclude the fourth question from the CRC Risk Assessment Tool ("Do you have a FDR with colorectal polyps diagnosed before age 50?") was based on the assumption that the information provided on polyp history in family members may not be reliable. When asked in succession, the three-question Risk Assessment Tool is able to capture 77% (sensitivity) of high-risk individuals. Figure 3 depicts the three-question CRC Risk Assessment Tool as it would appear in the pre-procedural evaluation of patients undergoing colonoscopy.

Data from 5,335 subjects was collected during the prospective validation phase. 1,069 (20%) individuals met high-risk criteria. The mean age and gender of subjects in the derivation and validation cohorts was similar (Table 1). The percentages of individuals fulfilling each high-risk criterion in the derivation and validation cohort are shown in Table 2. The sensitivity of each criterion in the specified sequence was similar in both cohorts (Figure 2). The three-question Risk Assessment Tool's ability to capture 77% of individuals meeting high-risk criteria was reproduced in the validation cohort.

In a second external validation cohort, 285 individuals with Lynch Syndrome (as defined by clinical genetic testing and the presence of a mutation in the MMR genes *MLH1* or *MSH2)*, were identified. The percentages of gene mutation carriers fulfilling each of the questions comprising the CRC Risk Assessment Tool are shown in Figure 4. When asked in succession, the three-question Tool was able to accurately identify 95% of gene mutation carriers. There was little incremental value gained with proceeding to the fourth question of the validated Risk Assessment Tool which captured less than one percent of additional mutation carriers. Data pertaining to the total number of polyps in mutation carriers was not available and the cumulative sensitivity of the CRC Risk Assessment Tool could not be determined for this and subsequent criterion. Overall, only 5% of individuals with Lynch Syndrome would have been missed by the three-question CRC Risk Assessment Tool.

DISCUSSION

We have developed a three-question CRC Risk Assessment Tool that will aid health care providers in identifying individuals who need further assessment for hereditary CRC and can be incorporated into the routine pre-procedural evaluation for patients undergoing colonoscopy. The simple algorithm is comprised of the highest yield questions related to personal and family history associated with an inherited form of CRC. Our goal was not to identify new risk factors or contest the importance of existing clinical criteria for Lynch Syndrome or inherited polyposis syndromes but to derive a quick, practical approach to improve CRC risk assessment in identifying high-risk individuals.

To date, a limited number of scoring systems have been developed for CRC risk assessment. Selvachandran et al. devised a four-page pre-procedural questionnaire that attempted to stratify patients at risk of having CRC or other significant pathology based on distal colonic symptoms.[23] A weighted numerical score was subjectively derived by the investigators and relied solely on patient symptomatology. The impact of family history on risk stratification was not incorporated into the questionnaire or scoring system. Church et al. took family history into account and designed a multivariable scoring system based solely on the presence of CRC in first- and second-degree relatives in order to make risk more readily identifiable.[24] A disadvantage of such scoring systems is their complexity and the need of an additional nomogram reference to convert a point score to appreciable CRC risk.

To the best of our knowledge, this study is the first study to generate a simple non-numerical algorithm for CRC risk assessment. The main advantage of using the recursive partitioning approach is that it generates algorithms that are easy to apply in a variety of different clinical settings and can even be self-administered. This method has frequently been used to predict outcomes in other medical conditions, including neurological, oncological, as well as cardiac disorders, including acute coronary syndromes and congestive heart failure.[25-29] In our analysis, recursive partitioning was used in a novel manner to create a hierarchy of importance among multiple criteria that are used to define the clinical outcome of high CRC risk.

