

Risk factors and prediction algorithm for advanced neoplasia on screening colonoscopy for average-risk individuals

Offir Ukashi^{ID}, Barak Pflantzer, Yiftach Barash, Eyal Klang, Shlomo Segev, Doron Yablecovitch, Uri Kopylov^{ID}, Shomron Ben-Horin and Ido Laish

Ther Adv Gastroenterol

2022, Vol. 15: 1–11

DOI: 10.1177/
17562848221101291

© The Author(s), 2022.
Article reuse guidelines:
sagepub.com/journals-
permissions

Abstract

Background: Screening with colonoscopy for all average-risk population is probably not cost-effective due to the limited sources and over-generalization of the risk, and risk stratification can be used to optimize colorectal cancer screening.

Objectives: We aimed to assess risk factors for advanced neoplasia (AN) and a classification tree algorithm to predict the risk.

Design: This is a retrospective cross-sectional study.

Methods: This study was composed of consecutive asymptomatic average-risk individuals undergoing first screening colonoscopy between 2008 and 2019. Detailed characteristics including background diseases, habits, and medications were collected. We used multivariable logistic regression to investigate the associations between clinical variables and the presence of AN and built a classification algorithm to predict AN.

Results: A total of 3856 patients were included (73.2% male, median age 55). Adenoma and AN detection rate were 15.8% and 3.4%, respectively. On multivariable analysis, predictors of AN [odds ratio (OR), 95% confidence interval (CI)] were age (1.04, 1.01–1.06, $p=0.003$), male sex (2.69, 1.56–4.64, $p<0.001$), and smoking (1.97, 1.38–2.81, $p<0.001$). A classification tree algorithm showed that smoking was the most important risk factor for prediction of AN (4.9% versus 2.4%, $p<0.001$), followed by age with a cutoff value of 60 in the smokers (8.4% versus 3.8%, $p=0.001$) and 50 in the non-smokers (2.9% versus 0.9%, $p=0.004$).

Conclusion: Smoking habits, old age, and male gender are highly associated with an increased risk for AN and should be incorporated in the individualized risk-assessment to adapt a screening program.

Keywords: advanced adenoma, advanced neoplasia, colorectal cancer, classification tree algorithm, screening colonoscopy

Received: 16 January 2022; revised manuscript accepted: 29 April 2022.

Introduction

Colorectal carcinoma (CRC) is the second most common cause of cancer-related death and the third most commonly detected cancer worldwide.¹ The typically slow progression from a benign pre-neoplastic lesion to a malignant carcinoma² allows for its prevention by removing precursor lesions or at least its early detection.^{3–5} Several screening programs have been implemented around the world in

average-risk population and have been found to decrease both the incidence and mortality of CRC,^{6–8} with the most commonly used modalities being colonoscopy and stool-based tests such as annual fecal immunohistochemistry testing (FIT).⁹ However, despite its proven effectiveness, the adherence rate among the eligible population in the United States is still less than 65%.^{10,11} On the other hand, screening with colonoscopy for all

Correspondence to:

Offir Ukashi

Department of
Gastroenterology, Sheba
Medical Center, Tel
Hashomer, Ha'ela 1,
Ramat gan, Israel

Department of Internal
Medicine A, Sheba Medical
Center, Tel Hashomer,
Israel

Sackler School of
Medicine, Tel-Aviv
University, Tel-Aviv, Israel
offirukashi@gmail.com

Barak Pflantzer

Department of Internal
Medicine A, Sheba Medical
Center, Tel Hashomer,
Israel

The Hebrew University of
Jerusalem, Jerusalem,
Israel

Yiftach Barash

Eyal Klang
Sackler School of
Medicine, Tel-Aviv
University, Tel-Aviv, Israel

Division of Diagnostic
Imaging, Sheba Medical
Center, Tel Hashomer,
Israel

DeepVision Lab, Sheba
Medical Center, Tel
Hashomer, Israel

Shlomo Segev

Sackler School of
Medicine, Tel-Aviv
University, Tel-Aviv, Israel

Institute of Medical
Screening, Sheba Medical
Center, Tel Hashomer,
Israel

Doron Yablecovitch

Uri Kopylov

Shomron Ben-Horin
Ido Laish

Department of
Gastroenterology, Sheba
Medical Center, Tel
Hashomer, Israel

Sackler School of
Medicine, Tel-Aviv
University, Tel-Aviv, Israel

average-risk population is probably not cost-effective due to the limited sources in most countries, the over-generalization of the risk, and the potential adverse events.¹² This calls for personalizing the risk in this population in order to identify patients with higher or lower risk for advanced neoplasia (AN) and to accordingly adjust the most cost-effective screening program.

Increased age,¹³ male sex,^{13,14} smoking history,^{15,16} meat and alcohol consumption,¹⁶ increased body mass index (BMI),^{17,18} and diabetes mellitus¹⁹ are independent risk factors to CRC, but none have been incorporated into screening guidelines. Numerous prediction models have been created using these risk factors to calculate the risk of AN formation,^{20–23} but low sensitivity precludes their sole use for screening. Moreover, only three studies excluded patients with prior colonoscopy and family history of CRC^{24–26}; both are dominant determinants that overshadow other baseline risk factors for stratification.

In addition to prediction model, a decision-tree algorithm is another model to show association of risk factors and their accumulated effect, as it can highlight hidden relationships between variables which might otherwise be overlooked. We aimed in the study to assess patient-related risk factors for AN in average-risk population undergoing screening colonoscopy and to construct a feasible decision-tree algorithm model based on these variables, in order to predict and individualize the risk.

