ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ESMO OPEN CANCER HORIZONS

ORIGINAL RESEARCH

# A user-friendly objective prediction model in predicting colorectal cancer based on 234 044 Asian adults in a prospective cohort

C. H. Chen[1,2,3], M. K. Tsai[4,5], C. Wen[6] & C. P. Wen[4,5,7*]

[1]Digestive Disease Center, Changbing Show-Chwan Memorial Hospital, Lukang; [2]Department of Food Science and Technology, Hungkuang University, Taichung; [3]Medical College Provisional Office, Chung-Hsing University, Taichung; [4]Institute of Population Health Sciences, National Health Research Institutes, Zhunan; [5]MJ Health Management Institution, Taipei, Taiwan; [6]Long Beach VAMC Hospital, University of Irvine Medical Center, Irvine, USA; [7]China Medical University Hospital, Taichung, Taiwan

Available online xxx

**Background:** Prediction models of colorectal cancer (CRC) had limited application for not being user-friendly. Whether fecal immunochemical tests (FITs) can help predict CRC has been overlooked.

**Patients and methods:** With 1972 CRCs identified, 234 044 adults aged ≥40 years were successively enrolled between 1994 and 2008. Prediction models were developed by questionnaire/medical screening and quantitative FIT. NNS (number needed to scope to find one cancer) is time dependent, spanning entire study period. Significant 'risk factors' were family history, body mass index, smoking, drinking, inactivity, hypertension, diabetes, carcinoembryonic antigen, and C-reactive protein.

**Results:** Positive FIT (≥20 μg/g hemoglobin/feces) had cancer risk 10-fold larger than negative FIT, and within each age group, another 10-fold difference. The C statistic of FIT (0.81) with age and sex alone was superior to the 'common risk-factors' model (0.73). NNS, stratified by age and by FIT values, demonstrated a scorecard of cancer risks, like 1/15 or 1/25, in 5 years. When FIT was negative, cancer risk was small (1/750-1/3000 annually). The larger the FIT, the sooner the appearance of CRC. For every 80-μg/g increase of FIT, there were 1.5-year earlier development of CRC incidence and 1-year earlier development of CRC mortality, respectively. Given the same FIT value, CRC events appeared in the proximal colon sooner than the distal colon.

**Conclusions:** A simple user-friendly model based on a single FIT value to predict CRC risk was developed. When positive, NNS offered a simple quantitative value, with a better precision than most risk factors, even combined. When FIT is negative, risk is very small, but requiring a repeat every other year to rule out false negative. FIT values correlated well with CRC prognosis, with worst for proximal CRC.

**Key words:** prediction model, colorectal cancer, fecal immunochemical tests

## INTRODUCTION

With >1.9 million new colorectal cancer (CRC) cases and 935 000 CRC-related deaths in 2020, CRC has ranked third in incidence and second in cancer-related mortality globally.[1] The incidence and mortality of CRC in the United States even increased for subjects younger than 50 years, with 17 930 cases and 3640 deaths.[2] The incidence of CRC is expected to increase with escalating Human Development Index, but these may be mitigated by appropriate screening and healthy behavior.[3] In Taiwan, the age-standardized incidence rate of CRC has increased from 25.46/100 000 persons in 2002 to 42.9/100 000 persons in 2017.[4] To reduce the large cancer burden of CRC, early detection with follow-up intervention has become a mandatory public health goal.

Prediction model is a simple tool to identify the high-risk individuals so that colonoscopy or risk reduction efforts can be implemented in a timely manner.[5] There have been numerous CRC prediction models published in the literature,[6-20] but their applications have been limited because there are no sufficiently strong risk factors identified to make prediction models worth the clinical effort.[21-23]

Fecal immunochemical test (FIT), an alternative to colonoscopy, is commonly listed as one of the two first-tier screening tests for CRC. When FIT is viewed as a screening test for CRC, it has a very high false-positive rate, >90%,[24] implying that most positive FITs do not lead to CRC. The reality is that FIT is not a stand-alone screening

*Correspondence to: Dr Chi-Pang Wen, Institute of Population Health Sciences, National Health Research Institutes, 35 Keyan Road, Zhunan, Miaoli 350, Taiwan; China Medical University Hospital, Taichung, Taiwan. Tel: +886-37-246-166*36318; Fax: +886-37-586-261

E-mail: cwengood@nhri.org.tw (C. P. Wen).

2059-7029/© 2021 The Authors. Published by Elsevier Ltd on behalf of European Society for Medical Oncology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

test but relied on identifying higher-risk individuals for colonoscopy. A positive FIT provides a measure of risks, just like a positive family history, and should be treated as such. However, most FIT used in the published models was not quantitative but dichotomous, yielding positive or negative.[25-27] The application of the quantitative FIT values in CRC prediction model has not been fully studied.

A CRC prediction model with quantitative FIT test results was developed in 234 044 individuals nearly followed up for 15 years. A step-wise inclusion of risk factors from general common history, CRC-specific history, FIT results, and additional blood indices was made in developing prediction models. The objective of the study is to assess the value of applying quantitative FIT in improving the model.

## MATERIALS AND METHODS

### Patients and ethical considerations

The cohort consisted of 234 044 participants aged ≥40 years without known cancers, 109 646 men and 124 398 women, in MJ Health Management Institution (MJ) for self-paying medical screening in Taiwan.[28] The cohort were enrolled between 1994 and 2008, and they were followed till the end of 2008. A total of 1972 cases of CRC were identified in a mean follow-up duration of 7.4 years. Informed consent from each participant and the ethical review from the Institutional Review Board at National Health Research Institutes of Taiwan have been approved.

### Data extraction and management

Each participant submitted a detailed history questionnaire before receiving physical examinations, and overnight fasting blood samples were assayed. The participant preserved an uncontaminated fecal sample in a plastic container with a liquid buffer and returned the sample within 1 day. The samples were stored in a 4°C refrigerator and were analyzed by the automated OC sensor instrument (Eiken Chemical Company, Tokyo, Japan) within a week. Positive FIT is defined as ≥20 μg/g hemoglobin/feces and negative FIT, <20 μg/g hemoglobin/feces. The exercise volume [metabolic equivalents-hours (MET-h) per week], the product of intensity (MET) of different activity and duration of exercise (h), was classified as inactive (<3.75 MET-h), low active (3.75-7.49 MET-h), and fully active (≥7.50 MET-h).[29]

Each dataset from the cohort was randomly and equally split into a training set to guide the building of the risk model and a validation set to assess the predictive performance of the models. Gender, family history of CRC (none, yes), physical activity (inactive, low, fully active), the servings of fruit and vegetable intake (low, medium, high), black color stool passage (none, occasional, always), bowel habit change (none, yes), diabetes (none, yes), hypertension (none, yes), anemia (none, mild, severe), carcinoembryonic antigen (CEA) (<5 ng/ml, ≥5 ng/ml), C-reactive protein (CRP) (<0.5 mg/dl, ≥0.5 mg/dl), smoking status (none, <30 pack-year, ≥30 pack-year), drinking status (none, <100 g/week, ≥100 g/week), age (40-49 years, 50-59 years, 60-69 years, ≥70 years), and body mass index (BMI) (<18.5 kg/m$^2$, 18.5-24.9 kg/m$^2$, ≥25 kg/m$^2$) were classified into categorical variables. The cut-off points for FIT were chosen by setting the reference group at different starting points through multiple iterations (Supplementary Appendix Table S1, available at https://doi.org/10.1016/j.esmoop.2021.100288).

### Outcome

Every individual's identification number was matched with the National Death File for CRC mortality and the National Cancer Registry file for CRC incidence, respectively.[30,31] Instead of immediate colonoscopy, a comprehensive cancer registry system would be more appropriate to develop a prediction model for expected CRC incidence in 5-10 years. Furthermore, no colonoscopy was indicated for negative FIT, and cancer registry would be the standard for CRC incidence. Finally, the National Death File-based mortality would be the ultimate criteria in evaluating CRC mortality. The time interval from FIT to the development of CRC incidence or mortality would be measured once CRC occurred.