Our study has several important limitations. For the identification of high-risk individuals we relied on the patients' self-report of their personal and family history of colorectal cancer and polyps, as well as extracolonic cancers, as the gold standard. Although prior studies have shown that histories reported by patients are accurate[30] and correct in up to 80-89% for colon cancer in FDRs,[31-33] the provided information may be inaccurate, particularly polyp history of family members. Nevertheless, physicians commonly accept a patient's report of personal and family history and in practice rarely have means to verify these medical histories, especially in an open-access endoscopy setting. The optimal use of the CRC Risk Assessment Tool is as a first step in determining if an individual may be at high risk for a

hereditary CRC syndrome. If the Risk Assessment Tool "flags" such a patient, a detailed personal and multigenerational family history (pedigree) of neoplasia should be elicited and formal genetic assessment should be considered. Such an instrument should aid in medical decision-making regarding patients suspected of being at high-risk for CRC and should enhance, not replace, physician assessment.

A second potential limitation of this study is related to response bias and the volunteer effect. This is often a limitation of survey studies where it is difficult to tell whether or not the responses are typical of those that would be received of all persons who are eligible to participate. One may argue that the cohort from which the CRC Risk Assessment Tool was developed may have been skewed to include individuals that may have been more knowledgeable of their personal and family medical histories. However, the high percent of patients fulfilling high-risk criteria in the validation cohort does not support this claim since the nine questions used to define high-risk were a mandatory component of each patient's electronic medical record and were completed before the procedure. Overall, the results from the validation phase lend credence to the generalizability of the instrument as its performance was reproducible with sustained discriminatory ability. While subjects in the derivation cohort were of a higher socioeconomic status (36.5% had household incomes greater than $100,000 and 65.8% had college or postgraduate degrees), and were mostly Caucasian, further testing would be needed to determine the tool's performance in a more diverse patient population.

A third limitation of this study may be related to the criteria selected to define high-risk. The original nine questions defining high-risk were adapted from the modified and revised Bethesda guidelines[8-11] as well as the AGA review on hereditary CRC and genetic testing.[14] We deliberately chose to include less restrictive criteria that have been associated with Lynch Syndrome to minimize the likelihood of missing high-risk subjects. For example, a personal history of an adenoma diagnosed at a young age is considered a criterion that when used alone, has little capacity to identify an individual with a MMR gene. However, there is little information about the natural history of these young-onset polyps and there are currently no specific recommendations to guide the clinical follow-up of such patients. In our opinion, it is reasonable for a young patient diagnosed with one or more adenomas to undergo additional clinical evaluation such as family history assessment, as well as review of the polyps removed to assess size, number and presence of histologically advanced lesions. The recognition of new polyposis phenotypes highlights the need for future study of individuals with young-onset adenomas and until additional information is made available, the current evaluation remains complex and an individualized process.

Finally, we do not have data on which of the patients identified by the CRC Risk Assessment Tool were found to ultimately have a hereditary cancer syndrome, as defined by mutation analysis, which was not our outcome of interest. This tool was not developed to predict which patients may be mutation carriers of genes commonly associated with CRC. A number of sophisticated prediction models have been specifically developed for such a purpose[22, 34-37] and may be used in conjunction with the CRC Risk Assessment Tool to streamline the follow-up evaluation of those individuals defined as high-risk during the time of colonoscopy.

Overall, the individuals who are identified by the CRC Risk Assessment Tool as high-risk should have one of two evaluations following colonoscopy. Those patients with a personal history of CRC at age less than 50 years or a FDR with young-onset CRC or a Lynch syndrome-associated cancer should be referred for a genetic evaluation and possible predictive genetic testing or molecular tumor analysis. Patients with a personal history of young-onset adenomas or more than three relatives with CRC should undergo further

clinical assessment to ascertain the histology, size, location, and number of polyps and to obtain a multigenerational family history of neoplasia. In all cases identified by the CRC Risk Assessment Tool, the percent likelihood of detecting a MMR mutation can be estimated by the user-friendly Web-based prediction model, $PREMM_{1,2}$.[22] A $PREMM_{1,2}$ score of <5% has been considered to be a reliable cut-off to exclude those patients who do not need further risk assessment because of the model's 100% negative predictive value for detecting germline *MLH1* or *MSH2* gene mutations.[38]