Materials and methods

The reporting of this study conforms to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.²⁷

Study design

We conducted a retrospective cross-sectional analysis of database records for average-risk patients (without known risk factors for CRC) who entered a screening program for CRC in the Institute of Medical Screening (IMS) affiliated to Sheba Medical Center (SMC), between January 2008 and December 2019. Patients were selected consecutively. The database contained demographic and clinical data as detailed below, and the results of colonoscopy and histopathology reports. The study was approved by the local Institutional Review Board.

Study population

Patients aged 40 to 75 years, undergoing first-time screening/diagnostic colonoscopy in SMC within the program, including patients with non-specific abdominal symptoms (abdominal pain, change in bowel habits) were included. We excluded patients with a previous colonoscopy in SMC, weight loss and objective bleeding symptoms (iron deficiency anemia, overt or occult blood in stool) as indications for colonoscopy, medical history of complicated diverticulitis, personal history of inflammatory bowel disease or inherited syndrome with predisposition to CRC, and family history of CRC in first-degree relatives. Patients with failed/incomplete colonoscopy, either due to poor preparation or technical difficulties, and whenever pathology report was not available, were also excluded.

Extracted data

The following clinical data were collected from the electronic health records in the year in which the index colonoscopy was performed, based on free-text physician records, and included demographics, body mass index (BMI), current smoking and alcohol habits, physical exercise habits, background comorbidities (ischemic heart disease, hypertension, congestive heart failure, diabetes), and chronic medications use / at least 3 months [aspirin, statins, or hormone replacement therapy (HRT)]. Age (<50, 50–59, 60–69, ≥70 years) and BMI (<20, 20–25, >25 kg/m²) were further subcategorized into groups. Colonoscopy reports were evaluated for the number of polyps, their size and location (proximal or distal to the splenic flexure), as well as pathology reports.

Policy and definitions

According to our policy, all polyps were removed and sent for pathological assessment (unless lost in the lumen). Advanced adenoma (AA) was defined as adenoma that was at least 10 mm in diameter, had high-grade dysplasia, had villous or tubulovillous histologic characteristics, or any combination of these. Non-advanced adenoma (NAA) had none of the aforementioned histologic characteristics. Sessile serrated adenoma/polyp (SSA/P) were categorized separately. Carcinoma was considered when there was at least a submucosal invasion (T1). Advanced neoplasia was defined as cancer or AA (SSA/P ≥ 10 mm or traditional serrated adenoma were not included).

Inflammatory and hyperplastic polyps were considered as normal findings. When multiple polyps were detected and removed, colorectal findings were categorized on the basis of the most advanced lesion identified at screening.

Statistical analysis

Categorical variables are expressed as proportions. Continuous variables were evaluated for normal distribution and presented as median and interquartile range (IQR) or as mean and standard deviation (if normally distributed). Patients' characteristics comparisons between AN group and non-AN group were performed using two samples *t* test or Mann–Whitney *U* test for continuous variables, while chi-square test or Fisher's exact test were used for categorical variables. Multivariable logistic regression was applied to identify independent predictors to AN. Demographic factors, physical parameters, comorbidities, chronic medications, and laboratory values were included in the multivariable analysis as potential predictors. Backward selection method (likelihood ratio) was performed and $p > 0.1$ was used as criteria for variable removal. Classification and regression tree (CART)²⁸ analysis was applied in order to identify subgroups of individuals with increased risk for AN. All statistical tests were two-sided and $p < 0.05$ was considered as statistically significant. Statistical analyses were performed using SPSS software (IBM SPSS statistics for windows, version 25, IBM corp. Armonk, NY, USA 2017).

Study ethics and patient consent

This study was carried out in accordance with the ethical guidelines of the Declaration of Helsinki. The study was approved by the Sheba Medical Center ethics committee. Approval was granted for Helsinki protocol SMC-6987-20 on 26 May 2020. Since this was a de-identified retrospective analysis, no informed consent was obtained.

Results

Patient characteristics

Our cohort was composed of 4799 patients who were followed up by our institutional IMS in the study period. After exclusion of patients with previous colonoscopy ($n=510$), failed colonoscopy ($n=103$), missed pathology report ($n=145$), family history of CRC ($n=129$), medical history of

complicated diverticulitis ($n=3$), and overt or occult bleeding as etiology for colonoscopy ($n=53$), 3856 patients were included in the final cohort (Figure 1).

Baseline characteristics of the study population are shown in Table 1. The median age of the cohort was 55 year (IQR 51–60) and almost half of them (47.9%) were in the 50–59 years age group. Male sex was more prevalent (73.2%) compared with female sex. Most of the study population were at least overweight (BMI > 25 kg/m²) and around a third of the population were smokers. Diabetes (8.4%) and hypertension (18.7%) were the most common comorbidities among the study population, while statins were the most prevalent chronic medication (8.3%). Data regarding BMI measures and exercise habits were missing for $\leq 10\%$ and it was quite balanced among AN and non-AN groups (13/129 and 403/3727, 13/129 and 361/3727, respectively).

