Gastroenterology 2014;147:1317–1326

Difference in Performance of Fecal Immunochemical Tests With the Same Hemoglobin Cutoff Concentration in a Nationwide Colorectal Cancer Screening Program

Tsung-Hsien Chiang,[1,2,3] Shu-Lin Chuang,[4] Sam Li-Sheng Chen,[5] Han-Mo Chiu,[1,4] Amy Ming-Fang Yen,[5] Sherry Yueh-Hsia Chiu,[6] Jean Ching-Yuan Fann,[7] Chu-Kuang Chou,[1,8] Yi-Chia Lee,[1,4] Ming-Shiang Wu,[1] and Hsiu-Hsi Chen[4]

[1]Department of Internal Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan; [2]Department of Integrated Diagnostics and Therapeutics, National Taiwan University Hospital, Taipei, Taiwan; [3]Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan; [4]Graduate Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan; [5]School of Oral Hygiene, College of Oral Medicine, Taipei Medical University, Taipei, Taiwan; [6]Department and Graduate Institute of Health Care Management, Chang Gung University, Tao-Yuan, Taiwan; [7]Department of Health Industry Management, Kainan University, Tao-Yuan, Taiwan; and [8]Division of Gastroenterology and Hepatology, Chia-Yi Christian Hospital, Chia-Yi, Taiwan

BACKGROUND & AIMS: We investigated whether 2 quantitative fecal immunochemical tests (FITs) with the same cutoff concentration of fecal hemoglobin perform equivalently in identifying patients with colorectal cancer (CRC). **METHODS:** A total of 956,005 Taiwanese subjects, 50 to 69 years old, participated in a nationwide CRC screening program to compare results from 2 FITs; 78% were tested using the OC-Sensor (n = 747,076; Eiken Chemical Co, Tokyo, Japan) and 22% were tested using the HM-Jack (n = 208,929; Kyowa Medex Co Ltd, Tokyo, Japan), from 2004 through 2009. The cutoff concentration for a positive finding was 20 μg hemoglobin/g feces, based on a standardized reporting unit system. The tests were compared using short-term and long-term indicators of performance. **RESULTS:** The OC-Sensor test detected CRC in 0.21% of patients, with a positive predictive value of 6.8%. The HM-Jack test detected CRC in 0.17% of patients, with a positive predictive value of 5.2%. The rate of interval cancer rate was 30.7/100,000 person-years among subjects receiving the OC-Sensor test and 40.6/100,000 person-years among those receiving the HM-Jack test; there was significant difference in test sensitivity (80% vs 68%, $P = .005$) that was related to the detectability of proximal CRC. After adjusting for differences in city/county, age, sex, ambient temperature, and colonoscopy quality, significant differences were observed between the tests in the positive predictive value for cancer detection (adjusted relative risk = 1.29; 95% confidence interval, 1.14–1.46) and the rates of interval cancer (0.75; 95% confidence interval, 0.62–0.92). Although each test was estimated to reduce CRC mortality by approximately 10%, no significant difference in mortality was observed when the 2 groups were compared. **CONCLUSIONS:** Different brands of quantitative FITs, even with the same cutoff hemoglobin concentration, perform differently in mass screening. Population-level data should be gathered to verify the credibility of quantitative laboratory findings.

Keywords: Population Screening; Colorectal Cancer; Screening Test Sensitivity; Interval Cancer.

Colorectal cancer (CRC) poses a major threat to global health. Because the widespread use of fecal occult-blood tests has the potential to decrease mortality from CRC,[1] use of these tests is commonly adopted as the preferred strategy for prevention. The traditional guaiac-based test is being increasingly replaced by the fecal immunochemical test (FIT), not only because the specificity of the FIT is higher, which tends to reduce false-positive cases, but also because the sampling method of the FIT is more patient-friendly. In addition, because FIT findings can be quantitated, the cutoff value for a positive test can be adjusted to accommodate budget and manpower limitations for a target population.[2–4]