Our findings have implications regarding the quality of care provided to patients presenting for open-access colonoscopy. While much attention has been placed on the technical performance of colonoscopy for optimizing CRC screening, complete risk assessment and subsequent surveillance recommendations for patients and family members may be compromised in this type of setting. A limited encounter procedure may miss the subset of patients who are at highest risk of developing CRC. In this study, 74% of the high-risk patients identified in the derivation cohort were scheduled for colonoscopy without a prior visit with a gastroenterologist. While endoscopists have been criticized for the over-performance of colonoscopy in average and moderate risk individuals, there are no data available regarding the inappropriate under-use of colonoscopy in possible high-risk individuals. It has been established that referrals for genetic evaluation are low, even in individuals with personal history of CRC who fulfill Bethesda criteria and have had close evaluation of personal and family history. Colonoscopy performed through an open-access system may also contribute to the missed opportunity for genetic evaluation referral. Genetic counseling and predictive genetic testing is of particular importance in the care of affected individuals and has been shown to increase life expectancy due to screening compliance or appropriate recommendations for surveillance or prophylactic surgery[39] among identified mutation carriers.[40-42] Genetic testing also provides affected individuals with information to help diagnose other family members, gain reassurance, engage in more informed health decision-making, and plan for the future.

In summary, the high morbidity and mortality in patients with hereditary CRC syndromes provides compelling evidence for the need to apply a methodological approach to assessing CRC risk. In the age of open-access colonoscopy, gastroenterologists need to differentiate patients who have a moderately increased risk of CRC from those who may be at high risk, need specialized cancer surveillance and genetic counseling. The CRC Risk Assessment Tool is a simple, robust tool that is easy to use and has good, reproducible ability to identify individuals who may be at high risk of hereditary CRC and would benefit from a more thorough risk assessment to precisely determine appropriate CRC and extracolonic cancer surveillance recommendations.

Acknowledgments

REFERENCES

1. Imperiale T. Toward Risk Stratification for Screening and Surveillance of Colorectal Neoplasia: One Small Step for the Colonoscopist. Gastroenterology. 2007; 133:1364–1376. [PubMed: 17919505]

2. Lieberman DA, de Garmo PL, Fleischer DE, et al. Patterns of endoscopy use in the United States. Gastroenterology. 2000; 118:619–624. [PubMed: 10702214]

3. Harewood GC, Lieberman DA. Colonoscopy practice patterns since introduction of medicare coverage for average-risk screening. Clin Gastroenterol Hepatol. 2004; 2:72–7. [PubMed: 15017635]

4. Byers T, Levin B, Rothenberger D, et al. American Cancer Society guidelines for screening and surveillance for early detection of colorectal polyps and cancer: update 1997. CA Cancer J Clin. 1997; 47:154–160. [PubMed: 9152173]

5. Jarvinen HJ, Aarnio M, Mustonen H, et al. Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer. Gastroenterology. 2000; 118:829–834. [PubMed: 10784581]

6. de Jong AE, Hendriks YM, Kleibeuker JH, et al. Decrease in mortality in Lynch syndrome families because of surveillance. Gastroenterology. 2006; 130(3):665–671. [PubMed: 16530507]

7. Heiskanen I, Luostarinen T, Jarvinen HJ. Impact of screening examinations on survival in familial adenomatous polyposis. Scand J Gastroenterol. 2000; 35(12):1284–1287. [PubMed: 11199368]

8. Vasen HFA, Mecklin JP, Meera Khan P, et al. The international collaborative group on HNPCC. Dis Colon Rectum. 1991; 34:424–425. [PubMed: 2022152]

9. Vasen HFA, Watson P, Mecklin JP, et al. New criteria for HNPCC proposed by the ICG-HNPCC. Gastro. 1999; 116:1453–1456.

10. Rodriquez-Bigas MA, Boland CR, Hamilton SR, et al. An NCI workshop on HNPCC: meeting highlights and Bethesda guidelines. J Natl Cancer Inst. 1997; 89:1758–1762. [PubMed: 9392616]

11. Umar A, Boland CR, Terdiman JP, et al. Revised Bethesda Guidelines for hereditary Nonpolyposis colorectal cancer (Lynch Syndrome) and microsatellite instability. J Natl Cancer Inst. 2004; 96:261–268. [PubMed: 14970275]

12. Stoffel E, Grover S, Russo L, et al. Physician Awareness of Cancer Screening Recommendations for Hereditary Nonpolyposis Colorectal Cancer is Limited. Gastroenterology. 2003; 124(4):A646.

