A prognostic model for advanced colorectal neoplasia recurrence

Item Type	Article
Authors	Liu, Lin; Messer, Karen; Baron, John A.; Lieberman, David A.; Jacobs, Elizabeth T.; Cross, Amanda J.; Murphy, Gwen; Martinez, Maria Elena; Gupta, Samir
Citation	A prognostic model for advanced colorectal neoplasia recurrence 2016, 27 (10):1175 Cancer Causes & Control
DOI	10.1007/s10552-016-0795-5
Publisher	SPRINGER
Journal	Cancer Causes & Control
Rights	© Springer International Publishing Switzerland (outside the USA) 2016.
Download date	26/08/2022 03:36:15
Item License	http://rightsstatements.org/vocab/InC/1.0/
Version	Final accepted manuscript
Link to Item	http://hdl.handle.net/10150/621531

TITLE PAGE

Title: A Prognostic Model for Advanced Colorectal Neoplasia Recurrence

Short title: Model for Advanced Colorectal Neoplasia Recurrence

Liu Lin[1,2], Messer Karen[1,2], Baron John A[3], Lieberman David A[4], Jacobs Elizabeth T[5], Cross Amanda J[6,7], Murphy Gwen[8], Martinez Maria Elena[1,2], Gupta Samir [2,9,10]*.

Affiliations:
1. Department of Family Medicine and Public Health, Division of Biostatistics and Bioinformatics, University of California San Diego, La Jolla, CA
2. University of California San Diego Moores Cancer Center, La Jolla, CA
3. Department of Medicine, University of North Carolina School of Medicine, Chapel Hill, NC
4. Division of Gastroenterology and Hepatology, Veterans Affairs Medical Center and Oregon Health and Science University, Portland, Oregon
5. Arizona Cancer Center, Arizona College of Public Health, University of Arizona, Tucson, AZ
6. Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, UK
7. Cancer Screening and Prevention Research Group, Department of Surgery and Cancer, Imperial College London, UK
8. Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Rockville, MD
9. Veteran Affairs San Diego Healthcare System, San Diego, CA
10. Department of Internal Medicine, Division of Gastroenterology, University of California San Diego, La Jolla, CA

*Corresponding Author:
Samir Gupta, MD, MSCS
3350 La Jolla Village Dr, MC 111D
San Diego, CA 92161
Email: s1gupta@ucsd.edu
Phone: (858) 552-8585x2475   Fax: (858) 552-4327

Abbreviations:

AN - advanced neoplasia

AUC - Area under the Receiver Operating Characteristic Curve

BMI - body mass index

CI - confidence intervals

CRC – colorectal cancer

LASSO - L1-regularized logistic regression model

NRI - net reclassification improvement

ROC - Receiver Operating Characteristic Curve

USMSTF - US Multisociety Task Force

Authorship Statement

Guarantor of article: Samir Gupta, MD, MSCS

Specific Author Contributions:

Study concept and design: LL; KM; MEM; SG

Acquisition of data: All

Analysis and interpretation of data: All

Drafting of the manuscript: LL; KM; SG

Critical revision of the manuscript for important intellectual content: All

Statistical analysis: LL; KM

Obtained funding: LL; KM; JAB; DAL; AJC; MG; MEM; SG

All authors approved the final version of the article, including the authorship list.

ABSTRACT

Purpose: Following colonoscopic polypectomy, US Multisociety Task Force (USMSTF) guidelines stratify patients based on risk for subsequent advanced neoplasia (AN) using number, size, and histology of resected polyps, but have only moderate sensitivity and specificity. We hypothesized that a state-of-the-art statistical prediction model might improve identification of patients at high risk for future AN, and address these challenges.

Methods: Data were pooled from 7 prospective studies which had follow-up ascertainment of metachronous AN within 3-5 years of baseline polypectomy (combined n=8,228). Pooled data were randomly split into training (n=5,483) and validation (n=2,745) sets. A prognostic model was developed using best practices. Two risk cut-points were identified in the training data which achieved a 10 percentage point improvement in sensitivity and specificity, respectively, over current USMSTF guidelines. Clinical benefit of USMSTF vs. model-based risk stratification was then estimated using validation data.

Results: The final model included polyp location, prior history and patient age as well as number, size and histology of resected polyps. The first risk cut-point improved sensitivity but with loss of specificity. The second risk cut-point improved specificity without loss of sensitivity (specificity 46.2% model vs. 42.1% guidelines, p<0.001; sensitivity 75.8% model vs. 74.0% guidelines, p=0.64). Estimated AUC was 65% (95% CI 62%, 69%).

Conclusion: This model-based approach allows flexibility in trading sensitivity and specificity, which can optimize colonoscopy over- vs under-use rates. Only modest improvements in prognostic power are possible using currently available clinical data. Research considering additional factors such as adenoma detection rate for risk prediction appears warranted.

Keywords: polyp surveillance; risk stratification; epidemiology; colorectal cancer; colorectal polyps

Introduction

        Colorectal cancer (CRC) is a leading cause of cancer death worldwide. For example, annually in the United States, 130,000 individuals develop CRC, and about 50,000 die of the disease annually[1]. CRC can be prevented by the identification and removal of colorectal polyps[2]. After polypectomy, recommendations are routinely provided for repeat surveillance colonoscopy, with the goal of reducing risk of CRC-associated morbidity and mortality based on practice guidelines[3]. For example in the US, current surveillance intervals are largely based on US Multisociety Task Force (USMSTF) guidelines, which stratify individuals for risk of future advanced neoplasia (AN) based on the number, size, and histology of resected polyps.