In the current free-market system, different brands of FIT may be chosen for screening, especially when an organized service screening is conducted on a nationwide scale. However, different brands of FIT are commonly found to have different cutoff values because FIT units are usually expressed as the hemoglobin concentration in sampling bottle buffers, which are not exchangeable. Interpretation of test results has therefore become unnecessarily complex. Difficulties in the interpretation of test findings are currently faced in Taiwan, where a nationwide CRC screening program has been in place since 2004, with biennial FIT performed for the eligible population aged 50 to 69 years.[5] The FITs most commonly used in Taiwan are the OC-Sensor (Eiken Chemical Co, Tokyo, Japan) and the HM-Jack (Kyowa Medex Co Ltd, Tokyo, Japan) tests, which have cutoff concentrations of 100 and 8 ng hemoglobin/mL buffer, respectively.

To address problems in interpretation of test findings, an expert working group recently mandated that a

Abbreviations used in this paper: **CI, confidence interval; CRC, colorectal cancer; FIT, fecal immunochemical test; ISO, International Organization for Standardization; RR, relative risk; SR, screening rate.**

0016-5085

http://dx.doi.org/10.1053/j.gastro.2014.08.043

CLINICAL AT

standardized reporting unit system be developed that uses the hemoglobin concentration in feces instead of that in the buffer. The cutoff concentrations of the OC-Sensor and the HM-Jack tests could therefore be transformed into 20 μg hemoglobin/g feces.[6] However, no evidence currently exists to support the proposal that the same cutoff concentration in feces claimed by different laboratories results in equivalent performance as seen in population-based screening programs. To test this hypothesis, both short-term and long-term indicators of performance are needed; the former includes the positive predictive value, cancer detection rate, interval cancer rate, and test sensitivity, and the latter is based mainly on the CRC-specific mortality rate.[7]

Without a large population-based longitudinal follow-up cohort, a thorough evaluation employing all of these indicators is difficult. However, a nationwide cohort composed of nearly 1 million CRC-screened subjects recently became available in Taiwan. This cohort was therefore utilized in the present study to ascertain whether 2 different brands of FIT, which claim to have identical cutoff hemoglobin concentrations in feces, perform equivalently for mass screening. Both short-term and long-term indicators of performance were measured to test this hypothesis.

Methods

Screening Design

Beginning in 2004, the Taiwanese Nationwide CRC Screening Program invited residents aged 50 to 69 years to receive a biennial FIT.[5] The main purpose of mass screening was to reduce mortality from CRC. To cover approximately 5.5 million eligible residents in a total of 25 municipalities, the Health Promotion Administration, Ministry of Health and Welfare (formerly Bureau of Health Promotion) set the coverage rate every 2 years for each municipality according to the screening budget and manpower capacity. Mass screening, including the processes of invitation, distribution of FIT, and testing of fecal sample, the referral for colonoscopic examination, and the histopathologic diagnosis were performed in a stepwise manner at local public health units, clinics, and hospitals in each municipality, with approximately 810 screening sites participating in the program. All screening results were transmitted via a virtual private network to a central database to periodically generate standardized indicators such that central and local governments could monitor the screening performance.

Fecal Immunochemical Test Testing

The 1-day method was adopted, and participants were advised to return the specimens for testing immediately after they were taken. Quantitative FIT testing was performed at approximately 125 qualified laboratories. In addition to recording a positive or negative result, numerical data were stored in the database for possible adjustment of the cutoff hemoglobin concentration. Test results were reported to all participants by mail and/or telephone. The choice of FIT was based on the open bidding process at local Public Health Bureaus and hospitals. Two major brands of FIT accounted for approximately 82.4% of all FITs in use; these were the OC-Sensor and the HM-Jack tests with the respective cutoff concentrations of 100 and 8 ng hemoglobin/mL buffer. The cutoff concentrations were determined by the Health Promotion Administration and based on the following calculation[6]:

$$\mu\text{g hemoglobin/g feces} = \frac{(\text{ng hemoglobin/mL}) \times (\text{volume of the device buffer in mL})}{(\text{mass of feces collected in mg})}$$

Because the mass of feces collected and volume of the device buffer were claimed as 10 mg and 2 mL, respectively, for OC-Sensor and 0.5 mg and 1.25 mL, respectively, for HM-Jack, the cutoff hemoglobin concentrations in buffer for both tests were equivalent to 20 μg hemoglobin/g feces.