13. Stoffel EM, Garber JE, Grover S, et al. Cancer Surveillance is often inadequate in individuals at high-risk for colorectal cancer. J Med Genet. 2003; 40(5):e54. [PubMed: 12746408]

14. Giardiello FM, Brensinger JD, Petersen GM. AGA technical review on hereditary colorectal cancer and genetic testing. Gastroenterology. 2001; 121:195–213. [PubMed: 11438508]

15. Ramsey SD, Clark L, Etzioni R, et al. Cost-effectiveness of microsatellite instability screening as a method of detecting hereditary nonpolyposis colorectal cancer. Ann Intern Med. 2001; 135:577–588. [PubMed: 11601929]

16. Kievit W, de Bruin JH, Adang EM, et al. Cost-effectiveness of a new strategy to identify HNPCC patients. Gut. 2005; 54:97–102. [PubMed: 15591512]

17. Grover S, Stoffel EM, Syngal S, et al. Physician Assessment of family cancer history and referral for genetic evaluation in colorectal cancer patients. Clin Gastroenterol Hepatol. 2004; 2:813–819. [PubMed: 15354282]

18. Raedle J, Trojan J, Brieger A, et al. Bethesda Guidelines: relation to microsatellite instability and MLH1 promoter methylation in patients with colorectal cancer. Ann Intern Med. 2001; 135:566–576. [PubMed: 11601928]

19. Furukawa T, Konishi F, Shitoh K, et al. Evaluation of screening strategy for detecting hereditary nonpolyposis colorectal carcinoma. Cancer. 2002; 94:911–920. [PubMed: 11920458]

20. Pinol V, Castells A, Andreu M, et al. Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer. JAMA. 2005; 293:1986–1994. [PubMed: 15855432]

21. Breiman, L.; Freidman, JH.; Olshen, RA., et al. Classification and Regression Trees. Wadsworth International Group; Belmont, CA: 1984.

22. Balmaña J, Stockwell DH, Steyerberg EW, et al. Prediction of MLH1 and MSH2 Mutations in Lynch Syndrome. JAMA. 2006; 296:1469–1478. [PubMed: 17003395]

23. Selvachandran SN, Hodder RJ, Ballal MS, et al. Prediction of colorectal cancer by a patient consultation questionnaire and scoring system: a prospective study. Lancet. 2002; 360:278–283. [PubMed: 12147370]

24. Church JM. A scoring system for the strength of a family history of colorectal cancer. Dis Colon Rectum. 2005; 48:889–96. [PubMed: 15785893]

25. Goldman L, Weinberg M, Weisberg M, et al. A computer-derived protocol to aid in the diagnosis of emergency room patients with acute chest pain. N Eng J Med. 1982; 307:588–596.

26. Goldman L, Cook EF, Brand DA, et al. A computer protocol to predict myocardial infarction in emergency department patients in chest pain. N Engl J Med. 1988; 318:797–803. [PubMed: 3280998]

27. Temkin NR, Holubkov R, Machamer JE, et al. Classification and regression trees (CART) for prediction of function at 1 year following head trauma. J Neurosurg. 1995; 82:764–771. [PubMed: 7714600]

28. Garzotto M, Beer TM, Hudson G, et al. Improved detection of prostate cancer using classification and regression tree analysis. J Clin Oncol. 2005; 23:4322–4329. [PubMed: 15781880]

29. Fonarow GC, Asams KF, Abraham WT, et al. Risk Stratification for In-Hospital Mortality in Acutely Decompensated Heart Failure: Classification and regression tree Analysis. JAMA. 2005; 293:572–580. [PubMed: 15687312]