A step-wise Cox proportional hazards regression analysis was carried out to identify risk predictors that were significantly associated with increased risk of CRC in multivariable models. We included physical inactivity and low intake of fruit and vegetables as risk factors of CRC since its risk approximated statistical significance. Hazard ratios (HRs) and 95% confidence interval (CI) were estimated for each variable. NNS (number needed to scope to find one cancer during study period) used in this study was approximately a reciprocal of age-adjusted incidence rate/1000 person-years.

### Exploratory analysis

The modeling started with CRC-nonspecific questionnaire (model 1), followed by combination of CRC-specific questionnaire (model 2). Model 3 started with FIT values (age, gender, and FIT). Model 4 was extended with the combination of all questionnaire and FIT values, and model 5 was composed of questionnaire, blood tests, and FIT.

### Statistical methods

A receiver operating characteristic curve for censored survival data and the area under the curve (AUC) were constructed to assess the discriminatory accuracy of full dataset, training set, and internal validation set for predicting the development of CRC.[32] We assessed internal calibration by determining the agreement between estimated and observed events in 5 years and 10 years, respectively.[33] Moreover, we also presented the slope of the ratio between estimated events and observed events (E/O ratio) in each model.[34] The 5- and 10-year absolute risks were calculated from baseline probability and relative risk profile by the Cox proportional hazards regression model, using the standard equation for survival data with censored observations.[35] The risk scores for each predictor

were derived from regression coefficients in the Cox proportional hazards regression model following the reported procedures.[36]

After developing the prediction models, we also calculated the Akaike information criterion (AIC) which takes into account both the statistical goodness of fit and the simplicity of the model to evaluate the accuracy of the model complexity.[37,38] With the lower AIC values, the model will have higher predictive accuracy.

All statistical tests were two-sided, and a $P$ value $<0.05$ was considered statistically significant. Stata 10.0 (Stata Corp, College Station, TX) and SAS 9.2 (SAS Institute Inc, Cary, NC) were used for statistical analyses and modeling.

## RESULTS

Table 1 showed the positive risk predictors of CRC that were statistically significant by univariate analysis in five models. All risk-factor likely variables were screened for possible inclusion as significant risk factors by controlling for age and gender only (Supplementary Appendix Table S1, available at https://doi.org/10.1016/j.esmoop.2021.100288).

### *Risk of CRC without considering FIT values*

Age and male sex were associated with increased risk of CRC in all the two models without considering FIT values. In model 1, the other positive predictors for CRC were smoking $\geq$30 pack-year, drinking status $<$100 g/week or $\geq$100 g/week, and BMI $\geq$25 kg/m$^2$. In model 2, the additional positive predictors for CRC were positive family history of CRC and bowel habit change.

### *Risk of CRC incorporating FIT values*

Model 3 (age, gender, and FIT) was developed with considering FIT values, and increasing FIT values were associated with increasing CRC risk in a dose-dependent manner by age (Figure 1A). Age and male sex were persistently associated with increased risk of CRC.

Figure 1B showed cancer risk of 5 years expressed in NNS. When FIT was positive, NNS by age and by FIT values could be directly converted into 5-year risk. For example, at FIT values at 80 μg/g, NNS was 15 for age 60-69 years and 25 for 50-59 years, implying the cancer risk being 1/15 and 1/25 in 5 years, respectively.

In model 4, the other positive predictors for CRC were positive family history of CRC, always black color stool passage, bowel habit change, and low intake of fruit and vegetables. In model 5, the other positive predictors for CRC were positive family history of CRC, always black color stool passage, bowel habit change, hypertension, anemia, CEA $\geq$5 ng/ml, CRP $\geq$0.5 mg/dl, and low intake of fruit and vegetables. The HRs of CRC consistently increased with the increment of FIT values after age stratification throughout models 3-5.

### *Risk score assignments and absolute risk of CRC*

The risk score was assigned as integer points to each risk level and calculated as a weighted distance from each level to the reference level of that particular risk factor (Supplementary Appendix Table S2, available at https://doi.org/10.1016/j.esmoop.2021.100288). For example, in model 1, the reference level of age (40-49 years) was assigned a risk score of zero, and increasing scores were assigned to increasing levels of age (age 50-59 years, score = 3; 60-69 years, score = 5; age $\geq$70 years, score = 6).

### *The goodness of fit of the risk prediction models*

We examined the C statistic based on 10-year prediction of CRC incidence, and the C statistic in each model was comparable for full dataset, training set, and internal validation set (Table 2). Furthermore, we found the accuracy evaluated by AIC was consistent with the C statistic based on AUC. Compared with model 1, the increase of C statistic was not significant in model 2 after adding specific CRC questionnaire. However, the C statistic was significantly increased in model 3 after considering age, gender, and FIT.

Compared with model 3, the C statistic could be increased in model 4 after adding all questionnaire and full model 5 after adding all questionnaire and blood test, respectively. However, it was of interest to note that the C statistic after adding blood test or questionnaire was not significantly enhanced after considering FIT.

### *Risk analysis based on CRC location*

It was noted that CRC in the distal colon, rather than CRC in the proximal colon, had relatively higher C statistic throughout FIT-based models 3-5. However, the rectal cancer had the lowest C statistic. Furthermore, the C statistic would markedly diminish from 0.812 for model 3 to 0.713 for FIT-only model without considering age and gender.

A larger FIT implied an earlier development of total CRC case as well as total CRC mortality, with 1.5-year earlier development of CRC incidence and 1-year earlier development of CRC mortality for a difference of 80 μg/g of FIT (Figure 2). With the same FIT value, we found that the latency for incidence or mortality was relatively shorter in the proximal colon cancer.

### *AUC of CRC in the five models from the full dataset*

The AUC was 0.72 for model 1, 0.73 for model 2, 0.81 for model 3, 0.82 for model 4, and 0.83 for model 5, respectively (Supplementary Appendix Figure S1, available at https://doi.org/10.1016/j.esmoop.2021.100288). The cross-validated calibration plot by risk deciles showed excellent calibration agreement of observed events with predicted events in the 5-year (Figure 3A) and 10-year (Figure 3B) timeframes across models.

### *Application of risk score and predictive power of FIT*

Through these models, we used 10 hypothetical examples to predict probability of developing CRC in 5 years