To monitor quality control within individual laboratories, the Health Promotion Administration has authorized the Taiwan Society of Laboratory Medicine to provide these laboratories with hemoglobin solutions and hemoglobin-spiked, stool-like matrix samples to test occult blood using both FITs every 6 months. Participating laboratories were required to analyze these test materials and return the findings for evaluation. Only accredited laboratories with findings that met the requirements of the International Organization for Standardization 15189 could participate in the nationwide program.

Confirmatory Diagnosis

A participant with a positive test was referred to one of approximately 485 hospitals for the confirmatory diagnosis with either a total colonoscopy or sigmoidoscopy plus barium enema. Details regarding size, location, and histopathology for colonic neoplasms were recorded. The histopathology of a colorectal neoplasm was classified according to the criteria of the World Health Organization.[8]

Performance Indicators of Fecal Immunochemical Test

Test performance was evaluated based on data from the prevalence screening. Short-term indicators included positive predictive value for cancer detection (number with cancer/total number of diagnostic endoscopies) and cancer detection rate (number with cancer/tested population). The detection of advanced adenoma, which was defined as an adenoma of $\geq$10 mm in diameter or having a villous component or high-grade dysplasia, was included in the calculations for the above indicators. The per-person analysis was used for both the CRC (ie, an individual discovered with metachronous cancers counted as one individual with cancer) and advanced adenoma (ie, the most advanced finding being an advanced adenoma). Short-term indicators also included the interval cancer rate (number of invasive cancers diagnosed after a negative FIT and $<$2 years to the next screen/total person-years at risk). To ascertain the occurrence of incident CRC, the screening database was linked with the Taiwan Cancer Registry, a nationwide program with high coverage (99%; each hospital mandated to report all cases of CRC) and high accuracy (percentage of death-certificate–only cases of $<$1% for CRC).[9] The indicator of test sensitivity was generated from the number of interval cancers using the proportional incidence method based on age- and sex-specific incidence rates derived from the Taiwan Cancer Registry. Adjustments were also made for the variation of sojourn

time during which CRC remained in the preclinical detectable phase.[10,11] The following equation was used:

$$\text{Sensitivity} = \frac{1 - I_T/I}{1 - \frac{1}{T}\int_0^T F(t)dt}$$

where I_T is the incidence of interval cancer in time T after the first screening, I is the baseline incidence in the absence of screening, and $F(t)$ is the probability distribution function of the preclinical detectable phase. The calculation is detailed in Supplementary Table 1.

As the number of eligible population was large, the phase-in approach was used by the nationwide screening program for gradual expansion of the coverage rate year by year. Person-years for each individual were calculated from the date of entry to the end of follow-up, which was defined as the earlier of the occurrence of an event or the end of the study in December 31, 2009.

Statistical Analysis

Differences in baseline characteristics between the 2 screened populations were determined by applying the Student t or χ^2 test. For the univariate analyses of test performance, the 2- sample proportion test was used to compare the 2 FITs with respect to the positive rate, referral rate for confirmatory diagnosis, positive predictive value, and cancer and advanced adenoma detection rates. For the comparisons of interval cancer rate and test sensitivity, the Poisson method was used. Because advanced age and male sex are known to be risk factors for colorectal neoplasms,[12] results stratified according to these 2 factors are also reported.

It was considered essential to validate the results of FIT performance by adjusting for influences other than brand of FIT, such as age, sex, referral rate for confirmatory endoscopy, city/county, ambient temperature during sampling, transport and storage before analysis, and the quality of colonoscopy (for positive predictive value and detection rate), each of which could lead to a difference in the detection of CRC between the 2 screened populations. To this end, a multivariable Poisson regression model with the outcome variables of positive predictive values for advanced adenoma detection and cancer detection, advanced adenoma and cancer detection rates, and interval cancer rate, respectively, was applied with results expressed as the adjusted relative risk (RR) and the corresponding 95% confidence interval (CI). Average monthly ambient temperature data were obtained from the Central Weather Bureau.