30. Murff HJ, Spigel DR, Syngal S. Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history. JAMA. 2004; 292:1480–1489. [PubMed: 15383520]

31. Douglas FS, O'Dair LC, Robinson M, et al. The accuracy of diagnoses as reported in families with cancer: a retrospective study. J Med Genet. 1999; 36:309–312. [PubMed: 10227399]

32. Love RR, Evans AM, Josten DM. The accuracy of patient reports of a family history of cancer. J Chronic Dis. 1985; 38:289–293. [PubMed: 3998045]

33. Sijmons RH, Boonstra AE, Reefhuis J, et al. Accuracy of family history of cancer: clinical genetic implications. Eur J Hum Genet. 2000; 8:181–186. [PubMed: 10780783]

34. Wijnen JT, Vasen HF, Khan PM, et al. Clinical findings with implications for genetic testing in families with clustering of colorectal cancer. N Engl J Med. 1998; 339:511–518. [PubMed: 9709044]

35. Lipton LR, Johnson V, Cummings C, et al. Refining the Amsterdam Criteria and Bethesda Guidelines: testing algorithms for the prediction of mismatch repair mutation status in the familial cancer clinic. J Clin Oncol. 2004; 22:4934–4943. [PubMed: 15611508]

36. Barnetson RA, Tenesa A, Farrington SM, et al. Identification and survival of carriers of mutations in DNA mismatch repair genes in colon cancers. N Enl J Med. 2006; 354:2751–2763.

37. Chen S, Wang W, Lee S, et al. Prediction of germline mutations and cancer risk in the Lynch syndrome. JAMA. 2006; 296:1479–1487. [PubMed: 17003396]

38. Balaguer F, Balmaña J, Castellví-Bel S, et al. Validation and extension of the PREMM1,2 model in a population-based cohort of colorectal cancer patients. Gastroenterology. 2008; 134:39–46. [PubMed: 18061181]

39. Syngal S, Weeks JC, Schrag D, et al. Benefits of colonoscopic surveillance and prophylactic colectomy in mutation carriers for hereditary nonpolyposis colorectal cancer. Ann Intern Med. 1998; 129:787–796. [PubMed: 9841584]

40. Sweet KM, Bradley TL, Westman JA. Identification and referral of families at high risk for cancer susceptibility. J Clin Oncol. 2002; 20:528–537. [PubMed: 11786583]

41. Lindor NM, Rabe K, Petersen GM, et al. Lower cancer incidence in Amsterdam I criteria families without mismatch repair deficiency: familial colorectal cancer type X. JAMA. 2005; 293:1979–1985. [PubMed: 15855431]

42. Halbert CH, Lynch J, Lynch J, et al. Colon cancer screening practices following genetic testing for hereditary nonpolyposis colon cancer (HNPCC) mutations. Arch Int Med. 2004; 164(17):1881–7. [PubMed: 15451763]

STUDY HIGHLIGHTS

1) What is current knowledge

- Diagnostic criteria to identify individuals with hereditary colorectal cancer (CRC) syndromes is complex and underutilized

- "Open-access" colonoscopy makes it challenging to identify patients who may be at high risk for CRC

- Patients who may be at high risk for CRC need genetic counseling and testing, intensive endoscopic surveillance, and screening for extracolonic tumors

2) What is new here

- We have developed a simple, three-question CRC Risk Assessment Tool to help identify the majority of patients who may be at high risk for CRC

- The CRC Risk Assessment Tool can be easily incorporated into routine preprocedural evaluations for patients undergoing colonoscopy

- The CRC Risk Assessment Tool can accurately detect 95% of known MMR gene mutation carriers with Lynch Syndrome

