

## Supplementary Materials

**Supplementary Table 1.** Scoring algorithm to calculate point values of ANN-based FIT-specific scoring systems for colorectal neoplasia by two-sample FIT results

Risk factors	Subjects positive in either FIT				Subjects negative in both FIT			
	For CRC		For CRC and adenoma		For CRC		For CRC and adenoma	
	Contribution (%)	Score †	Contribution (%)	Score †	Contribution (%)	Score †	Contribution (%)	Score †
Age at screening (years)	22.3		19.7		20.2		37.5	
Age group								
50-54	--	0	--	0	--	0	--	0
55-59	--	0.5	--	0.75	--	0.75	--	0.875
60-64	--	1.0	--	1.5	--	1.5	--	1.75
65-69	--	1.5	--	2.25	--	2.25	--	2.625
70-74	--	2.0	--	3.0	--	3.0	--	3.5
Sex								
Women	--	0	--	0	--	0	--	0
Men	15.9	2.0	24.4	4.0	11.5	2.0	27.2	2.5
Area of residence								
Rural	--	0	--	0	--	0	23.4	2.0
Urban	14.3	1.0	19.5	3.0	14.2	2.0	--	0
Chronic diarrhea								
Never	--		--		--	0	--	
Ever	--		--		16.6	2.0	--	
Chronic constipation								
Never	--		8.3	1.0	--		--	
Ever	--		--	0	--		--	
Chronic appendicitis / appendectomy								

Never	--		9.5	2.0	--		--	
Ever	--		--	0	--		--	
Mucus or bloody stool								
Never	--	0	--		--	0	--	
Ever	35.1	4.0	--		9.2	1.0	--	
Serious unhappy events								
Never	--		--	0	--		--	
Ever	--		4.5	1.0	--		--	
CRC in 1 <sup>st</sup> degree relatives								
Never	--	0	--	0	--	0	--	0
Ever	12.4	1.0	4.3	1.0	15.0	2.0	11.9	1.0
Colorectal polyps								
Never	--		9.8	2.0	--		--	
Ever	--		--	0	--		--	
Diagnosis of any cancer								
Never	--		--		--	0	--	
Ever	--		--		13.3	2.0	--	
<b>Total</b>	<b>--</b>	<b>0-10.0</b>	<b>--</b>	<b>0-17.0</b>	<b>--</b>	<b>0-14.0</b>	<b>--</b>	<b>0-9.0</b>

CRC, colorectal cancer; FIT, fecal immunochemical test; ANN, artificial neural network.

† Computed by multiplying the total score in the logistic regression-based FIT-specific scoring system by the contributions of predictors on the outcome in the corresponding ANN model.

**Supplementary Table 2.** Scoring algorithm to calculate point values of LR-based FIT-specific scoring systems for colorectal neoplasia by the results of first FIT

Risk factors	For CRC in all subjects					For CRC and adenoma in subjects positive in the first FIT and attending colonoscopy		
	Reference value (W <sub>ij</sub> )	Positive in the first FIT		Negative in the first FIT		Reference value (W <sub>ij</sub> )	β (95%CI) †	Score ‡
		β (95%CI) †	Score ‡	β (95%CI) †	Score ‡			
Age at screening (years)		0.055 (0.044, 0.065)		0.058 (0.048, 0.068)			0.033 (0.027, 0.039)	
Age group								
50-54	52 (W <sub>ref</sub> )	--	0	--	0	52 (W <sub>ref</sub> )	--	0
55-59	57	--	1.0	--	1.0	57	--	1.0
60-64	62	--	2.0	--	2.0	62	--	2.0
65-69	67	--	3.0	--	3.0	67	--	3.0
70-74	72	--	4.0	--	4.0	72	--	4.0
Sex								
Women	0 (W <sub>ref</sub> )	--	0	--	0	0 (W <sub>ref</sub> )	--	0
Men	1	0.406 (0.287, 0.524)	1.0	0.401 (0.284, 0.519)	1.0	1	0.575 (0.507, 0.644)	3.0
Area of residence								
Rural	0 (W <sub>ref</sub> )	--	0	--	0	0 (W <sub>ref</sub> )	--	0
Urban	1	0.379 (0.251, 0.507)	1.0	0.269 (0.151, 0.387)	1.0	1	0.197 (0.111, 0.282)	1.0
Chronic diarrhea								
Never	0 (W <sub>ref</sub> )	--		--	0	--	--	
Ever	1	--		0.463 (0.252, 0.674)	2.0	--	--	
Mucus or bloody stool								
Never	0 (W <sub>ref</sub> )	--	0	--	0	0 (W <sub>ref</sub> )	--	0
Ever	1	0.827 (0.606, 1.048)	3.0	0.224 (-0.110, 0.559)	1.0	1	0.273 (0.106, 0.441)	2.0
Chronic constipation								
Never	--	--		--		1	0.293 (0.160, 0.426)	2.0