For the long-term indicator of CRC mortality, the screening database was linked with the National Mortality Registry of Taiwan to ascertain CRC-specific death during the period of 2004–2009 in order to calculate the CRC-specific mortality rate (number of deaths attributed to the colorectal cancer/total person-years at risk) for both FITs. The death certificate in Taiwan was issued by the physician in charge who judged the disease or condition directly responsible for the death and recorded this information; the certificate was reviewed and coded at the central government according to the ICD-9. The major error rate (ie, incorrect causal sequence reported or only mechanism of death reported) was approximately 9%.[13] The Cox proportional hazards model was used to estimate the relative mortality between the 2 tests by taking into account the differences in CRC-specific death and follow-up time between the 2 screened populations, right censoring at the last day of follow-up, or the competing cause of death. The results were expressed as hazard ratios and the corresponding 95% CIs.

In addition to the relative mortality between the 2 FITs, the absolute mortality reduction for each FIT was estimated and compared with nonparticipants with the adjustment of self-selection bias.[14] The following equation was applied:

$$\begin{aligned}\text{RR}_{\text{adjusted for self-selection bias}} = {} & \text{Screening rate (SR)} \\ & \times \text{RR}_{\text{participants/uninvited}} + (1 - \text{SR}) \\ & \times \text{RR}_{\text{non-participants/uninvited}}\end{aligned}$$

The calculation is detailed in Supplementary Tables 2–4.

Because the stage and location of screen-detected and interval cancers are of clinical significance,[15] a subsidiary analysis was performed and a comparison was made between the 2 tests using the χ^2 test. Cancer was staged according to the American Joint Committee on Cancer 7th staging system.[16] The colon above the level of the splenic flexure (including the splenic flexure) was defined as the proximal colon. When concurrent proximal and distal cancers were present, subjects were placed into the distal colon category. All statistical analyses were performed using SAS version 9.2 (SAS Institute, Cary, NC). All P values were 2-sided and $P < .05$ was considered to indicate statistical significance.

Results

Baseline Characteristics

Between January 1, 2004 and December 31, 2009, a total of 956,005 subjects underwent screening. Among them, 747,076 (78%) and 208,929 (22%) received the OC-Sensor and HM-Jack tests, respectively; their baseline data according to demographic characteristics, geography, and temperature, and characteristics of the confirmatory diagnosis are presented in Table 1. Small differences, which were statistically significant owing to the large sample size, were observed with respect to sex, follow-up time, confirmatory examination tool, colonoscopy adenoma detection rate, and colonoscopy advanced adenoma detection rate. Differences were more prominent in the geographic areas and the hospital levels where confirmatory diagnoses were performed.

Positivity Rate and Referral Rate

As shown in Table 2, positivity rates were similar between the 2 tests (3.8% vs 3.9%), but the confirmatory examination rate was higher for those who received HM-Jack (80.9% vs 85.3%). As expected, positivity rates were higher for males and those of older age as compared with the total population group. These findings were unchanged regardless of adjustments for sex and age distributions (data not shown).

The effect of ambient temperature on FIT positivity was also evaluated. For the temperature ranges of 10–14°C, 15–19°C, 20–24°C, and ≥25°C, the positivity rates for OC-Sensor were 5.6%, 4.4%, 3.9%, and 3.6%, respectively,