        Although this approach is standard-of-care after colonoscopy, evidence suggests that current guidelines can be improved[4]. Indeed, the sensitivity and specificity of USMSTF guidelines for predicting metachronous AN are estimated to be 59 to 81% and 43 to 58%, respectively[5-9]. Thus, many patients who later develop AN are classified as low risk at baseline, and many who remain free of AN on follow up are classified as high risk at baseline. In different patient groups this results in over- or under-use of surveillance colonoscopy, exposing many individuals to the risks and costs of unnecessary colonoscopy, while others miss an opportunity for early detection and prevention. Improved risk stratification strategies are needed.

        Incorporation of additional risk factors beyond the number, size, and histology of resected polyps might improve risk stratification. For example, age, sex, body mass index (BMI), and polyp location have each been associated with risk for metachronous AN, but have not been formally incorporated into risk stratification guidelines[5, 7, 8, 10-12]. Our aim was to determine whether a statistical model incorporating additional clinical factors could improve post polypectomy risk stratification compared to USMSTF guidelines, utilizing a large pooled dataset of over 8,000 individuals who underwent polypectomy and subsequent surveillance colonoscopy.

Materials and Methods

Design and participants

Data were pooled from 7 prospective studies[13-19] of patients with sporadic colorectal adenoma in North America that included 8,228 individuals with polypectomy and repeat surveillance colonoscopy within 3 to 5 years, as previously described[5]. These studies ascertained baseline data on patient and adenoma characteristics considered as predictors in the analyses and assessed the number, size, histopathology of resected polyps, adenomas and CRCs detected at follow-up.

Primary outcome

The primary outcome was metachronous AN within 3 to 5 years of polypectomy, defined by any of the following: adenoma with size $\geq$1cm, high-grade dysplasia, and/or tubulovillous or villous histology or adenocarcinoma. AN occurring within six months of qualifying colonoscopy was counted as part of baseline rather than follow up findings.

USMSTF guidelines[3]

The USMSTF guideline low risk group includes patients with 1 or 2 small (<1 cm), tubular adenoma(s). The USMSTF guideline high risk group includes patients with 3 to 10 adenomas, or who have any adenoma with size $\geq$ 1cm, >25% villous features, or high-grade dysplasia. The USMSTF guideline also specifies a very high (highest) risk group with more than 10 adenomas. Very few subjects (n=11) in our data were in highest risk group, therefore, we combined high-risk and highest-risk subjects into one group.

Baseline risk factors (predictors)

In addition to adenoma size, number, and histology, we considered adenoma location (distal colorectum, proximal only, proximal and distal) at qualifying colonoscopy, age, sex, race/ethnicity, family history of CRC, cigarette smoking (current, former, never), body mass index (BMI), and history of prior polyp. Additionally, unlike current USMSTF guidelines, we considered number and size of adenoma(s) as continuous rather than categorical variables. To address missing data, we created an unknown category for variables for which more than 4% of patients were missing information including family history of CRC, polyp location, and history of previous polyp. Polyps with non-adenomatous histology were grouped into one category. Presence of high-grade dysplasia was not characterized by two studies; therefore this variable was characterized as unknown for patients from these two studies[14, 16].

Statistical analysis

All analyses were performed using R*[20].* Patients were randomly assigned 2:1 to training (n=5,483) and validation (n=2,745) data sets. Training data were used to develop our

prognostic model and identify risk stratification cut-points. Validation data were used to compare sensitivity, specificity and estimated clinical benefit between the model-based risk stratification (as developed in the training data) and existing USMSTF guidelines.

Development of the prognostic model: Risk factors that were significantly (p<0.15) associated with the outcome in univariate analysis were considered as potential predictors in a multivariable logistic regression model with the outcome AN. We used two complementary methods for variable selection that are considered to be superior to traditional stepwise model selection approaches[21, 22]. First, an L1-regularized logistic regression model (LASSO)[23, 24] was used to assess the order of entry of variables into the model, using the *glmpath* package in R. To assess stability of the entry order, a sensitivity analysis using bootstrap LASSO based on 1,000 samples was conducted. Second, Bayesian model averaging (BMA)[25] was used for variable selection, implemented in the *BMA* package in R. The performance of the models selected by BMA was compared using generalized R^2 and Brier's scores, discrimination was assessed using the Area under the Receiver Operating Characteristic Curve (AUC), and calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test[26]. Among the top performing BMA models, we prioritized the model which was most consistent with the variables selected by LASSO.

Identification of cut-points for risk stratification: We then used the predicted probability of AN from the selected best model to determine a cut-point above which a patient would be identified as at high risk for AN. We made an *a priori* plan to identify two risk stratification cut-points. Defining sensitivity as the proportion of individuals with metachronous AN at follow-up who were classified as high risk at baseline (Table 1), we targeted the first cut-point to achieve a 10 percentage point improvement in sensitivity compared to the current USMSTF guidelines. Defining specificity as the proportion of individuals without metachronous AN at follow-up who were classified as low risk at baseline (Table 1), we targeted the other cut-point to achieve a 10 percentage point improvement in specificity[27-29]. The population sensitivity and specificity of USMSTF guidelines were estimated on the entire set of pooled data. The cut-point determination included all subjects in the training data with defined values for all predictors in the final model (including those predictors assigned an unknown category for missing values) and used the *OptimalCutpoints* package in R.