Colonoscopy findings

The colonoscopic and pathologic findings of the cohort population are summarized in Table 2. Out of the 3856 enrolled patients, 16% had abnormal colonoscopy findings; of them, 471 patients (12.2%) had non-AA adenoma, 16 (0.4%) had serrated adenoma (4/16 were SSA/P ≥ 10 mm), 122 (3.2%) had AA, and 7 patients (0.2%) were diagnosed with CRC. Most of the adenomas (almost 61%) were located at the distal part of the colon. The ADR (adenoma detection rate) increased with advancing age groups: 12.8%, 16.9%, 21.6%, and 30.3% among patients < 50 , 50–60, 60–70 and > 70 years, respectively. Among patients with AA, 79 patients (2% of study population) had polyp larger than 1 cm, 73 of the patients (1.9%) had villous histology, and only 17 (0.4%) had high-grade dysplasia histology of the polyp. There were seven cases with CRC, of them two were diagnosed with stage 1, three with stage 2, and two with stage 3 disease.

Predictors for AN

Compared with patients without AN, patients with AN were significantly older (57 *versus* 54 years, $p = 0.001$) and had higher percentage of male sex (88.4% *versus* 72.7%, $p < 0.001$), smoking habits (53.5% *versus* 35.8%, $p < 0.001$), and ischemic heart disease (9.3% *versus* 4.5%, $p = 0.012$), while other comorbidities (e.g. HTN, DM) were

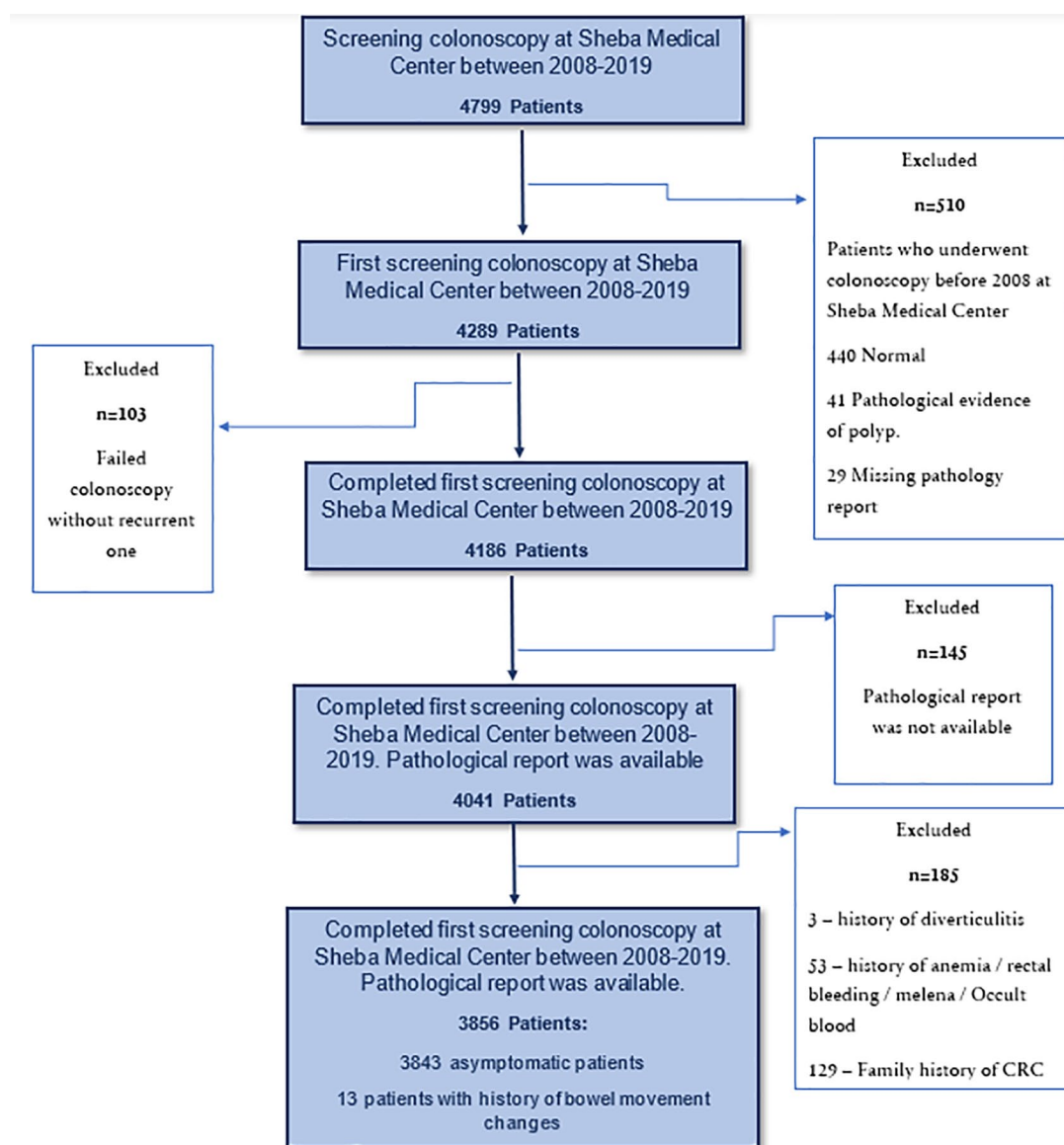


Figure 1. Study flowchart.

comparable between the groups (Table 3). Furthermore, within the group of AN, the risk further increased with advancing age (2.1%, 3.4%, 4.2%, and 5.0% in <50, 50–59, 60–69, and ≥70 years age groups, respectively) and with higher BMI (1.3%, 2.3%, and 4.0% in BMI <20, 20–25, and >25 kg/m², respectively), as depicted in Figure 2.