**Table 1. Risk factors identified for colorectal cancer incidence in cohort**

| | *N* | % | *n* of incidence | % | HR | 95% CI |
|---|---|---|---|---|---|---|
| Nonspecific questionnaire (model 1) | | | | | | |
| Age, years | | | | | | |
| 40-49 | 92 062 | 39.34 | 272 | 13.79 | **1.00** | |
| 50-59 | 74 845 | 31.98 | 554 | 28.09 | **2.51** | 2.17-2.90 |
| 60-69 | 49 050 | 20.96 | 753 | 38.18 | **4.86** | 4.23-5.59 |
| ≥70 | 18 087 | 7.73 | 393 | 19.93 | **7.52** | 6.44-8.78 |
| Gender | | | | | | |
| Female | 124 398 | 53.15 | 878 | 44.52 | **1.00** | |
| Male | 109 646 | 46.85 | 1094 | 55.48 | **1.45** | 1.32-1.58 |
| Smoking status | | | | | | |
| None | 126 851 | 72.61 | 808 | 64.28 | **1.00** | |
| <30 pack-year | 30 044 | 17.20 | 175 | 13.92 | **1.15** | 0.98-1.36 |
| ≥30 pack-year | 17 814 | 10.20 | 274 | 21.80 | **1.46** | 1.27-1.67 |
| Drinking status | | | | | | |
| None | 130 604 | 73.59 | 857 | 66.33 | **1.00** | |
| <100 g/week | 38 091 | 21.46 | 352 | 27.24 | **1.18** | 1.03-1.35 |
| ≥100 g/week | 8787 | 4.95 | 83 | 6.42 | **1.42** | 1.12-1.80 |
| Physical activity | | | | | | |
| Inactive | 131 795 | 67.69 | 920 | 64.61 | **1.09** | 0.94-1.26 |
| Low | 47 827 | 24.57 | 342 | 24.02 | **1.06** | 0.95-1.20 |
| Fully active | 15 071 | 7.74 | 162 | 11.38 | **1.00** | |
| Fruit and vegetables[a] | | | | | | |
| Low | 93 875 | 49.63 | 542 | 39.71 | **1.13** | 0.95-1.35 |
| Medium | 78 424 | 41.46 | 645 | 47.25 | **1.06** | 0.90-1.26 |
| High | 16 849 | 8.91 | 178 | 13.04 | **1.00** | |
| BMI, kg/m$^2$ | | | | | | |
| <18.5 | 7185 | 3.07 | 49 | 2.49 | **0.82** | 0.61-1.09 |
| 18.5-24.9 | 140 367 | 60.07 | 1091 | 55.44 | **1.00** | |
| ≥25 | 86 128 | 36.86 | 828 | 42.07 | **1.16** | 1.06-1.27 |
| Specific questionnaire (model 2) | | | | | | |
| Family history[b] | | | | | | |
| None | 190 486 | 96.34 | 1365 | 94.40 | **1.00** | |
| Yes | 7244 | 3.66 | 81 | 5.60 | **1.99** | 1.59-2.49 |
| Black color stool[c] | | | | | | |
| None | 166 016 | 98.58 | 1194 | 97.47 | **1.00** | |
| Occasional | 2108 | 1.25 | 25 | 2.04 | **1.37** | 0.92-2.04 |
| Always | 286 | 0.17 | 6 | 0.49 | **2.45** | 1.10-5.47 |
| Bowel habit change[d] | | | | | | |
| None | 163 689 | 96.52 | 1172 | 94.59 | **1.00** | |
| Yes | 5906 | 3.48 | 67 | 5.41 | **1.73** | 1.35-2.22 |
| Blood test | | | | | | |
| Diabetes[e] | | | | | | |
| None | 210 052 | 89.94 | 1688 | 85.69 | **1.00** | |
| Yes | 23 495 | 10.06 | 282 | 14.31 | **1.25** | 1.10-1.42 |
| Hypertension[f] | | | | | | |
| None | 157 186 | 67.24 | 1025 | 26.02 | **1.00** | |
| Yes | 76 570 | 32.76 | 944 | 32.91 | **1.24** | 1.13-1.36 |
| Anemia[g] | | | | | | |
| None | 202 090 | 86.46 | 1570 | 54.74 | **1.00** | |
| Mild | 28 807 | 12.32 | 354 | 12.34 | **1.16** | 1.00-1.34 |
| Severe | 2844 | 1.22 | 46 | 1.17 | **2.25** | 1.80-2.81 |
| CEA | | | | | | |
| <5 | 214 790 | 95.28 | 1735 | 88.66 | **1.00** | |
| ≥5 | 10 650 | 4.72 | 222 | 11.34 | **1.85** | 1.61-2.14 |
| CRP | | | | | | |
| <0.5 | 182 549 | 89.47 | 1307 | 82.93 | **1.00** | |
| ≥0.5 | 21 485 | 10.53 | 269 | 17.07 | **1.39** | 1.22-1.59 |
| FIT (model 3) | | | | | | |
| Age 40-49 years | | | | | | |
| <20 | 87 286 | 94.81 | 181 | 66.54 | **1.00** | |
| 20-39 | 574 | 0.62 | 5 | 1.84 | **3.02** | 1.24-7.34 |
| 40-59 | 942 | 1.02 | 15 | 5.51 | **6.83** | 4.04-11.57 |
| 60-79 | 490 | 0.53 | 8 | 2.94 | **7.31** | 3.60-14.85 |
| 80-119 | 540 | 0.59 | 14 | 5.15 | **11.76** | 6.83-20.26 |
| 120-159 | 286 | 0.31 | 4 | 1.47 | **6.36** | 2.36-17.13 |
| 160-199 | 172 | 0.19 | 9 | 3.31 | **25.81** | 13.22-50.41 |
| 200-299 | 318 | 0.35 | 14 | 5.15 | **19.49** | 11.32-33.58 |
| 300-499 | 283 | 0.31 | 11 | 4.04 | **19.27** | 10.48-35.42 |
| ≥500 | 123 | 0.13 | 9 | 3.31 | **41.99** | 21.50-82.03 |

*Continued*

**Table 1. Continued**

| | *N* | % | *n* of incidence | % | HR | 95% CI |
|---|---|---|---|---|---|---|
| Age 50-59 years | | | | | | |
| <20 | 70 032 | 93.57 | 332 | 59.93 | **2.26** | 1.89-2.71 |
| 20-39 | 553 | 0.74 | 8 | 1.44 | **5.19** | 2.56-10.55 |
| 40-59 | 961 | 1.28 | 38 | 6.86 | **17.51** | 12.34-24.84 |
| 60-79 | 521 | 0.70 | 24 | 4.33 | **20.52** | 13.41-31.41 |
| 80-119 | 590 | 0.79 | 31 | 5.60 | **24.02** | 16.41-35.17 |
| 120-159 | 279 | 0.37 | 25 | 4.51 | **43.45** | 28.60-66.01 |
| 160-199 | 190 | 0.25 | 18 | 3.25 | **48.42** | 29.83-78.60 |
| 200-299 | 340 | 0.45 | 28 | 5.05 | **41.80** | 28.08-62.24 |
| 300-499 | 258 | 0.34 | 23 | 4.15 | **48.59** | 31.48-75.00 |
| ≥500 | 143 | 0.19 | 21 | 3.79 | **84.55** | 53.79-132.90 |
| Age 60-69 years | | | | | | |
| <20 | 45 195 | 92.14 | 425 | 56.44 | **4.21** | 3.54-5.01 |
| 20-39 | 544 | 1.11 | 21 | 2.79 | **14.10** | 8.97-22.15 |
| 40-59 | 772 | 1.57 | 47 | 6.24 | **26.70** | 19.37-36.81 |
| 60-79 | 452 | 0.92 | 29 | 3.85 | **29.21** | 19.73-43.23 |
| 80-119 | 463 | 0.94 | 40 | 5.31 | **41.55** | 29.50-58.53 |
| 120-159 | 260 | 0.53 | 38 | 5.05 | **73.78** | 52.00-104.68 |
| 160-199 | 153 | 0.31 | 30 | 3.98 | **104.00** | 70.66-153.06 |
| 200-299 | 313 | 0.64 | 53 | 7.04 | **83.81** | 61.69-113.88 |
| 300-499 | 207 | 0.42 | 39 | 5.18 | **105.47** | 74.59-149.15 |
| ≥500 | 115 | 0.23 | 26 | 3.45 | **139.54** | 92.46-210.58 |
| Age ≥70 years | | | | | | |
| <20 | 16 406 | 90.71 | 217 | 55.22 | **6.49** | 5.33-7.90 |
| 20-39 | 256 | 1.42 | 13 | 3.31 | **21.04** | 11.98-36.94 |
| 40-59 | 325 | 1.80 | 20 | 5.09 | **29.58** | 18.64-46.95 |
| 60-79 | 174 | 0.96 | 17 | 4.33 | **52.39** | 31.86-86.15 |
| 80-119 | 200 | 1.11 | 16 | 4.07 | **41.66** | 24.98-69.47 |
| 120-159 | 120 | 0.66 | 18 | 4.58 | **85.56** | 52.71-138.90 |
| 160-199 | 91 | 0.50 | 16 | 4.07 | **109.39** | 65.59-182.45 |
| 200-299 | 160 | 0.88 | 38 | 9.67 | **147.52** | 103.95-209.35 |
| 300-499 | 94 | 0.52 | 25 | 6.36 | **183.18** | 120.53-278.42 |
| ≥500 | 49 | 0.27 | 10 | 2.54 | **160.95** | 85.08-304.47 |

BMI, body mass index; CEA, carcinoembryonic antigen; CI, confidence interval; CRP, C-reactive protein; FIT, fecal immunochemical test; HR, hazard ratio.
Values in bold were hazard ratios for colorectal cancer.
[a] The number of servings of fruit and vegetables a day was defined as low (≤2), medium (3-4), and high (≥5), respectively.
[b] First-degree relative was used to define positive family history for CRC.
[c] History of black stool passage was grouped as none, occasional (1-2 times/week), and always (≥3 times/week) in the past month.
[d] Bowel habit change included those with regular loose stool, constipation, and high frequency of bowel movement which continued throughout the past month.
[e] Diabetes for those with self-reported history, fasting blood glucose level ≥126 mg/dl, or on hypoglycemic agents.
[f] Hypertension for those with self-reported history, systolic blood pressure ≥140 mmHg, diastolic blood pressure ≥90 mmHg, or on anti-hypertensive agents.
[g] Severe anemia with hemoglobin <10 g/dl and mild anemia with hemoglobin between 10 mg and 14 g/dl for men and 10 mg and 12 g/dl for women.