Model validation and comparison of estimated clinical benefit: Model validation was performed on all subjects that could be classified by both the model-based method and the USMSTF guidelines in the validation data set. Model discrimination was assessed by the AUC. Model calibration was assessed by comparing the predicted risk and observed risk of AN for 10

deciles of risk groups. Potential clinical benefit (Table 1) was assessed in the validation data using the model coefficients and cut-points identified in the training data, by estimated sensitivity and specificity for metachronous AN and estimated rates of over- and under-use of colonoscopy. Overuse of surveillance colonoscopy was defined as the proportion of those classified as high risk at baseline who did not develop metachronous AN (i.e., 1 - positive predictive value). Underuse of surveillance colonoscopy was defined as the proportion of those classified as low risk at baseline who developed AN (1 - negative predictive value). Improvement in specificity and sensitivity using the predictive model on the validation data was assessed by McNemar's test. Differences in overuse and underuse between the predictive model and current guidelines were assessed by 95% confidence intervals (CI) based on 1,000 bootstrap samples. Clinical benefit of using the predictive model over current guidelines was also assessed using net reclassification improvement (NRI)[30].

Results

Analytic Cohort Selection

The final dataset included 8,228 patients, randomly split into 5,483 training and 2,745 validation subjects. There were no major differences in baseline patient characteristics between the training and validation data sets. Using current USMSTF guidelines, 30.1% of the cohort was classified as low risk, 47.9% as high risk, and 22% could not be classified due to missing data on adenoma size, adenoma number, histology or high-grade dysplasia, which are required by USMSTF guidelines for risk classification. As described in the *Methods*, the unclassified subjects were included in predictive model development in the training data, but were excluded from model assessment and the comparison of the clinical benefit between the USMSTF guidelines and the predictive model in the validation data.

Model Development

In univariate analyses of the training dataset, age, sex, history of prior polyps, adenoma number, adenoma size, adenoma location, presence of tubulovillous or villous histology, and presence of high-grade dysplasia were significantly associated with risk for AN on follow up (Table 3). Including these variables in multivariable logistic regression, LASSO variable selection and bootstrap samples showed that sex and presence of high-grade dysplasia were less important predictors, consistently entering the model last. Considering the top two best models selected by BMA, we found that both models did not include sex and high-grade dysplasia which was consistent with the LASSO result. We also found that both models included age, history of prior polyp, adenoma number, size, presence of tubulovillous or villous histology, while one included adenoma location and the other did not. Since adenoma location was identified as an important predictor in LASSO, the model including adenoma location was selected as the final model. In the training dataset, the final model had AUC = 0.68 (95% CI: 0.66-0.70), with good calibration (goodness-of-fit test p=0.39).

Identification of cut points for risk stratification

Based on the entire set of pooled data, the sensitivity and specificity of USMSTF guidelines were estimated to be 78% and 41%, respectively. Using the selected model, two cut-points were identified to achieve the targeted sensitivity of 88% and the targeted specificity of 51%, respectively; that is, a 10 percentage point improvement in sensitivity and specificity over the guidelines. When targeting improvement in sensitivity, the cut-point of 0.075 for predicted probability of AN was selected, and the actual sensitivity achieved at this cut-point was 88.1% in the training data. When targeting improvement in specificity, the cut-point of 0.101 was selected, and the actual specificity achieved at this cut-point in the training data was 51.2%. These two

cut points were finalized prior to applying the final predictive model to the validation dataset.

Model Validation

When applied to the validation data, the AUC of the model was similar to that observed in the training data (AUC=0.65; 95% CI, 0.62-0.69, Figure 1). Model calibration was also similar, as assessed by Goodness-of-fit test (p=0.21), and ratios of predicted to observed risk of AN, which were generally close to 1.00 (range 0.68-1.17, see Table S1, published online). The sensitivity and specificity of USMSTF guidelines in the validation data were plotted on the same figure (Figure 1, red dot), and was observed to lie below the model-based ROC curve, indicating that the novel model has potential to improve upon risk stratification compared to current guidelines across a range of cut-points. For comparison, the sensitivity and specificity corresponding to the two *a priori* identified cut-points were also plotted in Figure 1, demonstrating that the first model cut-point (Figure 1, blue dot) had superior sensitivity compared to USMSTF guidelines, and also demonstrating that the second cut-point (Figure 1, green dot) had superior specificity compared to USMSTF guidelines.

Comparison of model-based and USMSTF guidelines for risk stratification

We found that USMSTF guideline sensitivity and specificity in the validation data for prediction of AN were estimated to be 74% and 42.1%, respectively, consistent with prior reports[5-8]. Further, we found that 83.2% of individuals classified as high risk at baseline by USMSTF guidelines in the validation dataset had no AN on follow up (consistent with surveillance colonoscopy overuse), and 8.9% classified as low risk developed AN (consistent with surveillance colonoscopy underuse). Using the two cut-points identified in the training data, we next compared the sensitivity and specificity, underuse and overuse rate between the model and the USMSTF guidelines in the validation data (Table 4).

When using the cut-point targeted to improve sensitivity (Figure 1, blue dot), we found estimated sensitivity for the model-based rule to be significantly higher than for the USMSTF guidelines (88.9% model vs. 74.0% guidelines, p<0.001), correctly identifying an additional net 43 subjects as high risk (14.9% of all patients with AN on follow up). However, there was a significant reduction in specificity when comparing the model-based and the USMSTF guidelines (27.7% model vs. 42.1% guidelines, p<0.001), incorrectly classifying an additional net 262 subjects as high risk (14.4% of all patients without AN on follow up). The model-based risk stratification was associated with a 2.9 percentage point reduction in estimated surveillance colonoscopy underuse, when compared to the USMSTF guidelines (6.0% model vs. 8.9% guidelines, p<0.05), without substantially changing estimated surveillance colonoscopy overuse among those classified as high risk (83.7% model vs 83.2% guidelines, p>0.05).