Using logistic regression analysis, we found age [adjusted odds ratio (AOR) 1.04, 95% confidence interval (CI) 1.01–1.06, $p = 0.003$], male sex (AOR 2.69, 95% CI 1.56–4.64, $p < 0.001$), and smoking (AOR 1.97, 95% CI 1.38–2.81, $p < 0.001$) to be independently associated with

increased risk for AN (Table 3). Performing sub-analyses across different age cutoffs, we found quite similar findings for patients ≥50 years: age (AOR 1.03, 95% CI 1.01–1.06, $p = 0.022$), male sex (AOR 2.67, 95% CI 1.55–4.61, $p < 0.001$), and smoking (AOR 1.88, 95% CI 1.31–2.69, $p = 0.001$) compared with the entire cohort. For patients ≥60 years, no variable but smoking (AOR 3.35, 95% CI 1.83–6.15, $p < 0.001$) was independently associated with higher risk for AN. Notably, we did not find neither of the chronic medications (in particular aspirin), nor other background comorbidities to be associated with those lesions.

Figure 3 depicts a classification tree algorithm that was applied to identify predictors for AN. The latter showed that smoking is the most important factor (4.9% *versus* 2.4%, $p < 0.001$), followed by age with cutoff value of 60 in the smokers (8.4% *versus* 3.8%, $p = 0.001$) and 50 in the non-smokers (2.9% *versus* 0.9%, $p = 0.004$). The model also identified male sex as a predictor for AN among smokers up to 60 years (4.6% *versus* 1.2%, $p = 0.015$).

Discussion

In this large tertiary center cohort of patients who underwent screening colonoscopy between 2008 and 2019, we found older age, male sex, and smoking to correlate with increased probability for AN. These are well established risk factors that have been used in many models of colon cancer risk.^{20–26} We also built a classification tree algorithm composed of four root nodes, predicting the likelihood to acquire AN based on the above co-variable combinations.

Numerous prediction models are available to estimate future CRC risk or current risk for AN, but all have limitations, especially lack of validation on independent populations and potential limited in generalizability because of the derived population.^{20–26} The most common in the United States is the ‘The National Cancer Institute’s (NCI’s) CRC Risk’ Assessment Tool,²⁹ which uses demographics, family history, lifestyle factors, and personal medical history to estimate future absolute CRC risk. We chose to exclude from our cohort any patients who underwent prior CRC screening and patients with a positive family history, since, as previously discussed,^{30,31} these are the most dominant variables in predicting the risk for neoplasia in follow-up colonoscopies and can overshadow other potential baseline risk factors to be implemented in risk prediction models. Only three studies previously, to the best of our knowledge, have presented risk models after excluding patients with both prior colonoscopy and family history of cancer.^{24–26} Although inconsistency regarding either enrollment of patients above the age of 20,²⁴ not in concordance with the guidelines-based recommendations for CRC screening, or excluding patients above the age of 60,²⁵ these studies have included age and smoking history, and to a lesser degree male gender, in their prediction models as well.

Table 1. Patient’s baseline characteristics.

	N = 3856
Demographics	
Age (years) median (IQR)	55 [51–60]
Age group	
<50, n (%)	899 (23.3%)
50–59, n (%)	1846 (47.9%)
60–69, n (%)	910 (23.6%)
≥70, n (%)	201 (5.2%)
Sex	
Female, n (%)	1032 (26.8%)
Male, n (%)	2824 (73.2%)
Physical parameters	
Body mass index (kg/m ²), mean (standard deviation) ^a	26.48(±3.78)
Body mass index group (kg/m ²) ^a	
<20, n (%)	78 (2.3%)
20–25, n (%)	1169 (34%)
>25, n (%)	2193 (63.7%)
Habits	
Current smoking, n (%)	1404 (36.4%)
Alcohol use, n (%)	22 (0.6%)
Exercise, n (%) ^a	2589 (74.4%)
Comorbidities	
Ischemic heart disease, n (%)	181 (4.7%)
Hypertension, n (%)	721 (18.7%)
Congestive heart failure, n (%)	1 (<1%)
Diabetes, n (%)	325 (8.4%)
Medications	
Amino salicylic acid, n (%)	191 (5%)
Statins, n (%)	321 (8.3%)
Hormone replacement therapy n (%)	34 (<1%)
IQR, interquartile range. ^a Data were missing for ≤10%.	

Table 2. Colonoscopic findings.

	N = 3856
Normal colonoscopy, <i>n</i> (%)	3240 (84%)
Non-advanced adenoma, <i>n</i> (%)	471 (12.2%)
Location	
Proximal, <i>n</i> (%)	190 (4.9%)
Distal, <i>n</i> (%)	223 (5.8%)
Proximal and distal, <i>n</i> (%)	58 (1.5%)
1–2 non-advanced adenoma, <i>n</i> (%)	429 (11.1%)
≥3 non-advanced adenoma, <i>n</i> (%)	42 (1.1%)
Serrated adenoma, <i>n</i> (%)	16 (0.4%)
Location	
Proximal, <i>n</i> (%)	6 (0.2%)
Distal, <i>n</i> (%)	10 (0.4%)
1–2 serrated adenoma, <i>n</i> (%)	14 (0.4%)
≥3 serrated adenoma, <i>n</i> (%)	2 (0.1%)
Advanced adenoma, <i>n</i> (%)	122 (3.2%)
Location	
Proximal, <i>n</i> (%)	42 (1.1%)
Distal, <i>n</i> (%)	76 (2%)
Proximal and distal, <i>n</i> (%)	4 (0.1%)
Polyp characteristics	
Size ≥ 1 cm, <i>n</i> (%)	79 (2%)
Villous histology, <i>n</i> (%)	73 (1.9%)
High-grade dysplasia, <i>n</i> (%)	17 (0.4%)
Colon adenocarcinoma, <i>n</i> (%)	7 (0.2%)
Staging	
Stage 1 (<i>n</i>)	2
Stage 2 (<i>n</i>)	3
Stage 3 (<i>n</i>)	2
Stage 4 (<i>n</i>)	-
IQR, interquartile range.	