615 Patients (109 High Risk)
1. Do you have a FDR with a Lynch syndrome-related cancer diagnosed before age 50?

YES N=31   NO

HIGH RISK
Sens: 28.4%

584 Patients (78 High Risk)
2. Do you have ≥ 3 relatives with CRC?

YES N=25   NO

HIGH RISK
Sens: 51.4%

559 Patients (53 High Risk)
3. Do you have a FDR with colorectal polyps diagnosed before age 50?

YES N=19   NO

HIGH RISK
Sens: 68.8%

540 Patients (34 High Risk)
4. Have you had colorectal polyps diagnosed before age 50?

YES N=13   NO

HIGH RISK
Sens: 80.7 %

527 Patients (21 High Risk)
5. Have you had a FDR diagnosed with CRC before age 50?

YES N=9   NO

HIGH RISK
Sens: 89.0%

518 Patients (12 High Risk)
6. Have you had a total of ≥ 10 colorectal polyps?

YES N=5   NO

HIGH RISK
Sens: 93.6%

513 Patients (7 High Risk)
7. Have you had endometrial cancer diagnosed before age 50?

YES N=4   NO

HIGH RISK
Sens: 97.2%

509 Patients (3 High Risk)
8. Have you had CRC diagnosed before 50?

YES N=2   NO

HIGH RISK
Sens: 99.1%

507 Patients (1 High Risk)
9. Have you had any Lynch syndrome-related cancer (other than endometrial cancer) diagnosed before age 50?

YES N=1   NO

HIGH RISK
Sens: 100%

506 Patients
Not High Risk

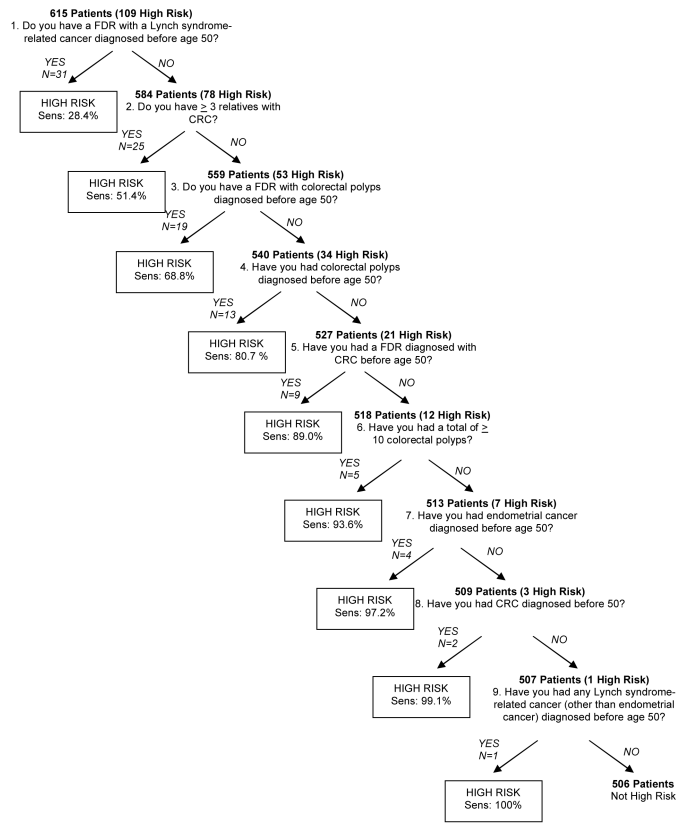**Figure 1.**
Development of the CRC Risk Assessment Tool
CRC=colorectal cancer, FDR=first-degree relative, Sens=cumulative sensitivity

615 (5335) Patients
1. Do you have a FDR with a Lynch syndrome-related cancer or CRC diagnosed before age 50?

YES
N=47(627)

HIGH RISK
Sens: 43.1%
Sens: 59.0%

NO

568 (4708) Patients
2. Have you had CRC or colorectal polyps diagnosed before age 50?

YES
N=19(108)

HIGH RISK
Sens: 61%
Sens: 69%

NO

549 (4600) Patients
3. Do you have ≥ 3 relatives with CRC?

YES
N=18(87)

HIGH RISK
Sens: 77%
Sens: 77%

NO

531 (4513) Patients
4. Do you have a FDR with colorectal polyps diagnosed before age 50?

YES
N=15(177)

HIGH RISK
Sens: 90.8%
Sens: 93.0%

NO

516 (4336) Patients
5. Have you had a total of ≥ 10 colorectal polyps?