(Supplementary Appendix Table S3, available at https://doi.org/10.1016/j.esmoop.2021.100288). The first nine examples were for age 55 years, and the last one, age 65 years. The first two examples had no FIT tests and the third one had negative FIT with both lifestyle history for smoking and drinking, medically confirmed hypertension and diabetes, and a positive family history of CRC. The 5-year risk was 2.8% when FIT value was 60 μg/g. When FIT increased from 60 to 80, 120, and 160, 5-year cancer risk increased up to 4.0%, 5.6%, and 7.9%, respectively. When FIT at 160 had all the positive lifestyle and family history along with hypertension and diabetes, the 5-year risk increased to 13.5% or NNS at 7.

## DISCUSSION

By incorporating a single FIT test stratified by age, a user-friendly CRC scorecard was developed in Figure 1B, with C statistic at 0.81, implying great accuracy. A simple score sheet showing the probability of finding a CRC can facilitate the risk communication for colonoscopy referral. In this study, for each positive FIT value, an NNS can be looked up from the scorecard (Figure 1B for 5-year risk). When FIT was negative, risks were too small to be concerned even with other risk factors.

However, we noted that the accuracy of FIT in predicting CRC risk was quite age dependent since age was also an important determinant for the development of CRC (Table 1). Without considering age, the C statistic of FIT-only model in predicting CRC incidence dropped dramatically. The possible explanations for the relatively lower accuracy of FIT in predicting the proximal CRC may include more rapidly growing tumor characteristics of right colon CRC, less irritating bleeding from the proximal colon, and easy degradation of hemoglobin due to longer transit time.[39,40] However, the common etiologies for rectal bleeding remain more variable such as diverticuli, hemorrhoids, and anal fissure to affect the accuracy of FIT in predicting rectal cancer.[40,41] Consistent with the literature, we found that increased FIT values correlated with the poor prognosis of CRC, either incidence or mortality, to indirectly reflect the advanced stage of the CRC.[41,42] Similarly, with the same level of FIT values, CRCs in the proximal colon tended to

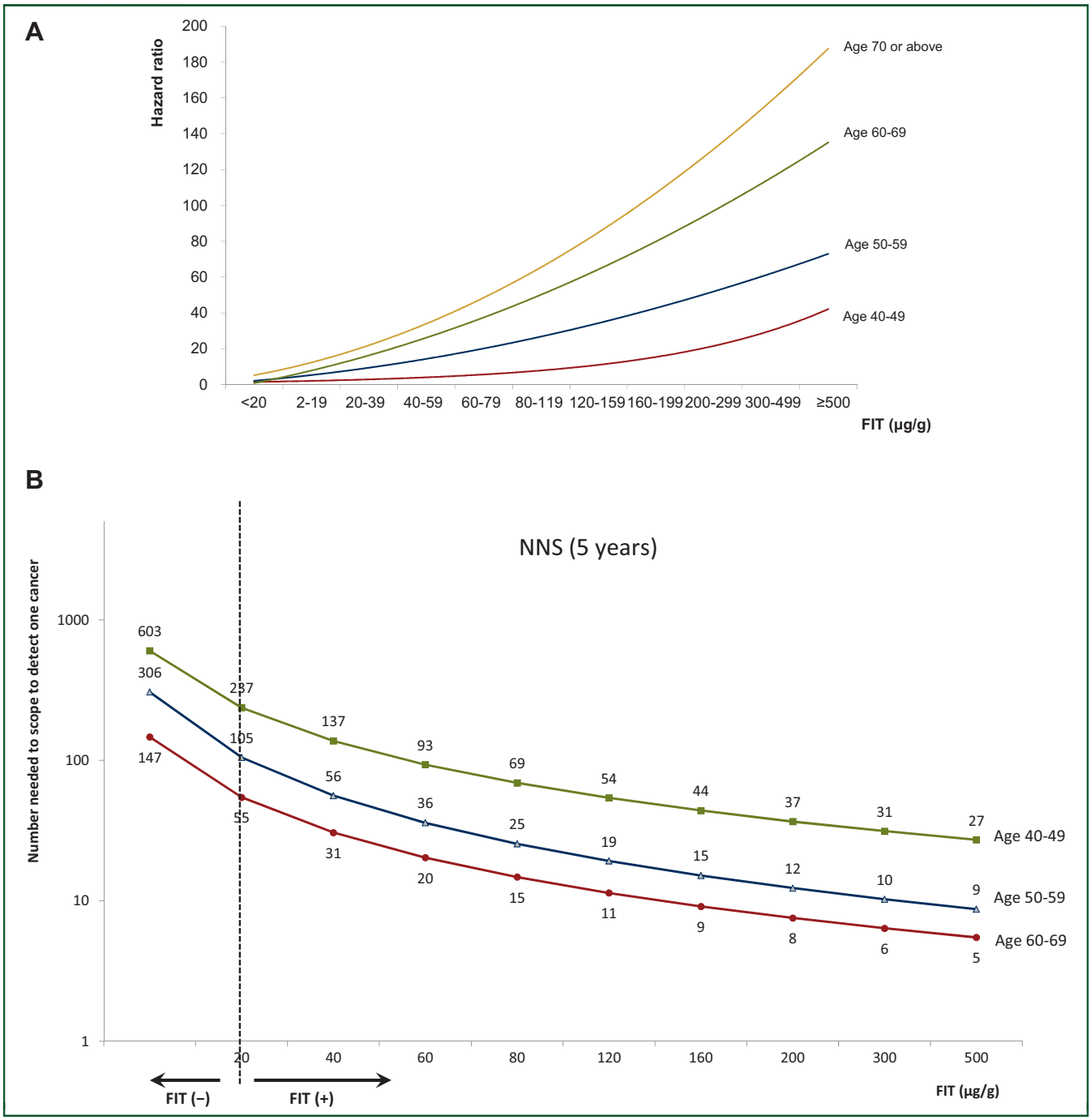


**Figure 1.** **(A) Hazard ratios of colorectal cancer incidence by age and fecal immunochemical test (FIT) groups. (B) NNS (number needed to scope for one cancer) by FIT values [<20 μg/g for FIT (−) and ≥20 μg/g for FIT (+)] in the initial round in the course of follow-up period.**

have poor prognosis due to the relatively lower sensitivity of FIT for the right-sided CRC.

Most prediction models reported had limited application because its usage is often cumbersome. In this study, the entire prediction process was made simple by having a single FIT test. A score sheet to quantify cancer risk was developed, when FIT was positive, showing the number of subjects needed to scope (NNS). Around 95% of FIT tests would show negative results, implying a negligible cancer risk level, around 1/750-1500. Positive FIT could identify individuals with cancer risk 10-fold larger than those with negative FIT. Within positive FIT, quantitative values could also differentiate another 10-fold difference in risk within each age group. While every positive FIT is recommended to undergo colonoscopy, cancer risks varied by 10-fold, making prediction model more clinically relevant for colonoscopy and urgency varied accordingly.