Table 1. Baseline Characteristics of the Screened Population

Characteristics	FIT 1[a] (n = 747,076)	FIT 2[b] (n = 208,929)	*P* value
Demographic characteristics			
Age, years, mean ± SD	58.42 ± 5.80	57.89 ± 5.76	.95
Sex, n (%)			
Male	286,408 (38.3)	80,821 (38.7)	.004
Female	460,668 (61.7)	128,108 (61.3)	
Follow-up time with the end point of mortality, *y*, mean ± SD	3.17 ± 1.55	3.00 ± 1.26	<.001
Geography and temperature			
Geographic area, n (%)			
Northern area	392,119 (52.5)	40,631 (19.4)	<.001
Central area	107,537 (14.4)	105,456 (50.5)	
Southern area	207,929 (27.8)	49,509 (23.7)	
Eastern area and offshore island	39,491 (5.3)	13,333 (6.4)	
Ambient temperature, °*C*, mean ± SD			
Overall	26.05 ± 3.19	26.06 ± 3.08	.99
Spring	21.86 ± 2.55	22.20 ± 2.08	.92
Summer	27.52 ± 1.82	27.31 ± 1.94	.94
Autumn	27.47 ± 1.72	27.73 ± 1.52	.91
Winter	21.72 ± 2.64	20.13 ± 2.47	.66
Confirmatory examination characteristics			
Time to confirmatory examination, months, mean ± SD	1.18 ± 1.79	1.02 ± 1.41	.94
Hospital level for confirmatory diagnosis, n (%)			<.001
Medical center	5958 (26.2)	1194 (17.2)	
Regional hospital	9792 (43.1)	3569 (51.6)	
Local hospital and clinic	5787 (25.4)	1888 (27.3)	
Nonspecified	1199 (5.3)	272 (3.9)	
Confirmatory examination tool, n (%)			<.001
Colonoscopy	19,599 (86.2)	6178 (89.2)	
Sigmoidoscopy ± barium enema	3063 (13.5)	744 (10.8)	
Missing data	74 (0.3)	1 (0.0)	
Screened-detected cancer, n (per 1000)	1546 (2.1)	359 (1.7)	<.001
Colonoscopic quality indicator (%)			
Cecal intubation rate[c]	79.8	79.6	.76
Adenoma detection rate[d]	45.6	43.0	<.001
Advanced adenoma detection rate[d]	15.0	12.7	<.001
Resection rate of <2 cm adenoma[e]	83.5	80.0	.61

[a]FIT 1 = OC-Sensor.
[b]FIT 2 = HM-Jack.
[c]Cecal intubation rate was defined as the number of subjects with cecal intubation/the number of subjects screened with colonoscopy.
[d](Advanced) adenoma detection rate was defined as the number of subjects with at least one detected (advanced) adenoma/the number of subjects positive to FIT having attended a colonoscopy.
[e]Resection rate of <2 cm adenoma was defined as the number of subjects with resection of adenoma/the number of subjects with at least one detected <2 cm adenoma having attended a colonoscopy.

and for HM-Jack were 5.5%, 3.8%, 4.7%, and 3.6%, respectively, revealing an inverse association ($P < .001$) between FIT positivity and ambient temperature.

Positive Predictive Value and Detection Rate

The OC-Sensor test detected CRC in 0.21% of patients, with a positive predictive value of 6.8%. The HM-Jack test detected CRC in 0.17% of patients, with a positive predictive value of 5.2%. The positive predictive value and the cancer detection rate were significantly higher with OC-Sensor than with HM-Jack (Table 3). Positive predictive values and cancer detection rates were also higher for male sex and older age groups as compared with the total population group. When advanced adenoma was used as the index lesion, a higher positive predictive value was seen with OC-Sensor as compared with HM-Jack, but advanced adenoma detection rates were similar between the 2 tests.

Interval Cancer Rate and Test Sensitivity

As shown in Table 4, the interval cancer rate for OC-Sensor was lower than that for HM-Jack (30.7 vs 40.6 per 100,000 person-years), resulting in a significant difference in test sensitivities (80% vs 68%; $P = .005$). The test sensitivity for each FIT was, however, similar among different subgroups stratified according to sex and age.

To consider adherence to the screening process, the 2-year sensitivity of the screening program was evaluated by including into the calculation of interval cancers those individuals who had positive FIT findings, followed by a negative assessment or no additional assessment.[17] Using

Table 2. Numbers of Tested Population, Positive Tests, and Confirmatory Diagnoses Stratified by the Age, Sex, and Brands of Quantitative Fecal Immunochemical Tests