We then investigated the model cut-point targeted to improve specificity (Figure 1, green dot). Estimated specificity for the model was significantly superior to USMSTF guidelines in the validation data, however it did not reach the 10 percentage point improvement that we targeted in the training data (46.2% model vs. 42.1% guidelines, p<0.001). At this cut-point, the model correctly identified an additional net 76 subjects as low risk (4.2% of all patients without AN on follow up). There was no reduction in sensitivity (75.8% model vs. 74.0% guidelines, p=0.64), and the model correctly identified an additional net 5 patients as at high risk (1.7% of all patients with AN on follow up). Model performance at this cut-point was associated with a modest absolute reduction in estimated surveillance colonoscopy overuse (81.8% model vs. 83.2% guidelines, p<0.05), along with a non-significant reduction in colonoscopy underuse (7.7% model vs. 8.9% guidelines, p>0.05).

Discussion

We have demonstrated that a novel predictive model for metachronous AN, together with an *a priori* established risk cut-point, can modestly improve specificity (46.2% model vs. 42.1% guidelines, p<0.001) with no deterioration in sensitivity (75.8% model vs. 74.0% guidelines, p=0.64), when compared to current USMSTF guidelines for post-polypectomy risk stratification. Our predictive model was developed using best statistical practices for predictive modeling on one of the most comprehensive data sets available. These practices include *a priori* specification of the model and cut-point development in the training data using optimal prediction strategies, and pre-specifying the testing plan in the validation data. Statistical theory thus supports that the model likely represents a near optimal prediction rule based on available data, which is an advantage of a model-based approach over more *ad hoc* approaches. The improved test characteristics we observed were likely due to selection of additional variables into the final model which are not used by USMSTF guidelines, and also by employing some variables as continuous rather than categorical predictors (e.g. considering the absolute number of adenomas in the predictive model rather than categorizing the number of adenomas as 3 or more, or less than 3, as in the USMSTF guidelines). Although the magnitude of the estimated benefit was rather modest, our work provides proof-of-concept that improved modeling techniques and additional variables, beyond adenoma number, size, and histology as used for USMSTF guideline-based management, have potential to improve risk stratification of individuals with colorectal polyps, and merit further investigation.

An additional advantage of model-based risk stratification is the ability to tune sensitivity and specificity to clinical needs more finely by choosing a clinically appropriate cut-point. Indeed, as seen from Figure 1, a risk cut-point to maximize sensitivity or specificity could be identified. Thus, the model could be used to select strategies for surveillance that either maximize sensitivity or specificity depending on the priorities of patients, physicians, or policy makers. For example, use of the cut-point targeted to improve sensitivity in practice would have resulted in detection of an additional 14.9% of all individuals with metachronous AN on follow up. In situations with strained colonoscopy resources, or for patients at increased risk for colonoscopy associated complications due to factors such as age, use of the cut-point targeted to improve specificity could have resulted in reduced colonoscopy use, without an increase in underuse for individuals who developed metachronous AN. These examples also illustrate how a model-based approach can be used to investigate the balance between estimated population level costs of surveillance colonoscopy overuse and underuse across a range of risk stratification cut-points.

A number of cohort and case control studies have identified risk factors for metachronous AN after initial polypectomy[6, 8, 10, 11, 31-45]. These include patient characteristics (age, sex, BMI, diabetes, family history of CRC), polyp characteristics (adenoma size, number, location), and histology (villous and/or high-grade dysplasia). We also found many of these characteristics to be associated with risk for metachronous AN in our analysis. However, unlike prior reports, we extended our identification of risk factors for metachronous AN to develop and validate a statistical model for predicting this outcome. To our knowledge, there have not been previous reports of comprehensive statistical models that take into account multiple characteristics, and compare performance to established practice guidelines. As such, our findings confirm and extend prior work seeking to improve risk stratification of individuals after baseline polypectomy. In addition, our model-based stratification rule was based on choosing appropriate cut-points for predicted probability of risk and the performance of the risk stratification was evaluated through sensitivity and specificity, a strategy not commonly reported though it has the advantage of statistical rigor and direct clinical relevance.

Notably, our targeted cut points, which aimed to achieve substantial improvements in either sensitivity or specificity, could not simultaneously improve sensitivity and specificity. This illustrates the limits of the potential improvements supported by the current data, and highlights the need for incorporation of data from additional risk factors, which might support further improvements in overall prognostic power. It is likely that additional variables will need to be identified in order to achieve the goal of substantially improving both specificity and sensitivity at the same time. Although we included key patient and polyp characteristics that were not taken into account by USMSTF guidelines in our predictive model, there are other variables that might be influential but were not considered. For example, the adenoma detection rate has been found to be significantly associated with interval CRC after colonoscopy[46, 47], but was not available for our study. Colonoscopy is an operator-dependent test with performance characteristics (such as adenoma detection rate) varying within endoscopist and other factors[48]. Thus, indicators of the quality of colonoscopy are natural further candidates that could be considered in risk classification[35]. Indeed, because of the observed close relationship between colonoscopist adenoma detection rate and risk for interval cancer after colonoscopy[46, 47], it is possilbe that this parameter might be a very powerful prognostic variable for metachronous neoplasia. Aspirin and dietary variables, such as red meat, processed meat and fiber have also been linked with risk for CRC, and may also be considered as candidate variables for risk stratification for metachronous advanced neoplasia[5]. It is likely that additional variables (such as adenoma detection rate and aspirin exposure) would have a

substantial impact on the ability of the model to correctly identify both high risk and low risk patients. For variables already considered in the predictive model, additional information (such as types of previous polyp, serrated or not) might also improve model performance.