Without considering multiple risk factors, such as family history or lifestyle risks, model 3 can effectively predict cancer risk to reach 0.81 in C statistic when FIT was positive. Age was also a powerful risk factor, and when combined with FIT values, easy-to-use NNS prediction curves by

**Table 2. The goodness of fit for different models**

| | Total colon | | | Proximal colon | | Distal colon | | Rectum | |
|---|---|---|---|---|---|---|---|---|---|
| | C statistic | 95% CI | AIC | C statistic | 95% CI | C statistic | 95% CI | C statistic | 95% CI |
| Model 1: nonspecific questionnaire (age, gender, smoking status, drinking status, physical activity, fruit and vegetables, BMI) | | | | | | | | | |
| Full dataset | 0.723 | 0.71-0.74 | 13 373.99 | 0.732 | 0.70-0.76 | 0.726 | 0.70-0.75 | 0.707 | 0.68-0.73 |
| Training set | 0.729 | 0.71-0.75 | 6536.857 | 0.734 | 0.69-0.78 | 0.711 | 0.67-0.75 | 0.703 | 0.67-0.74 |
| Validation set | 0.719 | 0.70-0.74 | 6852.514 | 0.742 | 0.70-0.78 | 0.757 | 0.72-0.79 | 0.715 | 0.68-0.75 |
| Model 2: All questionnaire (nonspecific questionnaire, family history, black color stool, bowel habit change) | | | | | | | | | |
| Full dataset | 0.726 | 0.71-0.74 | 11 983.854 | 0.735 | 0.71-0.76 | 0.728 | 0.70-0.76 | 0.710 | 0.69-0.73 |
| Training set | 0.731 | 0.71-0.75 | 5874.814 | 0.733 | 0.69-0.78 | 0.717 | 0.68-0.76 | 0.709 | 0.67-0.74 |
| Validation set | 0.724 | 0.70-0.74 | 6134.171 | 0.751 | 0.71-0.79 | 0.759 | 0.72-0.79 | 0.717 | 0.69-0.75 |
| Model 3: FIT, age and gender | | | | | | | | | |
| Full dataset | 0.812 | 0.80-0.83 | 11 800.232 | 0.817 | 0.79-0.84 | 0.848 | 0.82-0.87 | 0.778 | 0.75-0.80 |
| Training set | 0.815 | 0.80-0.84 | 5738.679 | 0.819 | 0.78-0.86 | 0.829 | 0.79-0.87 | 0.775 | 0.74-0.81 |
| Validation set | 0.813 | 0.79-0.83 | 6111.653 | 0.828 | 0.79-0.87 | 0.878 | 0.85-0.91 | 0.795 | 0.76-0.83 |
| Model 3: FIT only | 0.713 | 0.70-0.73 | 12 222.184 | 0.689 | 0.66-0.72 | 0.761 | 0.73-0.79 | 0.689 | 0.67-0.71 |
| Training set | 0.712 | 0.69-0.73 | 5953.527 | 0.698 | 0.65-0.74 | 0.740 | 0.70-0.78 | 0.687 | 0.65-0.72 |
| Validation set | 0.719 | 0.70-0.74 | 6276.06 | 0.680 | 0.63-0.73 | 0.790 | 0.75-0.83 | 0.691 | 0.66-0.72 |
| Model 4: All questionnaire + FIT | | | | | | | | | |
| Full dataset | 0.823 | 0.81-0.84 | 10 540.312 | 0.820 | 0.79-0.85 | 0.860 | 0.84-0.88 | 0.795 | 0.77-0.82 |
| Training set | 0.825 | 0.81-0.84 | 5194.875 | 0.824 | 0.78-0.87 | 0.845 | 0.81-0.88 | 0.796 | 0.76-0.83 |
| Validation set | 0.826 | 0.81-0.85 | 5414.879 | 0.840 | 0.81-0.87 | 0.891 | 0.86-0.92 | 0.811 | 0.78-0.84 |
| Model 5: All questionnaire + FIT + blood tests (diabetes, hypertension, anemia, CEA, CRP) | | | | | | | | | |
| Full dataset | 0.827 | 0.81-0.85 | 10 482.619 | 0.830 | 0.80-0.86 | 0.859 | 0.84-0.88 | 0.803 | 0.78-0.83 |
| Training set | 0.839 | 0.82-0.86 | 5181.593 | 0.833 | 0.79-0.87 | 0.848 | 0.81-0.88 | 0.809 | 0.78-0.84 |
| Validation set | 0.830 | 0.81-0.85 | 5378.192 | 0.853 | 0.82-0.88 | 0.888 | 0.86-0.92 | 0.817 | 0.79-0.85 |

AIC, Akaike information criterion; CEA, carcinoembryonic antigen; CI, confidence interval; CRP, C-reactive protein; FIT, fecal immunochemical test.

different age groups were developed (Figure 1B). Reading the size of NNS, similar to one from model 3, would give one an impression of the cancer risk and reflected the urgency for colonoscopy if FIT is positive. When risk factors were additionally considered with examples shown in Supplementary Appendix Table S3 (available at https://doi.org/10.1016/j.esmoop.2021.100288), cancer risk increased and accuracy improved (AUC changed from 0.81 to 0.83).

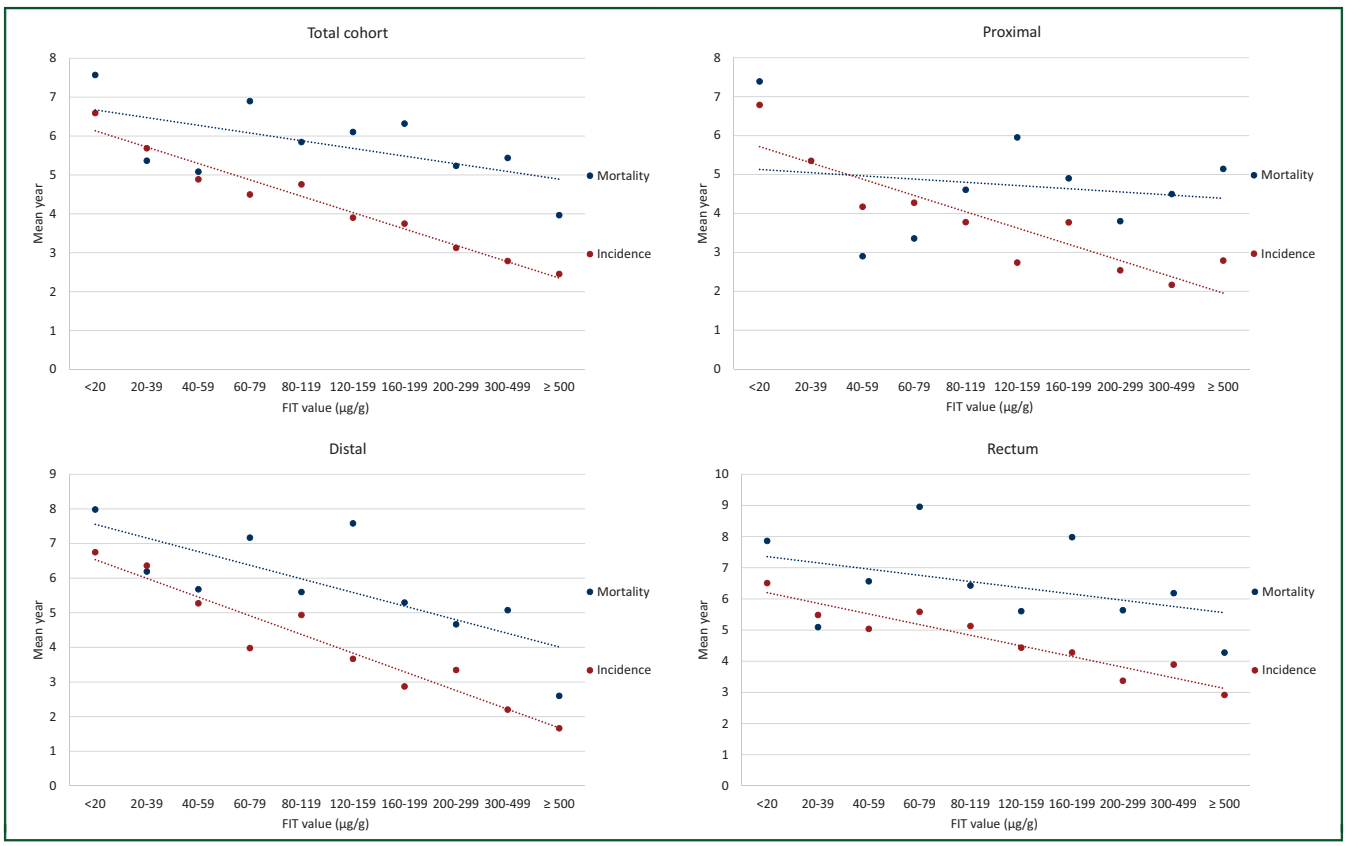


**Figure 2. Mean number of years from fecal immunochemical test (FIT) to the development of colorectal cancer (CRC) cases and CRC mortality based on the CRC location.**