Ever	--	--		--		0 (W <sub>ref</sub> )	--	0
Chronic appendicitis / appendectomy								
Never	--	--		--		1	0.114 (0.010, 0.219)	1.0
Ever	--	--		--		0 (W <sub>ref</sub> )	--	0
Serious unhappy events								
Never	--	--		--		0 (W <sub>ref</sub> )	--	0
Ever	--	--		--		1	0.443 (0.251, 0.636)	3.0
CRC in 1 <sup>st</sup> degree relatives								
Never	0 (W <sub>ref</sub> )	--	0	--	0	--	--	
Ever	1	0.420 (0.170, 0.670)	2.0	0.673 (0.429, 0.918)	2.0	--	--	
Diagnosis of any cancer								
Never	0 (W <sub>ref</sub> )	--		--	0	--	--	
Ever	1	--		0.650 (0.361, 0.938)	2.0	--	--	
<b>Total</b>	--	--	<b>0-11.0</b>	--	<b>0-13.0</b>	--	--	<b>0-16.0</b>

FIT, fecal immunochemical test; LR, logistic regression; CRC, colorectal cancer; CI, confidence interval.

<sup>†</sup>  $\beta$  (95%CI) derived from multivariable LR model.

<sup>‡</sup> LR-based risk score= $\beta \cdot (W_{ij} - W_{ref}) / B$ , in which constant B is the number of regression unit equivalent to 1 point in the final risk score, and was calculated by multiplying the  $\beta$  for age by 5, e.g. for the LR-based scoring system for CRC in FIT-positive subjects, constant B=0.055\*5=0.275. Based on an age-standardized method, the point values of other variables were obtained with their corresponding regression coefficients dividing by 0.275 and rounding to the nearest whole number: 0.406/0.275=1.0 for sex, 0.379/0.275=1.0 for area of residence, 0.827/0.275=3.0 for mucus or bloody stool, and 0.420/0.275=2.0 for CRC in 1<sup>st</sup> degree relatives.

**Supplementary Table 3.** Scoring algorithm to calculate point values of ANN-based FIT-specific scoring systems for colorectal neoplasia by the results of first FIT

Risk factors	Subjects positive in the first FIT				Subjects negative in the first FIT for CRC	
	For CRC		For CRC and adenoma			
	Contribution (%)	Score <sup>†</sup>	Contribution (%)	Score <sup>†</sup>	Contribution (%)	Score <sup>†</sup>
Age at screening (years)	17.8		21.9		29.6	
Age group						
50-54	--	0	--	0	--	0
55-59	--	0.5	--	1.0	--	1.0
60-64	--	1.0	--	2.0	--	2.0
65-69	--	1.5	--	3.0	--	3.0
70-74	--	2.0	--	4.0	--	4.0
Sex						
Women	--	0	--	0	--	0
Men	15.5	1.5	34.0	5.0	21.7	3.0
Area of residence						
Rural	--	0	--	0	--	0
Urban	14.3	1.5	11.4	2.0	13.2	1.5
Chronic diarrhea						
Never	--		--		--	0
Ever	--		--		11.8	1.5
Mucus or bloody stool						
Never	--	0	--	0	--	0
Ever	33.4	4.0	8.0	1.0	8.2	1.0
Chronic constipation						
Never	--		8.0	1.0	--	
Ever	--		--	0	--	
Chronic appendicitis / appendectomy						
Never	--		6.9	1.0	--	
Ever	--		--	0	--	
Serious unhappy events						
Never	--		--	0	--	
Ever	--		9.8	2.0	--	
CRC in 1 <sup>st</sup> degree relatives						
Never	--	0	--		--	0
Ever	19.0	2.0	--		8.6	1.0
Diagnosis of any cancer						
Never	--		--		--	0
Ever	--		--		6.9	1.0
<b>Total</b>	<b>--</b>	<b>0-11.0</b>	<b>--</b>	<b>0-16.0</b>	<b>--</b>	<b>0-13.0</b>

ANN, artificial neural network; FIT, fecal immunochemical test; CRC, colorectal cancer.

<sup>†</sup> Computed by multiplying the total score in the logistic regression-based FIT-specific scoring system by the contributions of predictors on the outcome in the corresponding ANN model.