	Tested population		Positive test		Positivity rate, %		Diagnostic examination		Referral rate for diagnostic examination, %	
Brands of FIT	FIT 1[a]	FIT 2[b]	FIT 1	FIT 2	FIT 1	FIT 2	FIT 1	FIT 2	FIT 1	FIT 2
Male										
50–59 years	169,711	48,645	6876	1866	4.1[c]	3.8[c]	5605	1575	81.5[c]	84.4[c]
60–69 years	116,697	32,176	6834	1927	5.9	6.0	5528	1639	80.9[d]	85.1[d]
Subtotal	286,408	80,821	13,710	3793	4.7	4.7	11,133	3214	81.2[d]	84.7[d]
Female										
50–59 years	302,278	85,471	8380	2555	2.8[d]	3.0[d]	6823	2192	81.4[d]	85.8[d]
60–69 years	158,390	42,637	6016	1773	3.8[d]	4.2[d]	4780	1517	79.5[d]	85.6[d]
Subtotal	460,668	128,108	14,396	4328	3.1[d]	3.4[d]	11,603	3709	80.6[d]	85.7[d]
Both sexes										
50–59 years	471,989	134,116	15,256	4421	3.2	3.3	12,428	3767	81.5[d]	85.2[d]
60–69 years	275,087	74,813	12,850	3700	4.7[d]	5.0[d]	10,308	3156	80.2[d]	85.3[d]
Total	747,076	208,929	28,106	8121	3.8	3.9	22,736	6923	80.9[d]	85.3[d]

[a]FIT 1, OC-Sensor.
[b]FIT 2, HM-Jack.
[c]$P < .05$ or [d]$P < .01$ in the comparison between FIT 1 and FIT 2.

this approach, a significant difference was again observed between the 2 FITs (OC-Sensor: 77%; 95% CI, 73%–81% vs HM-Jack: 67%; 95% CI, 60%–75%; $P = .027$).

Multivariate Analyses

Taking into account the differences in baseline characteristics of the 2 screened populations, multivariate analyses with the adjustments of demographics, geography, and temperature, and hospital levels (an indicator for the quality of confirmatory diagnosis as shown in Supplementary Table 5) were performed. As shown in Table 5, findings were remarkably similar to those obtained from the univariate analyses: a higher positive predictive value for cancer detection and a lower interval cancer rate were noted for OC-Sensor as compared with HM-Jack, with the exception that no significant difference in the cancer detection rate was observed. With respect to detection of advanced adenoma, the positive predictive value remained higher for OC-Sensor as compared with HM-Jack, but the advanced adenoma detection rate was similar for the 2 tests.

Colorectal Cancer Mortality

Regarding relative mortality rates between the 2 screened populations, the crude and adjusted (for age and

Table 3. Positive Predictive Values and Detection Rates for the Advanced Adenoma and Colorectal Cancer According to the Age, Sex, and the Brands of Quantitative Fecal Immunochemical Tests

	Positive predictive value, %				Detection rate (per 1000)			
	Advanced adenoma		Colorectal cancer		Advanced adenoma		Colorectal cancer	
Brands of FIT	FIT 1[a]	FIT 2[b]	FIT 1	FIT 2	FIT 1	FIT 2	FIT 1	FIT 2
Male								
50–59 years	17.1	15.8	5.9	4.8	5.6	5.1	1.9	1.5
60–69 years	17.5	16.3	9.4 [c]	7.0[c]	8.3	8.3	4.4	3.6
Subtotal	17.3	16.1	7.6[d]	5.9[d]	6.7	6.4	3.0[d]	2.4[d]
Female								
50–59 years	8.7[c]	6.7 [c]	5.3 [c]	4.1[c]	2.0	1.7	1.2	1.0
60–69 years	10.0	9.3	7.1 [c]	5.3[c]	3.0	3.3	2.2	1.9
Subtotal	9.2[c]	7.7 [c]	6.0[d]	4.6[d]	2.3	2.2	1.5	1.3
Both sexes								
50–59 years	12.5[d]	10.5[d]	5.5[c]	4.4[c]	3.3	3.0	1.5	1.2
60–69 years	14.0	12.9	8.3[d]	6.2[d]	5.3	5.5	3.1[c]	2.6[c]
Total	13.2[d]	11.6[d]	6.8[d]	5.2[d]	4.0	3.8	2.1[d]	1.7[d]

[a]FIT 1, OC-Sensor.
[b]FIT 2, HM-Jack.
[c]$P < .05$ or [d]$P < .01$ in the comparison between FIT 1 and FIT 2.