Smoking was selected to be the root node of our model, indicating that it had a considerable influence on the probability to detect AN during screening colonoscopy. Regardless of other patient characteristics, the likelihood to acquire AN was multiplied among smokers compared with non-smokers. Cigarette smoking may affect the colorectal mucosa by means of inducing genetic and epigenetic abnormalities, resulting in precancerous lesions followed by CRC development.³² Previously published molecular epidemiologic studies have shown that smoking is associated with MSI-high, CIMP-high, and BRAF-mutation enriched CRC subtypes.³² The increased risk is highly associated with pack-years and has a reverse association with smoking duration.³³ While epigenetic changes which are expressed by DNA methylation levels may be reversible, it seems that DNA adducts that might be made by exposure to tobacco carcinogen may result in irreparable DNA damage.³² Therefore, discouragement of smoking initiation and encouragement of smoking cessation is highly recommended for CRC prevention.

Our data also support age as a risk factor for AN with an additive effect when combined with smoking; thus, smokers older than 60 years had approximately threefold (and not only twofold) increased risk for AN than non-smokers above the age of 50. Age indeed has been used in all previous published risk prediction models. Male sex was also found to be associated with AN, and this risk was emphasized in young smokers; thus, a male smoker younger than 60 years had an approximately fourfold increased risk for advanced neoplasia compared with the same-age female. Niedermaier *et al.*³⁴ showed that more than half of the excess risk of AN among men compared with women cannot be explained by traditional modifiable risk factors, such as BMI, physical activity, smoking, or diet. Further investigation should be considered in order to clarify this issue.

We did not find any reverse association between chronic medications such as aspirin, HRT, and statins use and the risk for AN. While data regarding aspirin and HRT are well established,^{32,35} the protective effect of statins use on the risk of AN is still controversial.³⁵ Actually, a part of smoking, we did not observe any association between other

Table 3. Univariable and multivariable analyses for advanced neoplasia (AN).

	Patients with AN	Patients without AN	Univariable analysis	Multivariable analysis	
	<i>n</i> = 129	<i>n</i> = 3727	<i>p</i> value	AOR (95% CI)	<i>p</i> value
Demographics					
Age (years), median (IQR)	57 (52–62)	54 (51–60)	0.001	1.04 (1.01–1.06)	0.003
Age group, <i>n</i> (%)					
<0	19 (15%)	880 (23.5%)			
50–59	62 (48%)	1784 (48%)	0.073		
60–69	38 (29.5%)	872 (23.5%)	0.014		
≥70	10 (7.5%)	191 (5%)	0.026		
Sex					
Male, <i>n</i> (%)	114 (88.4%)	2710 (72.7%)	<0.001	2.69 (1.56–4.64)	<0.001
Physical parameters					
Body mass index (kg/m ²) mean (standard deviation) ^a	27.10 (±3.70)	26.46 (±3.78)	0.077		
Body mass index group (kg/m ²) ^a					
<20, <i>n</i> (%)	1 (<1%)	77 (2.1%)			
20–25, <i>n</i> (%)	27 (20.9%)	1142 (30.6%)	0.559		
>25, <i>n</i> (%)	88 (68.2%)	2105 (56.4%)	0.248		
Habits					
Smoking, <i>n</i> (%)	69 (53.5%)	1335 (35.8%)	<0.001	1.97 (1.38–2.81)	<0.001
Alcohol use, <i>n</i> (%)	1 (<1%)	21 (<1%)	0.754		
Exercise, <i>n</i> (%) ^a	79 (68.1%)	2510 (74.6%)	0.117		
Comorbidities					
Ischemic heart disease, <i>n</i> (%)	12 (9.3%)	169 (4.5%)	0.012		
Hypertension, <i>n</i> (%)	29 (22.5%)	692 (18.6%)	0.262		
Congestive heart failure, <i>n</i> (%)	0 (0%)	1 (<1%)	0.853		
Diabetes, <i>n</i> (%)	12 (9.3%)	313 (8.4%)	0.716		
Medications					
Amino salicylic acid, <i>n</i> (%)	7 (5.4%)	184 (4.9%)	0.801		
Statins, <i>n</i> (%)	1 (<1%)	33 (<1%)	0.464		
Hormone replacement therapy, <i>n</i> (%)	13 (10%)	308 (8.3%)	0.895		
AOR, adjusted odds ratio; CI, confidence interval; IQR, interquartile range. ^a Data were missing for ≤10%.					