YES
N=5(16)

HIGH RISK
Sens: 95.4%
Sens: 95.0%

NO

511 (4320) Patients
6. Have you had any Lynch syndrome-related cancers diagnosed before age 50?

YES
N=5(54)

HIGH RISK
Sens: 100%
Sens: 100%

NO

506 (4266) Patients
Not High Risk

**Three- question
CRC Risk Assessment Tool
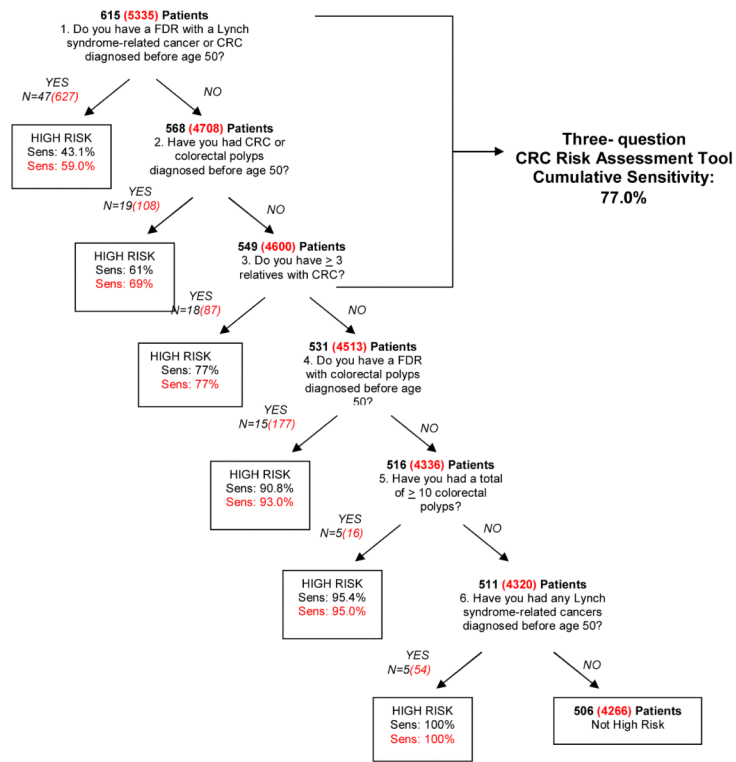Cumulative Sensitivity:
77.0%**

Figure 2.
Validation of the CRC Risk Assessment Tool
■ = derivation cohort results, ■ = validation cohort results CRC=colorectal cancer, FDR=first-degree relative, Sens=cumulative sensitivity

	YES	NO
• Do you have a **first-degree relative** (mother, father, brother, sister or child) with any of the following conditions diagnosed before age 50?		
Colon or rectal cancer	☐	☐
Cancer of the uterus, ovary, stomach, small intestine, urinary tract (kidney, ureter, bladder), bile ducts, pancreas, or brain	☐	☐
• Have **you** had any of the following conditions diagnosed before age 50?		
Colon or rectal cancer	☐	☐
Colon or rectal polyps	☐	☐
• Do you have 3 or more relatives with a history of colon or rectal cancer? (this includes parents, brothers, sisters, children, grandparents, aunts, uncles and cousins)	☐	☐

Figure 3.
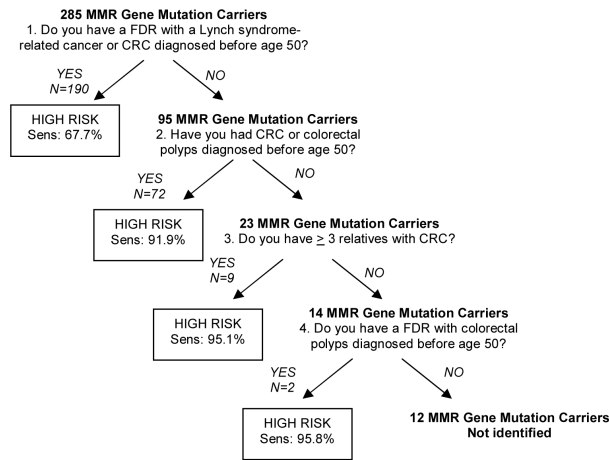The pre-procedural Colorectal Cancer Risk Assessment Tool