Several limitations may be considered in interpreting this work. First, colonoscopies contributing data to this study were performed between 1984 and 1998. Changes in colonoscopy quality may have occurred over time, such that individuals encountered in clinical practice now might be at different risk for metachronous AN than the individuals included in this study. Second, six of the seven studies contributing to this analysis were from prevention trials, and in two of the trials (one of aspirin, the other of calcium supplementation) a modest degree of intervention efficacy was shown. Further, studies contributing data were not limited to individuals undergoing first time colonoscopy. We are unable to ascertain whether these issues might have resulted in inclusion of a study population that was at higher or lower risk for advanced metachronous neoplasia compared to the general population. Thus, patients from included trials may not be representative of the general population of individuals with adenomas. Third, our training and validation sets were a random split of a larger dataset, and we did not have additional independent validation data sets available. Fourth, analyses may have been affected by the detail of variable ascertainment. For example, we used history of prior polyp as a predictor variable, but do not have detail on type of prior polyp. It is possible that within the prior polyp category, it is only those with prior history of AN who have increased risk. Fifth, some predictors had a high percentage of missing data (24% for high-grade dysplasia and 21% for history of previous polyps). We included an 'unknown' category for these variables in order to retain more subjects in the analysis and improve the study power. Note that our focus in this study is risk prediction, and optimal methods to address missing data in this context are an area of active statistical research. Further, family history data were collected differently across studies, and were aggregated in this analysis as history of CRC in one or more parents, siblings, or children. Details such as prior polyp type and relationship and age of family members with CRC were not available for analysis, and may have affected the performance of these variables in our models. All of these limitations might be addressed in the future by conducting analyses using large clinical datasets collected as part of usual care, or by developing prospective registries of patients undergoing polypectomy, with careful data collection.

In conclusion, our results suggest that a predictive model including polyp and patient characteristics can be used to improve post-polypectomy risk stratification. Such a model has potential to reduce the unnecessary colonoscopy risks and costs for low risk patients and

identify high risk individuals who might benefit from early surveillance colonoscopy. However, in order to substantially improve prognostic models for planning colonoscopy surveillance, additional clinical variables associated with colonoscopy outcomes appear to be needed. In ongoing work, we will incorporate quality factors and other additional patient and polyp factors to build on our current methodology for predicting metachronous AN after initial polypectomy. More research is needed to develop and validate prognostic models for planning colonoscopy surveillance, and this should eventually lead to prospective trials that test model-based strategies for surveillance.

References

1. Siegel RL, Miller KD, Jemal A (2015) Cancer statistics, 2015. CA Cancer J Clin 65:5-29.

2. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR et al (2010) Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. Lancet 375:1624-33.

3. Lieberman DA, Rex DK, Winawer SJ, Giardiello FM, Johnson DA et al (2012) Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology 143:844-57.

4. Martínez ME, Ahnen D, Greenberg ER (2013) One-year risk for advanced colorectal neoplasia. Ann Intern Med 158:639.

5. Martínez ME, Baron JA, Lieberman DA, Schatzkin A, Lanza E et al (2009) A pooled analysis of advanced colorectal neoplasia diagnoses after colonoscopic polypectomy. Gastroentereolgy 136:832-41.

6. Laiyemo AO, Murphy G, Albert PS, Sansbury LB, Wang Z et al (2008) Postpolypectomy colonoscopy surveillance guidelines: predictive accuracy for advanced adenoma at 4 years. Ann Intern Med 148:419-26.

7. Pinsky PF, Schoen RE, Weissfeld JL, Church T, Yokochi LA et al (2009) The yield of surveillance colonoscopy by adenoma history and time to examination. Clin Gastroenterol Hepatol 7:86-92.

8. Chung SJ, Kim YS, Yang SY, Song JH, Kim D et al (2011) Five-year risk for advanced colorectal neoplasia after initial colonoscopy according to the baseline risk stratification: a prospective study in 2452 asymptomatic Koreans. Gut 60:1537-43.

9. Stegeman I, de Wijkerslooth TR, Stoop EM, van Leerdam ME, Dekker E et al (2013) Colorectal cancer risk factors in the detection of advanced adenoma and colorectal cancer. Cancer Epidemiol 37:278-83.

10. Saini SD, Kim HM, Schoenfeld P (2006) Incidence of advanced adenomas at surveillance colonoscopy in patients with a personal history of colon adenomas: a meta-analysis and systematic review. Gastrointest Endosc 64:614-26.

11. van Heijningen EM, Lansdorp-Vogelaar I, Kuipers EJ, Dekker E, Lesterhuis W et al (2013) Features of adenoma and colonoscopy associated with recurrent colorectal neoplasia based on a large community-based study. Gastroenterology 144:1410-8.

12. Laiyemo AO, Pinsky PF, Marcus PM, Lanza E, Cross AJ et al (2009) Utilization and yield of surveillance colonoscopy in the continued follow-up study of the polyp prevention trial. Clin Gastroenterol Hepatol 7:562-7.

13. Schatzkin A, Lanza E, Corle D, Lance P, Iber F et al (2000) Lack of effect of a low-fat, high-fiber diet on the recurrence of colorectal adenomas. Polyp Prevention Trial Study Group. N Engl J Med 342:1149-55.

14. Baron JA, Cole BF, Sandler RS, Haile RW, Ahnen D et al (2003) A randomized trial of aspirin to prevent colorectal adenomas. N Engl J Med 348:891-9.

15. Alberts DS, Martínez ME, Roe DJ, Guillen-Rodriguez JM, Marshall JR et al (2000) Lack of effect of a high-fiber cereal supplement on the recurrence of colorectal adenomas. Phoenix Colon Cancer Prevention Physicians' Network. N Engl J Med 342:1156-62.