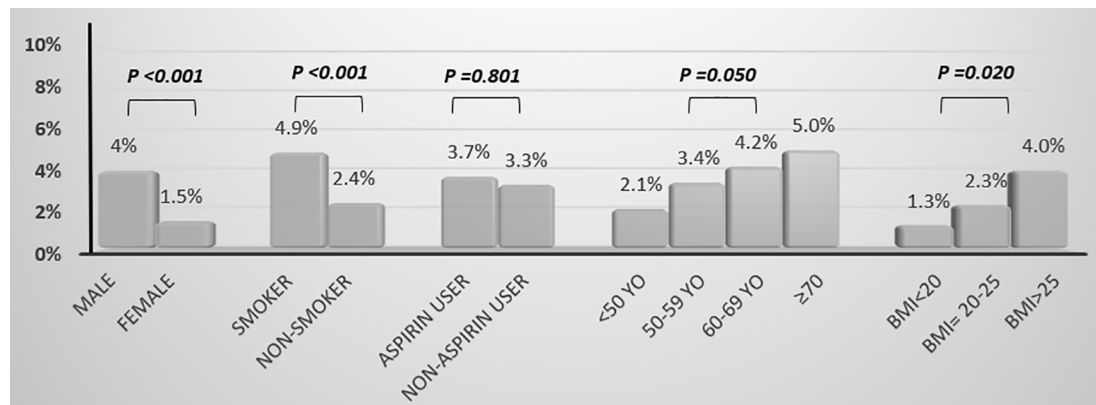


Figure 2. Advanced neoplasia rate among the cohort population, divided to sex, smoking status, aspirin use, age groups (years), and BMI (body mass index) groups (kg/m²).

background comorbidities (e.g. diabetes, congestive heart failure) or lifestyle habits (e.g. alcohol use or exercise) and the risk for AN as independent risk factors, although BMI was associated with univariable analysis. Possible explanation might be the low frequency of these risk factors (medication use, diabetes) and the low incidence of AA detection rate in our cohort, as discussed below. It is thus plausible that in a larger cohort size, with higher number of patients with these risk factors and with AN, more dominant risk factors (age, gender, or smoking habits) will not overshadow more modest ones.

The low prevalence of ADR and AA detection rate in our cohort group (15.8% and 3.2%, respectively) needs to be discussed in light of the accepted rate in the literature in screening populations (25–30% and 4–5%, respectively).^{29,36} This difference may be partially explained by the wide disparity of ages among our cohort patients. We thus observed increasing rates of ADR and AN detection rate with increasing age groups, and they were indeed in the range of that reported in the literature in older (>60 years) age groups (21.6–30.3% and 4.2–5%, respectively). Moreover, 145 patients with polyps were excluded from the cohort (Figure 1) since they had no available pathology report following polyp resection. Since adenoma is known to be the most common polyp, these patients would have raised the ADR by another 2–3% had their pathological reports been available. Third, the use of high-definition colonoscopies was not available until 2015 in our institution, potentially limiting the polyp detection rate in previous years. This may be consistent with the finding of the observed

low serrated polyp detection rate. Finally, data regarding prior colonoscopy out of SMC was not available, thus patients might have undergone screening colonoscopy before the considered index colonoscopy at our institute. If patients indeed had normal or NAA's findings at previous colonoscopy, this could have caused a lower observed ADR in our cohort relative to 'pure' first-time screening population.

Although colonoscopy is recommended worldwide for CRC screening in average-risk population, most countries in the world use FIT as the preferred screening modality, mainly due to limited resources and low adherence rate for colonoscopy.^{9,37} The use of prediction model or algorithm can be used to individualize the risk in order to identify patients with higher or lower risk for AN and to accordingly adjust the most cost-effective screening modality. Our classification tree algorithm shows (Figure 3) that smokers, especially male or above the age of 60, have a significant risk for AN (up to 8.4%), thus probably prompting a colonoscopy as the preferred screening modality, while non-smokers, or even young (<60 years) female smokers, have only 1.2–2.4% risk for AN, potentially prompting FIT as the more cost-effective screening modality. In addition to identifying patients with lower or higher risk, such an algorithm can be used as a shared decision-making tool between patient and physician, prompting improved screening compliance.

Our study has several limitations. First, the relatively low AA detection rate precluded our ability to possibly expose more statistically significant

risk factors for AN and include them in the classification tree algorithm. Second, the retrospective design of the study, with its inherent drawback of missing (e.g. pathology reports) or incomplete data, had either limited our ability to include other potentially important risk factors in the era of lifestyle habits (e.g. red meat and alcohol use, physical exercise), or could have biased, most probably underestimating, the observed ADR and AA detection rate. Third, this study was conducted at a single tertiary large medical center in high-socioeconomic status patients who underwent a structured screening program, which can limit the generalization of our findings, especially for more low-socioeconomic status population with potentially increased likeliness for lifestyle-related risk factors and a consequent ADR. Fourth, as mentioned above, the unavailable data regarding prior colonoscopy out of SMC could have caused an ascertainment bias toward either underestimation or overestimation of the observed ADR/AA detection rate.

However, at the same time, this structured and organized design has enabled us to process a high-quality and credible data regarding risk factors for AN in average-risk population, apparently without a significant bias of other overshadowing and dominant risk factors of family history or previous colonoscopy. We have shown that smoking habits, age, and male sex are highly associated with an increased likelihood for AN, and these factors should be incorporated to individualize the risk and, in the decision making of adapting a screening program.