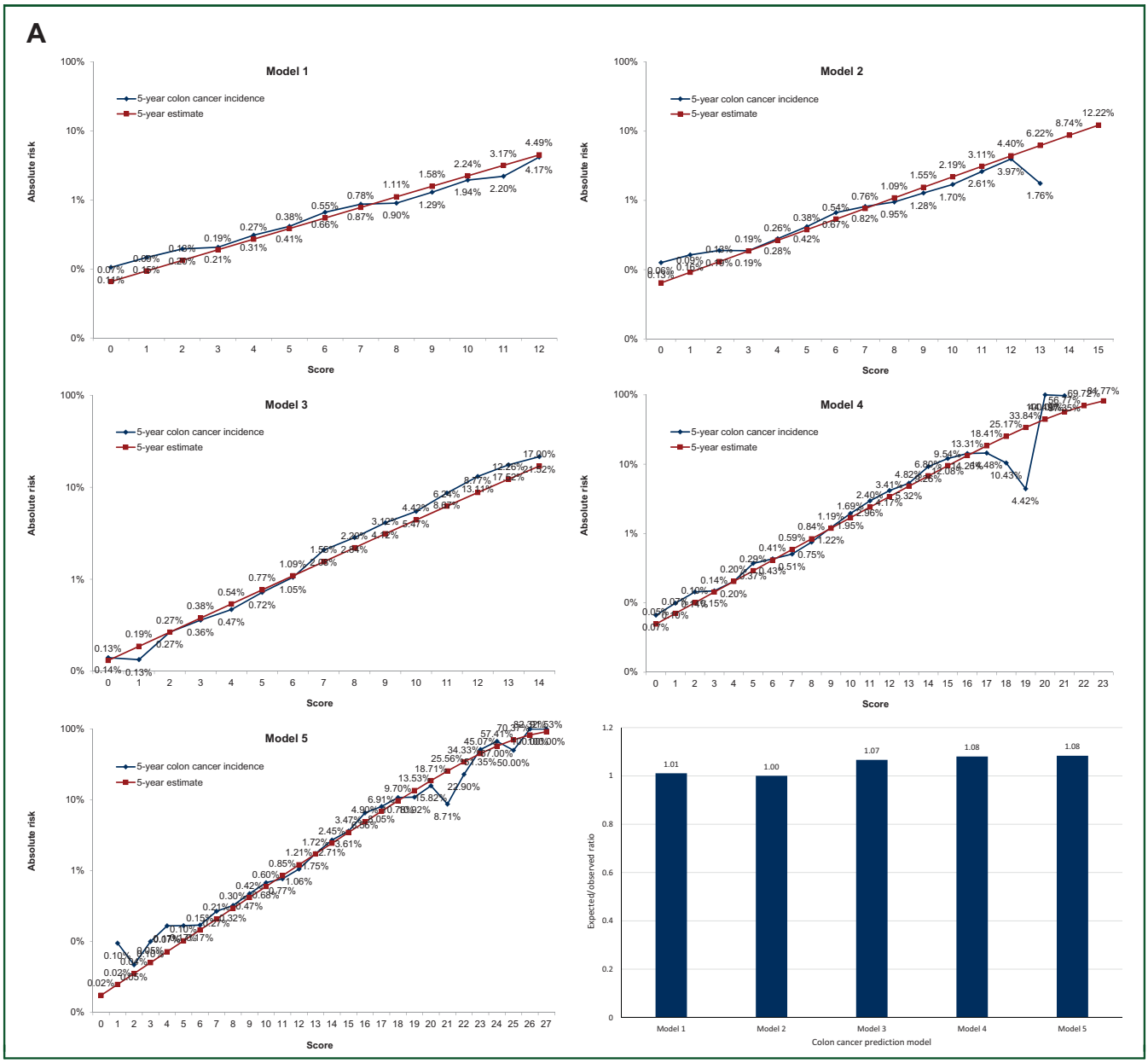


**Figure 3.** **(A) Internal calibration of the risk prediction models. Calibration determined the extent of agreement between predicted and observed events in 5 years and a cross-validated calibration plot was generated for the different models. (B) Internal calibration of the risk prediction models. Calibration determined the extent of agreement between predicted and observed events in 10 years and a cross-validated calibration plot was generated for the different models.**

When FIT is negative, the cancer risk is sufficiently low so that the addition of other risk factors remains too low to be concerned. It required a combination of eight or more risk factors to reach the lowest risk of a positive FIT. Since such individuals constituted <1% of the negative FIT (Supplementary Appendix Table S4, available at https://doi.org/10.1016/j.esmoop.2021.100288), cancer risk of negative FIT as a whole could be ignored without the need to check the model. This observation is compatible with the current guideline not pursuing colonoscopy for negative FIT, representing the 94%-96% of the cohort (Supplementary Appendix Table S1, available at https://doi.org/10.1016/j.esmoop.2021.100288). The 10 examples we gave illustrated the power of positive FIT, far larger than most known risk factors. When FIT was negative, common risk factors added limited cancer risks; but when FIT was positive, addition of risk factors increased the cancer risks.

To evaluate the performance of prediction models, four criteria were often used. Firstly, accuracy in discriminating ability to find a cancer. Most known risk factors for CRC had relative risks around 2-3-fold increase at most, and were dwarfed by the up to 10-100-fold increase shown by quantitative FIT values. The C statistic in our models ranged from 0.72 to 0.83, with better accuracy when FIT was added. The addition of known risk factors from history or blood tests only improved the C statistic minimally from 0.81 to 0.83. Secondly, calibration in prediction results