16. Baron JA, Beach M, Mandel JS, van Stolk RU, Haile RW et al (1999) Calcium supplements for the prevention of colorectal adenomas. Calcium Polyp Prevention Study Group. N Engl J Med 340:101-7.

17. Greenberg ER, Baron JA, Tosteson TD, Freeman DH, Jr., Beck GJ et al (1994) A clinical trial of antioxidant vitamins to prevent colorectal adenoma. Polyp Prevention Study Group. N Engl J Med 331:141-7.

18. Alberts DS, Martínez ME, Hess LM, Einspahr JG, Green SB et al (2005) Phase III trial of ursodeoxycholic acid to prevent colorectal adenoma recurrence. J Natl Cancer Inst 97:846-53.

19. Lieberman DA, Weiss DG, Bond JH, Ahnen DJ, Garewal H et al (2000) Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. N Engl J Med 343:162-8.

20. Dean CB, Nielsen JD (2007) Generalized linear mixed models: a review and some extensions. Lifetime Data Anal 13:497-512.

21. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD (2000) Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 19:1059-79.

22. Wang D, Zhang W, Bakhai A (2004) Comparison of Bayesian model averaging and stepwise methods for model selection in logistic regression. Stat Med 23:3451-67.

23. Friedman J, Hastie T, Tibshirani R (2010) Regularization Paths for Generalized Linear Models via Coordinate Descent. J Stat Softw 33:1-22.

24. Park MY, Hastie T (2007) L1-Regularization Path Algorithm for Generalized Linear Models. J R Stat Soc Series B Stat Methodol 69:659-77.

25. Volinsky CT, Madigan D, Raftery AE, Kronmal RA (1977) Bayesian Model Averaging in Proportional Hazard Models: Assessing the Risk of a Stroke. J R Stat Soc Ser C Appl Stat 46:433-48.

26. Hosmer D, Lemeshow, S (2000) Applied Logistic Regression. Wiley, New York

27. Vermont J, Bosson JL, François P, Robert C, Rueff A et al (1991) Strategies for graphical threshold determination. Comput Methods Programs Biomed 35:141-50.

28. Schäfer H (1989) Constructing a cut-off point for a quantitative diagnostic test. Stat Med 8:1381-91.

29. Gallop RJ, Crits-Christoph P, Muenz LR, Tu XM (2003) Determination and Interpretation of the Optimal Operating Point for ROC Curves Derived Through Generalized Linear Models. Understanding Statistics 2:219-242.

30. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS (2008) Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 27:157-72.

31. Seo JY, Chun J, Lee C, Hong KS, Im JP et al (2015) Novel risk stratification for recurrence after endoscopic resection of advanced colorectal adenoma. Gastrointest Endosc 81:655-64.

32. Fairley KJ, Li J, Komar M, Steigerwalt N, Erlich P (2014) Predicting the risk of recurrent adenoma and incident colorectal cancer based on findings of the baseline colonoscopy. Clin Transl Gastroenterol 5:e64.

33. van Enckevort CC, de Graaf AP, Hollema H, Sluiter WJ, Kleibeuker JH et al (2014) Predictors of colorectal neoplasia after polypectomy: based on initial and consecutive findings. Neth J Med 72:139-45.

34. Jang ES, Kim JW, Jung YJ, Jeong JB, Kim BG et al (2013) Clinical and endoscopic predictors of colorectal adenoma recurrence after colon polypectomy. Turk J Gastroenterol 24:476-82.

35. Brenner H, Chang-Claude J, Jansen L, Seiler CM, Hoffmeister M (2012) Role of colonoscopy and polyp characteristics in colorectal cancer after colonoscopic polyp detection: a population-based case-control study. Ann Intern Med 157:225-32.

36. Cottet V, Jooste V, Fournel I, Bouvier AM, Faivre J et al (2012) Long-term risk of colorectal cancer after adenoma removal: a population-based cohort study. Gut 61:1180-6.

37. de Jonge V, Sint Nicolaas J, van Leerdam ME, Kuipers EJ, Veldhuyzen van Zanten SJ (2011) Systematic literature review and pooled analyses of risk factors for finding adenomas at surveillance colonoscopy. Endoscopy 43:560-72.

38. Nusko G, Hahn EG, Mansmann U (2008) Risk of advanced metachronous colorectal adenoma during long-term follow-up. Int J Colorectal Dis 23:1065-71.

39. Bonithon-Kopp C, Piard F, Fenger C, Cabeza E, O'Morain C et al (2004) Colorectal adenoma characteristics as predictors of recurrence. Dis Colon Rectum 47:323-33.

40. Bertario L, Russo A, Sala P, Pizzetti P, Ballardini G et al (2003) Predictors of metachronous colorectal neoplasms in sporadic adenoma patients. Int J Cancer 105:82-7.

41. Nusko G, Mansmann U, Kirchner T, Hahn EG (2002) Risk related surveillance following colorectal polypectomy. Gut 51:424-8.

42. Gschwantler M, Kriwanek S, Langner E, Goritzer B, Schrutka-Kolbl C et al (2002) High-grade dysplasia and invasive carcinoma in colorectal adenomas: a multivariate analysis of the impact of adenoma and patient characteristics. Eur J Gastroenterol Hepatol 14:183-8.

43. Bertario L, Russo A, Sala P, Pizzetti P, Ballardini G et al (1999) Risk of colorectal cancer following colonoscopic polypectomy. Tumori 85:157-62.

44. Yang G, Zheng W, Sun QR, Shu XO, Li WD et al (1998) Pathologic features of initial adenomas as predictors for metachronous adenomas of the rectum. J Natl Cancer Inst 90:1661-5.