Ethics approval and consent to participate

This study was carried out in accordance with the ethical guidelines of the Declaration of Helsinki. The study was approved by the Sheba Medical Center ethics committee. Approval was granted for Helsinki protocol SMC-6987-20 on 26 May 2020. Since this was a de-identified retrospective analysis, no informed consent was obtained.

Consent for publication

Not applicable.

Author contribution(s)

Offir Ukashi: Conceptualization; Data curation; Writing – original draft.

Barak Pflantzer: Data curation; Writing – review & editing.

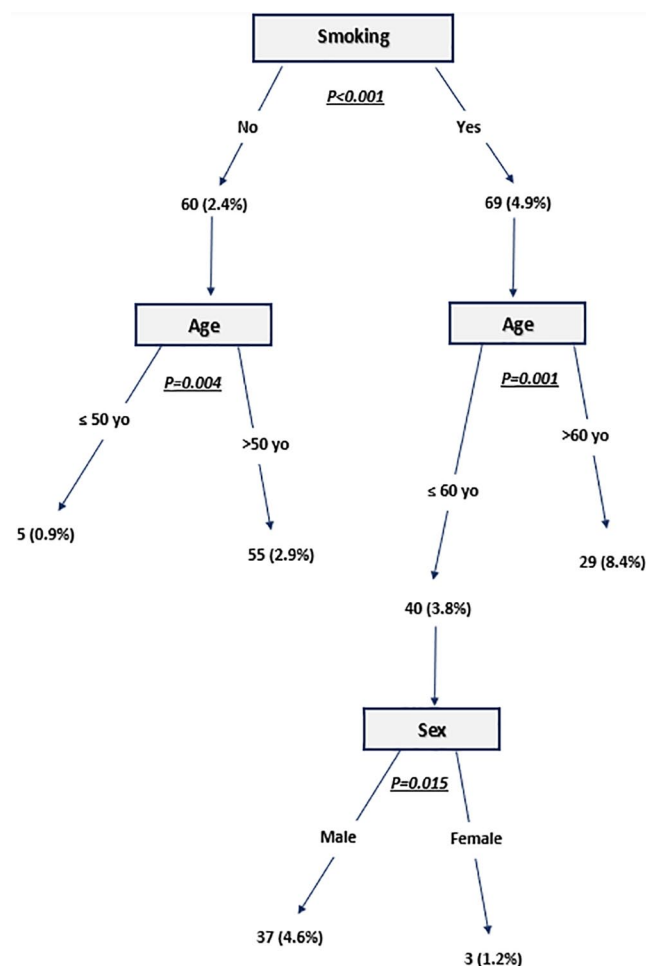


Figure 3. Classification tree algorithm for advanced neoplasia detection among average-risk population.

Yiftach Barash: Data curation; Writing – review & editing.

Eyal Klang: Data curation; Writing – review & editing.

Shlomo Segev: Methodology; Writing – review & editing.

Doron Yablecovitch: Methodology; Writing – review & editing.

Uri Kopylov: Methodology; Writing – review & editing.

Shomron Ben-Horin: Methodology; Writing – review & editing.

Ido Laish: Conceptualization; Methodology; Writing – original draft.

ORCID iDs

Offir Ukashi  <https://orcid.org/0000-0002-8601-6550>

Uri Kopylov  <https://orcid.org/0000-0002-7156-0588>

Acknowledgements

Not applicable.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest statement

The authors declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: UK: speaker and advisory fees – Abbvie, Janssen Takeda Medtronic; research support – Janssen Takeda Medtronic. SBH: consulting and advisory board fees and / or research support – Abbvie, MSD, Janssen, Takeda, and CellTrion. The remaining authors declare that there is no conflict of interest.

Availability of data and materials

Not applicable.

References

1. Sung H, Ferlay J, Siegel RL, *et al.* Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209–249.
2. Grady WM and Markowitz SD. The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening. *Dig Dis Sci* 2015; 60: 762–772.
3. Kahi CJ. Reviewing the evidence that polypectomy prevents cancer. *Gastrointest Endosc Clin N Am* 2019; 29: 577–585.
4. Brenner H, Chang-Claude J, Seiler CM, *et al.* Protection from colorectal cancer after colonoscopy: a population-based, case-control study. *Ann Intern Med* 2011; 154: 22–30.
5. Zauber AG, Winawer SJ, O'Brien MJ, *et al.* Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* 2012; 366: 687–696.
6. Levin TR, Corley DA, Jensen CD, *et al.* Effects of organized colorectal cancer screening on cancer incidence and mortality in a large community-based population. *Gastroenterology* 2018; 155: 1383–1391.e5.
7. Brenner H, Stock C and Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; 348: g2467.
8. Kahi CJ, Imperiale TF, Juliar BE, *et al.* Effect of screening colonoscopy on colorectal cancer incidence and mortality. *Clin Gastroenterol Hepatol* 2009; 7: 770–775; quiz 711.
9. Rabeneck L and Lansdorp-Vogelaar I. Assessment of a cancer screening program. *Best Pract Res Clin Gastroenterol* 2015; 29: 979–985.
10. Mankaney G, Rizk M, Sarvepalli S, *et al.* Patient-initiated colonoscopy scheduling effectively increases colorectal cancer screening adherence. *Dig Dis Sci* 2019; 64: 2497–2504.
11. Dougherty MK, Brenner AT, Crockett SD, *et al.* Evaluation of interventions intended to increase colorectal cancer screening rates in the United States: a systematic review and meta-analysis. *JAMA Intern Med* 2018; 178: 1645–1658.
12. Rutter CM, Johnson E, Miglioretti DL, *et al.* Adverse events after screening and follow-up colonoscopy. *Cancer Causes Control* 2012; 23: 289–296.
13. Regula J, Rupinski M, Kraszewska E, *et al.* Colonoscopy in colorectal-cancer screening for detection of advanced neoplasia. *N Engl J Med* 2006; 355: 1863–1872.
14. Nguyen SP, Bent S, Chen YH, *et al.* Gender as a risk factor for advanced neoplasia and colorectal cancer: a systematic review and meta-analysis. *Clin Gastroenterol Hepatol* 2009; 7: 676–681.
15. Botteri E, Iodice S, Raimondi S, *et al.* Cigarette smoking and adenomatous polyps: a meta-analysis. *Gastroenterology* 2008; 134: 388–395.
16. Steinmetz J, Spyckerelle Y, Guéguen R, *et al.* [Alcohol, tobacco and colorectal adenomas and cancer. Case-control study in a population with positive fecal occult blood tests]. *Presse Med* 2007; 36(9 Pt 1): 1174–1182.
17. Terry MB, Neugut AI, Bostick RM, *et al.* Risk factors for advanced colorectal adenomas: a pooled analysis. *Cancer Epidemiol Biomarkers Prev* 2002; 11: 622–629.