Figure 4.
Ability of CRC Risk Assessment Tool to Identify MMR Gene Mutation Carriers
MMR=mismatch repair, FDR=first-degree relative, CRC=colorectal cancer,
Sens=sensitivity

Table 1

Patient Demographics

Subjects	Derivation Cohort				Validation Cohort			
	TOTAL		HIGH RISK		TOTAL		HIGH RISK	
	N	(%)	N	(%)	N	(%)	N	(%)
Overall (N=615)	615	(100)	109	(17.7)	5335	(100)	1069	(20.0)
Gender								
Male	287	(46.7)	38	(34.9)	1928	(45.2)	397	(37.1)
Female	327	(53.3)	71	(65.1)	2338	(54.8)	672	(62.9)
Age (years)								
Mean ± SD	57.8 ± 11.1		56.0 ± 12.9		55.7 ± 14.0		55.0 ± 13.7	
Race/Ethnicity					*		*	
White	532	(86.5)	94	(86.2)				
African American	22	(3.6)	3	(2.8)				
Hispanic/Latin	30	(4.9)	6	(5.5)				
Other	17	(2.8)	5	(4.6)				
Missing	14	(2.3)	1	(0.9)				
Level of Education					*		*	
HS/some college	202	(32.8)	40	(36.7)				
College graduate	170	(27.6)	37	(33.9)				
Postgraduate	235	(38.2)	31	(28.5)				
Missing	8	(1.3)	1	(0.9)				
Household Income					*		*	
<$50,000	155	(25.2)	35	(32.1)				
$50,000-$99,999	160	(26.0)	29	(26.6)				
$100,000-149,999	95	(15.5)	16	(14.7)				
>$150,000	129	(21.0)	18	(16.5)				
Missing	76	(12.4)	11	(10.1)				

SD=standard deviation, HS=HighSchool,

*
=not available

Table 2

High-risk Patient Characteristics

HIGH RISK Criteria	Derivation Cohort N	(%)	Validation Cohort N	(%)
Total eligible subjects	615		5,335	
Total high-risk subjects	109 *†	(17.7)	1,069 *‡	(20.0)
PERSONAL HISTORY§:				
Polyp 50 years	19	(17.4)	128	(12.0)
10 cumulative polyps	7	(6.4)	42	(3.9)
CRC 50 years	6	(5.5)	33	(3.1)
Endometrial Cancer 50 years	3	(2.8)	35	(3.3)
Lynch syndrome-related cancer 50 years	2	(1.8)	54	(5.1)
Subjects fulfilling:				
one criterion	29	(26.6)	238	(22.3)
two criteria	4	(3.7)	26	(2.4)
three criteria	0	(0)	3	(0.3)
Total:	33	(30.3)	267	(25.0)
FAMILY HISTORY§:				
FDR with Lynch syndrome-related cancer 50 years	31	(28.4)	410	(38.4)
3 relatives with CRC	30	(27.5)	164	(15.3)
FDR with polyp 50 years	23	(21.1)	320	(30.0)
FDR with CRC 50 years	21	(19.3)	298	(27.9)
Subjects fulfilling:				
one criterion	66	(60.5)	701	(65.6)
two criteria	16	(14.7)	158	(14.7)
three criteria	1	(0.9)	37	(3.5)
four criteria	1	(0.9)	16	(1.5)
Total:	84	(77.0)	912	(85.3)

CRC=colorectal cancer, FDR=first-degree relative

†Eight patients were defined as high-risk based on both personal and family history

‡Twelve patients were defined as high-risk based on both personal and family history

§Percentages (%) related to Personal and Family History criteria correspond to proportion of high-risk individuals meeting each criterion

SD= standard deviation, HS=High School,

*=not available