45. Triantafyllou K, Papatheodoridis GV, Paspatis GA, Vasilakaki TH, Elemenoglou I et al (1997) Predictors of the early development of advanced metachronous colon adenomas. Hepatogastroenterology 44:533-8.

46. Corley DA, Jensen CD, Marks AR, Zhao WK, Lee JK et al (2014) Adenoma detection rate and risk of colorectal cancer and death. N Engl J Med 370:1298-306.

47. Kaminski MF, Regula J, Kraszewska E, Polkowski M, Wojciechowska U et al (2010) Quality indicators for colonoscopy and the risk of interval cancer. N Engl J Med 362:1795-803.

48. Leggett BA, Hewett DG (2015) Colorectal cancer screening. Intern Med J 45:6-15.

49. Imperiale TF, Ransohoff DF, Itzkowitz SH, Levin TR, Lavin P et al (2014) Multitarget stool DNA testing for colorectal-cancer screening. N Engl J Med 370:1287-97.

Figure Legend

Figure. ROC curve for the predictive model using validation data. This graph demonstrates
the potential advantage of using a continuous predictive model for risk stratification. Cut-points
for probability of metachronous advanced neoplasia, which were estimated by the predictive
model, can be identified to improve either sensitivity or specificity over the USMSTF guidelines.
The red dot represents the sensitivity and specificity estimated by the USMSTF guidelines in
the validation data. **The blue dot** represents the estimated sensitivity and specificity in the
validation data corresponding to the cut-point targeted for improved sensitivity over USMSTF
guidelines. **The green dot** represents the estimated sensitivity and specificity in the validation
data corresponding to the cut-point identified in the test data targeted to improve specificity over
the USMSTF guidelines.

TABLES

Table 1: Glossary of Risk Stratification Terms[49]
Sensitivity (true positive rate): proportion of individuals with metachronous advanced neoplasia at follow-up who was classified as "high risk" at baseline by classification strategy (i.e. predictive model or guidelines).
Specificity (true negative rate): proportion of individuals without metachronous advanced neoplasia at follow-up who was classified as "low risk" at baseline by classification strategy
Positive predictive value (PPV): proportion of individuals with metachronous advanced neoplasia at follow-up among those who were classified by a strategy as 'high risk' at baseline.
Negative predictive value (NPV): proportion of individuals without metachronous advanced neoplasia at follow-up among those who were classified by a strategy as 'low risk' at baseline.
Surveillance colonoscopy overuse rate (1-PPV): proportion of individuals without metachronous advanced neoplasia at follow-up among those who were classified by a strategy as 'high risk' at baseline.
Surveillance colonoscopy underuse rate (1-NPV): proportion of individuals with metachronous advanced neoplasia at follow-up among those who were classified by a strategy as 'low risk' at baseline.

Table 2: Patient characteristics for the study cohort, stratified by training vs. validation dataset assignment

	Total (n=8228) n (%)	Training (n=5483) n (%)	Validation (n=2745) n (%)	P-value[a]
Demographics				
Age (year)				
Mean (SD)	62.1 (9.42)	62.1 (9.43)	62.0 (9.41)	0.42
BMI				
Mean (SD)	27.5 (4.43)	27.5 (4.42)	27.6 (4.47)	0.60
Sex				
Female	2374 (28.9)	1566 (28.6)	808 (29.4)	0.42
Male	5845 (71.1)	3917 (71.4)	1937 (70.6)	0.42
Race/ethnicity				
White	7316 (88.9)	4870 (88.8)	2466 (89.1)	0.91
African American	461 (5.6)	311 (5.7)	150 (5.5)	
Others	451 (5.5)	302 (5.5)	149 (5.4)	
Family history of colorectal cancer	5818 (70.7)	3840 (70.0)	1978 (72.1)	0.04
No				
Yes	1879 (22.8)	1265 (23.1)	614 (22.4)	
Unknown	531 (6.5)	378 (6.9)	153 (5.6)	
Cigarette smoker				
Never	2805 (34.3)	1906 (34.9)	899 (32.9)	0.18
Former	4081 (49.9)	2695 (49.4)	1386 (50.8)	
Current	1299 (15.9)	853 (15.6)	446 (16.3)	
Personal history of previous polyp				
No	4447 (54)	2967 (54.1)	1480 (53.9)	0.18
Yes	2057 (25)	1342 (24.5)	715 (26)	
Unknown	1724 (21)	1174 (21.4)	550 (20)	
Adenoma characteristics				
Number (count)				
Mean (SD)	1.73 (1.28)	1.73 (1.26)	1.74 (1.33)	0.85
Size of largest (mm)				
Mean (SD)	8.14 (6.20)	8.23 (6.30)	7.95 (6.00)	0.02
Location [b]				
Distal colorectum	3867 (47)	2578 (47)	1289 (47)	0.31
Proximal only	2482 (30.2)	1657 (30.2)	825 (30.1)	
Proximal distal	1553 (18.9)	1017 (18.5)	536 (19.5)	
Unknown	326 (4.0)	231 (4.2)	95 (3.5)	
Histology				
Tubular	5576 (67.8)	3729 (68)	1847 (67.3)	0.75
Tubullovillous or villous	1728 (21)	1147 (20.9)	581 (21.2)	
Unknown and others	924 (11.2)	607 (11.1)	317 (11.5)	
High-grade dysplasia				
No	5630 (68.4)	3747 (68.3)	1883 (68.6)	0.91
Yes	599 (7.3)	404 (7.4)	195 (7.1)	
Unknown[c]	1999 (24.3)	1332 (24.3)	667 (24.3)	
Risk of advanced adenoma[d]				0.53
High risk	2477 (30.1)	1635 (29.8)	842 (30.7)	
Low risk	3939 (47.9)	2649 (48.3)	1290 (47.0)	
Unknown[c]	1812 (22.0)	1199 (21.9)	613 (22.3)	

[a] P-value is for comparing the distribution of variables between training and validation datasets.