18. Ben Q, An W, Jiang Y, *et al.* Body mass index increases risk for colorectal adenomas based on meta-analysis. *Gastroenterology* 2012; 142: 762–772.
19. Larsson SC, Orsini N and Wolk A. Diabetes mellitus and risk of colorectal cancer: a meta-analysis. *J Natl Cancer Inst* 2005; 97: 1679–1687.
20. Ma GK and Ladabaum U. Personalizing colorectal cancer screening: a systematic review of models to predict risk of colorectal neoplasia. *Clin Gastroenterol Hepatol* 2014; 12: 1624–1634.e1.
21. Driver JA, Gaziano JM, Gelber RP, *et al.* Development of a risk score for colorectal cancer in men. *Am J Med* 2007; 120: 257–263.
22. Yeoh KG, Ho KY, Chiu HM, *et al.* The Asia-Pacific Colorectal Screening score: a validated tool that stratifies risk for colorectal advanced neoplasia in asymptomatic Asian subjects. *Gut* 2011; 60: 1236–1241.
23. Kaminski MF, Polkowski M, Kraszewska E, *et al.* A score to estimate the likelihood of detecting advanced colorectal neoplasia at colonoscopy. *Gut* 2014; 63: 1112–1119.
24. Hong SN, Son HJ, Choi SK, *et al.* A prediction model for advanced colorectal neoplasia in an asymptomatic screening population. *PLoS ONE* 2017; 12: e0181040.
25. Murchie B, Tandon K, Hakim S, *et al.* A new scoring system to predict the risk for high-risk adenoma and comparison of existing risk calculators. *J Clin Gastroenterol* 2017; 51: 345–351.
26. Sharara AI, El Mokahal A, Harb AH, *et al.* Risk prediction rule for advanced neoplasia on screening colonoscopy for average-risk individuals. *World J Gastroenterol* 2020; 26: 5705–5717.
27. vonElm E, Altman DG, Egger M, *et al.* The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology* 2007; 18: 800–804.
28. Breiman L, Friedman JH, Olshen RA, *et al.* CART (1st ed.). R Classification and regression trees, 1984, <https://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=1619230> (accessed 2017).
29. Imperiale TF, Wagner DR, Lin CY, *et al.* Risk of advanced proximal neoplasms in asymptomatic adults according to the distal colorectal findings. *N Engl J Med* 2000; 343: 169–174.
30. Taylor DP, Burt RW, Williams MS, *et al.* Population-based family history-specific risks for colorectal cancer: a constellation approach. *Gastroenterology* 2010; 138: 877–885.
31. Sharara AI and Harb AH. Development and validation of a scoring system to identify individuals at high risk for advanced colorectal neoplasms who should undergo colonoscopy screening. *Clin Gastroenterol Hepatol* 2014; 12: 2135–2136.
32. Keum N and Giovannucci E. Global burden of colorectal cancer: emerging trends, risk factors and prevention strategies. *Nat Rev Gastroenterol Hepatol* 2019; 16: 713–732.
33. Liang PS, Chen T-Y and Giovannucci E. Cigarette smoking and colorectal cancer incidence and mortality: systematic review and meta-analysis. *Int J Cancer* 2009; 124: 2406–2415.
34. Niedermaier T, Heisser T, Gies A, *et al.* To what extent is male excess risk of advanced colorectal neoplasms explained by known risk factors? Results from a large German screening population. *Int J Cancer* 2021; 149: 1877–1886.
35. Katona BW and Weiss JM. Chemoprevention of colorectal cancer. *Gastroenterology* 2020; 158: 368–388.
36. Lieberman DA, Weiss DG, Bond JH, *et al.* Use of colonoscopy to screen asymptomatic adults for colorectal cancer. Veterans Affairs Cooperative Study Group 380. *N Engl J Med* 2000; 343: 162–168.
37. Rabeneck L, Chiu HM and Senore C. International perspective on the burden of colorectal cancer and public health effects. *Gastroenterology* 2020; 158: 447–452.

Visit SAGE journals online
[journals.sagepub.com/
 home/tag](https://journals.sagepub.com/home/tag)

 SAGE journals