**Supplementary Table 4.** Sensitivity analyses for discrimination and calibration of several primary screening tests in detection of colorectal neoplasia

Primary screening tests	In derivation set		In validation set	
	AUC (95%CI)	<i>P</i> for calibration *	AUC (95%CI)	<i>P</i> for calibration *
<b>Redefining FIT positive as positive in the first FIT</b>				
For CRC in all subjects				
LR-based FIT-specific system	0.779 (0.768, 0.789)	0.774	0.793 (0.769, 0.816)	0.723
ANN-based FIT-specific system	0.775 (0.765, 0.786)	0.964	0.785 (0.761, 0.809)	0.090
For CRC and adenoma in subjects attending colonoscopy †				
LR-based FIT-specific system	0.612 (0.603, 0.621)	0.096	0.608 (0.589, 0.626)	0.496
ANN-based FIT-specific system	0.610 (0.601, 0.619)	0.336	0.615 (0.596, 0.633)	0.481
<b>Supplementing missed/interval CRC within 1-year of screening</b>				
For CRC in all subjects				
LR-based FIT-specific system	0.850 (0.840, 0.859)	0.792	0.919 (0.906, 0.933)	0.881
ANN-based FIT-specific system	0.847 (0.837, 0.857)	0.600	0.918 (0.903, 0.932)	0.217
For CRC and adenoma in subjects attending colonoscopy ‡				
LR-based FIT-specific system	0.659 (0.653, 0.666)	0.008	0.659 (0.645, 0.672)	0.247
ANN-based FIT-specific system	0.648 (0.641, 0.654)	< 0.001	0.651 (0.638, 0.664)	0.023
<b>Using CRC and adenoma detected within 90 days as outcome ‡</b>				
LR-based FIT-specific system	0.654 (0.647, 0.661)	0.104	0.651 (0.637, 0.665)	0.441
ANN-based FIT-specific system	0.652 (0.645, 0.659)	0.006	0.646 (0.632, 0.660)	0.023

AUC, area under the receiver operating characteristic curve; CRC, colorectal cancer; CI, confidence interval; FIT, fecal immunochemical test; LR, logistic regression; ANN, artificial neural network. \* Based on the Hosmer-Lemeshow goodness-of-fit test; † Subjects positive in the first FIT and attending colonoscopy; ‡ Subjects positive in any FIT and attending colonoscopy.

**Supplementary Table 5.** Sensitivity analysis for performance of several primary screening tests in detection of colorectal neoplasia

Primary screening test	High-risk subjects, n (%)	Detection rate, n (%)	Sensitivity (95% CI), %	Specificity (95% CI), %	No. of subjects needed to screen for one case	No. of colonoscopy needed to detect one case
<b>Redefining FIT positive as positive in the first FIT</b>						
For CRC in all subjects (n = 807,109)						
LR-based system incorporating FIT result	84074 (10.4)	1529 (0.19)	54.6 (52.7, 56.4)	89.7 (89.7, 89.8)	528	55
LR-based FIT-specific system	106743 (13.2)	1607 (0.20)	57.3 (55.5, 59.1)	86.9 (86.9, 87.0)	502	66
ANN-based FIT-specific system	172664 (21.4)	1733 (0.21)	61.8 (60.0, 63.6)	78.8 (78.7, 78.8)	466	100
For CRC and adenoma in subjects attending colonoscopy (n = 32,233) †						
LR-based FIT-specific system	15252 (47.3)	3140 (9.74)	62.0 (60.7, 63.4)	55.4 (54.8, 56.0)	10	5
ANN-based FIT-specific system	15057 (46.7)	3118 (9.67)	61.6 (60.3, 62.9)	56.1 (55.5, 56.7)	10	5
<b>Supplementing missed/interval CRC within 1-year of screening</b>						
For CRC in all subjects (n = 807,109)						
LR-based system incorporating FIT result	111336 (13.8)	1570 (0.19)	76.0 (74.1, 77.8)	86.4 (86.3, 86.4)	514	71
LR-based FIT-specific system	151027 (18.7)	1623 (0.20)	78.6 (76.8, 80.3)	81.4 (81.4, 81.5)	497	93
ANN-based FIT-specific system	166580 (20.6)	1641 (0.20)	79.5 (77.7, 81.2)	79.5 (79.4, 79.6)	492	102
For CRC and adenoma in subjects attending colonoscopy (n = 71,023) ‡						
LR-based FIT-specific system	26557 (37.4)	4794 (6.75)	57.3 (56.2, 58.4)	65.3 (64.9, 65.6)	15	6
ANN-based FIT-specific system	27630 (38.9)	4710 (6.63)	56.3 (55.2, 57.4)	63.4 (63.0, 63.8)	15	6
<b>Using CRC and adenoma detected within 90 days as outcome (n=71,023) ‡</b>						
LR-based FIT-specific system	26290 (37.0)	4216 (5.94)	56.9 (55.8, 58.1)	65.3 (64.9, 65.7)	17	6

ANN-based FIT-specific system	27340 (38.5)	4272 (6.02)	57.7 (56.6, 58.8)	63.7 (63.4, 64.1)	17	6
-------------------------------	--------------	-------------	-------------------	-------------------	----	---

FIT, fecal immunochemical test; LR, logistic regression; ANN, artificial neural network; CRC, colorectal cancer; CI, confidence interval. <sup>†</sup> Subjects positive in the first FIT and attending colonoscopy; <sup>‡</sup> Subjects positive in any FIT and attending colonoscopy.