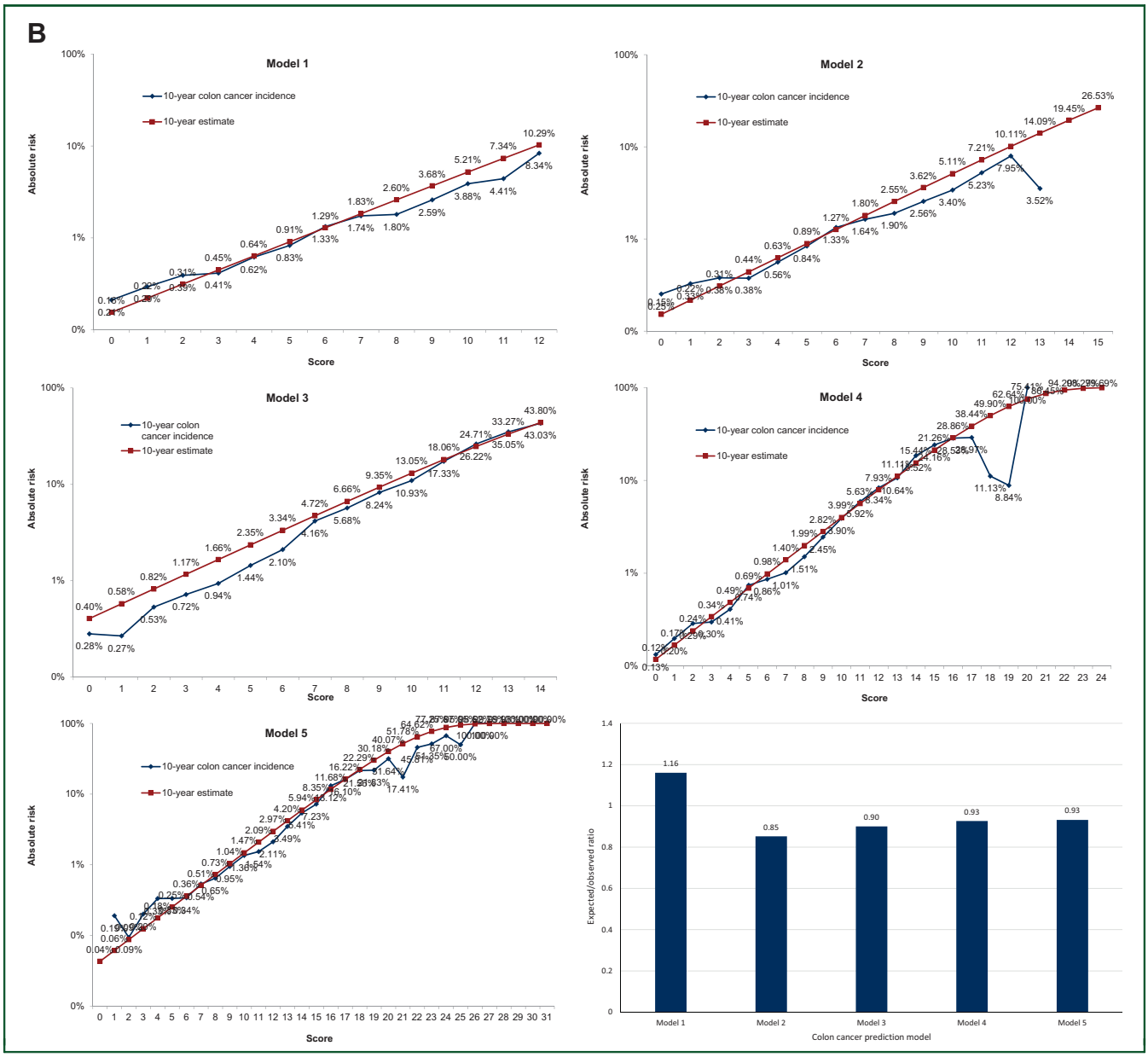


**Figure 3. Continued.**

matching observed event. All the E/O slopes of these five models approximated to 1, particularly for 5-year prediction. However, the E/O slope of model 3 was more balanced to be close to 1 for both 5-year or 10-year predictions. Under visual representation of the calibration curves, model 1 and model 3 also showed closer relationship between estimated and observed events.[34] Thirdly, validation in applicability from study population to different target population. It requires external validation done by independent database and investigators before the widespread use of our prediction model.[43] Moreover, the predicted probability provided by our models or the figures of our NNS nomogram may vary with local incidence of CRC and the different FIT laboratories.[44] However, the literature conducted between different ethnics has consistently supported the dose-response effect between the incidence of CRC and quantitative FIT values.[45-47] Fourthly, user convenience with simple, informative, and real-world relevance.[34,48-50] One characteristic of FIT is age dependent and the NNS nomogram we developed was equivalent to model 3, so that its accuracy, calibration and validation have been fully addressed and found satisfactory. In addition, it fit 'user convenience' well, as it is simple to use, highly informative particularly in decision making for immediately colonoscopy or not, and provides real-world relevance in the perception of probability of cancer risk.

The step-wise nature of the model allowed a progressive identification of high-risk individuals based on availability of data or resources. In this way, anyone without medical tests could use this model based on his/her own health history. We divided history model into CRC-specific and CRC-nonspecific risk factors. The model provided the list of



https://doi.org/10.1016/j.esmoop.2021.100288

CRC-specific symptoms, such as bowel habit change or black-colored stools, that model users should seek care immediately, and not to ignore or dismiss them. The model would also alert those with CRC-nonspecific risk factors, such as smoking, drinking, or exercise, and motivate them to change lifestyles. Even though these relative risks for CRC were only modestly elevated, their reductions will create far more health benefits than affecting CRC only.

FIT test is relatively inexpensive and non-intrusive, compared to sigmoidoscopy or colonoscopy, and the message from using the prediction model encourages everyone to have an FIT test. In this way, the model promotes screening to identify high-risk individuals at an earlier stage, and subsequent measures to reduce these risks. The added power from blood work-up was limited, except for finding serious anemia. Identification of serious anemia is an ominous sign, regardless of its relationship with CRC, and is an urgent call for seeking medical care. Both CEA and CRP are nonspecific tumor maker, but could add value to the prediction model. We were the first one to report this in a prediction model.

The strengths of this study, other than the large sample size, included the consistent use of FIT test based on OC sensor since 1994, single laboratory that analyzed the blood and stool test, the identification of cancer from National Cancer Registry with 97% completeness and an Asian cohort with FIT tests. Furthermore, the high concordance between the National Death File, the National Cancer Registry file, and the National Health Insurance database has been validated in the literature.[51,52]

There are important limitations to this prediction model. Firstly, CRC in this study was identified not by immediate colonoscopy but by cancer registry. Thus, there was a delay and time gap of FIT, positive or negative, and the development of CRC. However, a cancer registry system with complete case ascertainment would be preferred for developing a prediction model and assessing the survival. Colonoscopy was conducted mainly after positive FIT, but not all cancers expected in 5-10 years would be identified by immediate colonoscopy. We found the delay for the cancer to develop was proportionate to quantitative FIT values, with its interval shortened by larger values of FIT test. For every 80-μg/g increase of FIT, there were 1.5-year earlier development of CRC incidence and 1-year earlier development of CRC mortality, respectively. When FIT was negative, no colonoscopy was normally accompanied, and cancer cases were mainly identified by cancer registry, with an average delay of 6 years. Secondly, the study chose a cut point at 20 μg/g for positive FIT, and results would vary if different cut points were chosen. However, the risk was quantified and increased with increasing values. This model with quantitative FIT was much superior than that reported FIT based on simply dichotomy. Thirdly, the model came from a cohort with apparently healthy subjects, and the model is not applicable to those individuals with various severity of gastrointestinal symptoms, such as constipation or bowel habit changes.

### *Conclusion*

With C statistic at 0.81, a simple user-friendly model predicting CRC was developed from a single FIT value by considering age and sex. Age was also a powerful risk factor, and easy-to-use NNS prediction curves by different FIT value age groups could be used for risk communication. The FIT values correlate well with prognosis of CRC. With the same FIT value, CRC in the proximal colon tends to have poor survival. When positive, FIT provides a quantitative risk scorecard, expressed in NNS, better than many risk factors combined, but when FIT is negative, risk is too small to be concerned, even considering other risk factors.

## ACKNOWLEDGEMENTS

All or part of the data used in this research were authorized by and received from MJ Health Research Foundation (Authorization Code: MJHRFB2014001C). Any interpretation or conclusion described in this paper does not represent the views of MJ Health Research Foundation. MJ Health Survey Database and MJ BioData were available from MJ Health Research Foundation Web site: http://www.mjhrf.org. CPW and CHC had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## FUNDING

This work was supported in part by Taiwan Ministry of Health and Welfare Clinical Trial Center (MOHW108-TDU-B-212-133004). The funding source had no role in study design, data collection, data analysis, data interpretation, writing of the report, or the decision to submit the paper for publication.

## DISCLOSURE

The authors have declared no conflicts of interest.

## 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(3):209-249.
2. Siegel RL, Miller KD, Goding Sauer A, et al. Colorectal cancer statistics, 2020. *CA Cancer J Clin*. 2020;70(3):145-164.
3. Fidler MM, Soerjomataram I, Bray F. A global view on cancer incidence and national levels of the human development index. *Int J Cancer*. 2016;139(11):2436-2446.
4. Health Promotion Administration. Taiwan cancer registry. Available at https://www.hpa.gov.tw/EngPages/Detail.aspx?nodeid=1061&pid=6069. Accessed May 18, 2021.
5. Win AK, Macinnis RJ, Hopper JL, Jenkins MA. Risk prediction models for colorectal cancer: a review. *Cancer Epidemiol Biomarkers Prev*. 2012;21(3):398-410.
6. Cai QC, Yu ED, Xiao Y, et al. Derivation and validation of a prediction rule for estimating advanced colorectal neoplasm risk in average-risk Chinese. *Am J Epidemiol*. 2012;175(6):584-593.
7. Colditz GA, Atwood KA, Emmons K, et al. Harvard report on cancer prevention volume 4: Harvard Cancer Risk Index. Risk Index Working Group, Harvard Center for Cancer Prevention. *Cancer Causes Control*. 2000;11(6):477-488.