[b] Adenoma location is defined as proximal (cecum, ascending colon, hepatic flexure, transverse colon, and splenic flexure) or distal (descending colon, sigmoid colon, and rectum)

[c] Presence of high-grade dysplasia was not characterized by AFT and CPPS studies, this variable was characterized as unknown for patients from these two studies.

[d] Risk of advanced adenoma was classified using USMSTF guidelines, subjects were classified as *unknown* if missing adenoma size, adenoma number, histology or high-grade dysplasia.

Table 3: Factors associated with metachronous advanced neoplasia, training data

	Unadjusted[a]		Adjusted[b]	
	OR (95% CI)	P-value	OR (95% CI)	P-value
Demographics				
Age (per year)	1.04 (1.03-1.05)	<0.001	1.03 (1.02-1.04)	<0.001
BMI	1.01 (0.99-1.03)	0.17		
Sex				
Female	ref			
Male	1.24 (1.03-1.50)	0.02		
Race/ethnicity				
White	ref			
African American	1.00 (0.71-1.42)	0.99		
Other	0.88 (0.61-1.28)	0.51		
Family history of colorectal cancer				
No	ref			
Yes	1.04 (0.86-1.26)	0.68		
Unknown	1.07 (0.78-1.47)	0.68		
Cigarette smoker				
Never	ref			
Former	1.10 (0.92-1.31)	0.31		
Current	0.97 (0.75-1.25)	0.80		
Personal history of previous polyp				
No	ref		ref	
Yes	1.45 (1.21-1.75)	<0.001	1.50 (1.23-1.84)	0.001
Unknown	0.85 (0.68-1.07)	0.16	0.90 (0.71-1.14)	0.38
Adenoma characteristics				
Number (per adenoma)	1.25 (1.19-1.32)	<0.001	1.13 (1.06-1.21)	<0.001
Size (per mm)	1.05 (1.04-1.06)	<0.001	1.03 (1.02-1.04)	<0.001
Location				
Distal colorectum	ref		ref	
Proximal only	1.44 (1.19-1.76)	<0.001	1.41(1.14-1.74)	0.002
Proximal distal	2.25 (1.83-2.76)	<0.001	1.51 (1.17-1.95)	0.002
Unknown	1.08 (0.69-1.70)	0.73	0.96 (0.58-1.61)	0.88
Histology				
Tubular	ref		ref	
Tubullovillous or villous	2.11 (1.76-2.54)	<0.001	1.63 (1.32-2.00)	<0.001
Unknown/other	1.31 (1.01-1.71)	0.04	1.09 (0.83-1.45)	0.53
High-grade dysplasia				
No	ref			
Yes	1.81 (1.38-2.36)	<0.001		
Unknown CPPS	1.23 (0.96-1.58)	0.11		
Unknown AFT	0.85 (0.65-1.10)	0.22		

[a] Unadjusted OR from simple logistic regression model.
[b] Adjusted OR from the best selected multivariable logistic regression model. Note that the ORs are not available for risk factors that were not selected in the best model. Ref, reference group

Table 4: Net clinical benefit of the novel model vs. USMSTF guidelines using model cut points targeted to improve a) sensitivity, and b) specificity, applied to the validation dataset

Model Cut Point Strategy	Measure of Clinical Benefit	Validation Dataset Results		
		USMSTF Guidelines (%)	Predictive Model (%)	Difference (95% CI)[a]
Improve sensitivity compared to USMSTF guidelines				
Specified sensitivity within training data: 88%	Sensitivity	74.0	88.9	14.9 (9.5, 20.2)***
Resulting cut point = 0.075	Specificity	42.1	27.7	-14.4 (-16.7, -11.9)***
	Overuse	*83.2*	*83.7*	0.5 (-0.6,1.6)
	Underuse	*8.9*	*6.0*	-2.9 (-5.0, -0.9)[†]
Improve specificity compared to USMSTF guidelines				
Specified specificity within training data = 51%	Sensitivity	74.0	75.8	1.8 (-4.2, 7.4)
Resulting cut-point = 0.101	Specificity	42.1	46.2	4.1 (2.0, 6.7) ***
	Overuse	*83.2*	*81.8*	-1.4 (-2.7, -0.1) [†]
	Underuse	*8.9*	*7.7*	-1.2 (-3.0, 0.6)

[a] Difference is the estimate using predictive model minus the estimate using US guidelines and 95% CI is based on 1000 bootstrapped samples.
*P-value<0.05; **P-value <0.01; ***P-value <0.001; [†]Statistically significant at alpha =0.05 based on 95% bootstrap CI.

Table S1. Goodness-of-fit test results on validation data: predicted advanced neoplasia using predictive model vs. observed advanced neoplasia

Risk group	Total	Predicted AN (n)	Observed AN (n)	Predicted to observed ratio
1	212	10	10	1.00
2	211	13.2	14	0.94
3	211	15.8	20	0.79
4	212	18.9	17	1.11
5	211	21.8	32	0.68
6	211	24.8	24	1.03
7	212	29.1	31	0.94
8	211	34.2	36	0.95
9	211	42.5	50	0.85
10	211	64.5	55	1.17

Goodness-of-fit test statistic = 10.79 (based on deciles of risk); degrees of freedom = 8; $p = 0.214$.

Statement of Interests

1. Author's declaration of personal interests:
Relevant to this work, the Authors have no personal interests to declare.

2. Declaration of funding interests.