8. Digestive Disease Institute mIG, Cleveland Clinic. Colon cancer risk assessment. Available at http://digestive.ccf.org/scores/go. Accessed October 30, 2020.
9. Driver JA, Gaziano JM, Gelber RP, Lee IM, Buring JE, Kurth T. Development of a risk score for colorectal cancer in men. *Am J Med*. 2007;120(3):257-263.
10. Emmons KM, Wong M, Puleo E, Weinstein N, Fletcher R, Colditz G. Tailored computer-based cancer risk communication: correcting colorectal cancer risk perception. *J Health Commun*. 2004;9(2):127-141.
11. Imperiale TF, Wagner DR, Lin CY, Larkin GN, Rogge JD, Ransohoff DF. Using risk for advanced proximal colonic neoplasia to tailor endoscopic screening for colorectal cancer. *Ann Intern Med*. 2003;139(12):959-965.
12. Freedman AN, Slattery ML, Ballard-Barbash R, et al. Colorectal cancer risk prediction tool for white men and women without known susceptibility. *J Clin Oncol*. 2009;27(5):686-693.
13. Jeon J, Du M, Schoen RE, et al. Determining risk of colorectal cancer and starting age of screening based on lifestyle, environmental, and genetic factors. *Gastroenterology*. 2018;154(8):2152-2164.e19.
14. Ma E, Sasazuki S, Iwasaki M, et al. 10-Year risk of colorectal cancer: development and validation of a prediction model in middle-aged Japanese men. *Cancer Epidemiol*. 2010;34(5):534-541.
15. National Cancer Institute. Colorectal cancer risk assessment tool. Available at https://ccrisktool.cancer.gov/. Accessed October 30, 2020.
16. Shin A, Joo J, Yang H-R, et al. Risk prediction model for colorectal cancer: National Health Insurance Corporation study, Korea. *PLoS One*. 2014;9(2):e88079.
17. Steffen A, MacInnis RJ, Joshy G, Giles GG, Banks E, Roder D. Development and validation of a risk score predicting risk of colorectal cancer. *Cancer Epidemiol Biomarkers Prev*. 2014;23(11):2543-2552.
18. Taylor DP, Stoddard GJ, Burt RW, et al. How well does family history predict who will get colorectal cancer? Implications for cancer screening and counseling. *Genet Med*. 2011;13(5):385-391.
19. Wei EK, Colditz GA, Giovannucci EL, Fuchs CS, Rosner BA. Cumulative risk of colon cancer up to age 70 years by risk factor status using data from the Nurses' Health Study. *Am J Epidemiol*. 2009;170(7):863-872.
20. Wells BJ, Kattan MW, Cooper GS, Jackson L, Koroukian S. Colorectal cancer predicted risk online (CRC-PRO) calculator using data from the multi-ethnic cohort study. *J Am Board Fam Med*. 2014;27(1):42-55.
21. Liu Y, Colditz GA, Rosner BA, Dart H, Wei E, Waters EA. Comparison of performance between a short categorized lifestyle exposure-based colon cancer risk prediction tool and a model using continuous measures. *Cancer Prev Res (Phila)*. 2018;11(12):841-848.
22. Smith T, Muller DC, Moons KGM, et al. Comparison of prognostic models to predict the occurrence of colorectal cancer in asymptomatic individuals: a systematic literature review and external validation in the EPIC and UK Biobank prospective cohort studies. *Gut*. 2019;68(4):672-683.
23. Usher-Smith JA, Walter FM, Emery JD, Win AK, Griffin SJ. Risk prediction models for colorectal cancer: a systematic review. *Cancer Prev Res (Phila)*. 2016;9(1):13-26.
24. Cross AJ, Wooldrage K, Robbins EC, et al. Faecal immunochemical tests (FIT) versus colonoscopy for surveillance after screening and polypectomy: a diagnostic accuracy and cost-effectiveness study. *Gut*. 2019;68(9):1642.
25. Cooper JA, Parsons N, Stinton C, et al. Risk-adjusted colorectal cancer screening using the FIT and routine screening data: development of a risk prediction model. *Br J Cancer*. 2018;118(2):285-293.
26. Li W, Zhao LZ, Ma DW, et al. Predicting the risk for colorectal cancer with personal characteristics and fecal immunochemical test. *Medicine (Baltimore)*. 2018;97(18):e0529.
27. Zheng Y, Hua X, Win AK, et al. A new comprehensive colorectal cancer risk prediction model incorporating family history, personal characteristics, and environmental factors. *Cancer Epidemiol Biomarkers Prev*. 2020;29(3):549-557.
28. Wen CP, Cheng TY, Tsai MK, et al. All-cause mortality attributable to chronic kidney disease: a prospective cohort study based on 462 293 adults in Taiwan. *Lancet*. 2008;371(9631):2173-2182.
29. Wen CP, Wai JP, Tsai MK, et al. Minimum amount of physical activity for reduced mortality and extended life expectancy: a prospective cohort study. *Lancet*. 2011;378(9798):1244-1253.
30. Ministry of Health and Welfare. Cause of death statistics. Available at https://www.mohw.gov.tw/np-128-2.html. Accessed October 30, 2020.
31. Taiwan Cancer Registry. Taiwan cancer registry, introduction. Available at http://tcr.cph.ntu.edu.tw/main.php?Page=N1. Accessed October 30, 2020.
32. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics*. 2000;56(2):337-344.
33. D'Agostino RB, Nam B-H. Evaluation of the performance of survival analysis models: discrimination and calibration measures. In: Balakrishnan N, Rao CR, editors. *Handbook of Statistics*. Amsterdam: Elsevier; 2003:1-25.
34. Alba AC, Agoritsas T, Walsh M, et al. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. *J Am Med Assoc*. 2017;318(14):1377-1384.
35. Hosmer DW, Lemeshow S, May S. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. Washington, DC: Wiley-Interscience; 2008.
36. Vilkin A, Rozen P, Levi Z, et al. Performance characteristics and evaluation of an automated-developed and quantitative, immunochemical, fecal occult blood screening test. *Am J Gastroenterol*. 2005;100(11):2519-2525.
37. Akaike H. Information theory and an extension of the maximum likelihood principle. In: Petrov BN, Csaki F, editors. *2nd International Symposium on Information Theory*. Budapest, Hungary: Akademia Kiado; 1973:267-281.
38. Cavanaugh JE, Neath AA. The Akaike information criterion: background, derivation, properties, application, interpretation, and refinements. *WIREs Comput Stat*. 2019;11:e1460.
39. Doubeni CA, Levin TR. In screening for colorectal cancer, is the FIT right for the right side of the colon? *Ann Int Med*. 2018;169(9):650-651.
40. de Klerk CM, Vendrig LM, Bossuyt PM, Dekker E. Participants-related risk factors for false-positive and false-negative fecal immunochemical tests in colorectal cancer screening: systematic review and meta-analysis. *Am J Gastroenterol*. 2018;113:1778-1787.
41. Amitay EL, Cuk K, Niedermaier T, Weigl K, Brenner H. Factors associated with false-positive fecal immunochemical tests in a large german colorectal cancer screening study. *Int J Cancer*. 2019;144:2419-2427.
42. Yuan SY, Wu W, Fu J, et al. Quantitative immunochemical fecal occult blood test for neoplasia in colon cancer screening. *J Dig Dis*. 2019;20:78-82.
43. Collins GS, de Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14:40.
44. Gies A, Bhardwaj M, Stock C, Schrotz-King P, Brenner H. Quantitative fecal immunochemical tests for colorectal cancer screening. *Int J Cancer*. 2018;143(2):234-244.
45. Brenner H, Qian J, Werner S. Variation of diagnostic performance of fecal immunochemical testing for hemoglobin by sex and age: results from a large screening cohort. *Clin Epidemiol*. 2018;10:381-389.
46. Cubiella J, Digby J, Rodriguez-Alonso L, et al. The fecal hemoglobin concentration, age and sex test score: development and external validation of a simple prediction tool for colorectal cancer detection in symptomatic patients. *Int J Cancer*. 2017;140(10):2201-2211.
47. Selby K, Jensen CD, Lee JK, et al. Influence of varying quantitative fecal immunochemical test positivity thresholds on colorectal cancer detection: a community-based cohort study. *Ann Intern Med*. 2018;169(7):439-447.
48. Rosella LC, Corey P, Stukel TA, Mustard C, Hux J, Manuel DG. The influence of measurement error on calibration, discrimination, and overall estimation of a risk prediction model. *Popul Health Metr*. 2012;10(1):20.
49. Snell KI, Ensor J, Debray TP, Moons KG, Riley RD. Meta-analysis of prediction model performance across multiple studies: which scale

helps ensure between-study normality for the C-statistic and calibration measures? *Stat Methods Med Res*. 2018;27(11):3505-3522.
50. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230.
51. Kao WH, Hong JH, See LC, Yu HP, Hsu JT, Chou IJ, et al. Validity of cancer diagnosis in the National Health Insurance database compared with the linked National Cancer Registry in Taiwan. *Pharmacoepidemiol Drug Saf*. 2018;27(10):1060-1066.
52. Chan HC, Huang CC, Huang CC, et al. Predicting colon cancer-specific survival for the asian population using national cancer registry data from Taiwan. *Ann Surg Oncol*. 2021. https://doi.org/10.1245/s10434-021-10646-2. Epub ahead of print.