Table 4. Comparisons of the Number of Interval Cancer, Interval Cancer Rate, and Test Sensitivity Between 2 Quantitative Fecal Immunochemical Tests

	Person-year at risk[a]	No. of ICs	Incidence of IC (expected incidence in the absence of screening)[b]	Proportional incidence	1-proportional incidence, % (95% CI)	Test sensitivity, % (95% CI)[c]
FIT 1[d]						
Male						
50–59 years	328,335	85	25.9 (72.9)	0.36	64 (57–73)	75 (66–85)
60–69 years	235,674	141	59.8 (177.5)	0.34	66 (60–73)	79 (71–86)
Subtotal	564,009	226	40.1 (116.6)	0.34	66 (61–71)	77 (71–83)
Female						
50–59 years	607,842	97	16.0 (53.1)	0.30	70 (63–78)	82 (74–92)
60–69 years	328,778	137	41.7 (129.4)	0.32	68 (62–75)	79 (72–87)
Subtotal	936,620	234	25.0 (79.9)	0.31	69 (64–74)	80 (75–86)
Both sexes						
50–59 years	936,177	182	19.4 (62.9)	0.31	69 (64–75)	81 (74–88)
60–69 years	564,452	278	49.3 (152.6)	0.32	68 (63–72)	80 (74–85)
Total	1,500,629	460	30.7 (96.6)	0.32	68 (65–72)	80 (76–84)[e]
FIT 2[f]						
Male						
50–59 years	69,740	23	33.0 (72.9)	0.45	55 (42–72)	64 (49–84)
60–69 years	50,054	37	73.9 (177.5)	0.42	58 (47–72)	69 (56–84)
Subtotal	119,794	60	50.1 (116.6)	0.43	57 (48–67)	67 (57–79)
Female						
50–59 years	127,145	32	25.2 (53.1)	0.47	53 (41–67)	62 (48–78)
60–69 years	68,482	36	52.6 (129.4)	0.41	59 (48–73)	71 (58–87)
Subtotal	195,627	68	34.8 (79.8)	0.44	56 (48–66)	67 (57–78)
Both sexes						
50–59 years	196,885	55	27.9 (62.9)	0.44	56 (47–66)	65 (55–78)
60–69 years	118,536	73	61.6 (152.6)	0.40	60 (52–69)	71 (61–82)
Total	315,421	128	40.6 (96.6)	0.42	58 (52–65)	68 (61–76)[e]

IC, interval cancer.
[a]The interval cancer was defined as a cancer that developed in the interval of 2 years after a negative FIT result. For those who had >2 years of follow-up but did not receive the subsequent screening, their follow-up time was set at 2 years in the calculation of person-years at risk.
[b]Per 100,000 person-years.
[c]The test sensitivity was adjusted for sojourn time when colorectal cancer was in the preclinical detectable phase; the calculation is detailed in Supplementary Table 1.
[d]FIT 1, OC-Sensor.
[e]$P < .01$ in the comparison between FIT 1 and FIT 2.
[f]FIT 2, HM-Jack.

sex) hazard ratios were estimated to be 1.21 (95% CI, 0.91–1.61) and 1.22 (95% CI, 0.92–1.63), respectively, when OC-Sensor was compared with HM-Jack; the difference between the 2 groups was not significant. Regarding the absolute mortality reduction with the adjustment of self-selection bias, the results were 11% (95% CI, 6%–16%) and 13% (95% CI, 7%–18%), respectively, for the OC-Sensor and HM-Jack, as compared with nonparticipants, given the screening rate of 21.4% during the study period; the difference between the 2 FITs remained nonsignificant ($P = .20$).

Stage and Location of Screen-Detected Cancer and Interval Cancer

Findings are presented in Table 6. Regarding the cancer stage for the overall population, the proportions of stage 0–I CRC were 21.1%, 47.3%, and 35.5% for non–screen-detected cancer, screen-detected cancer, and interval cancer, respectively. Regarding the 2 FITs, stage 0–I CRC accounted for 47.5% and 46.1% of screen-detected cancers for OC-Sensor and HM-Jack, respectively; this difference was not significant ($P = .67$). With regard to interval cancer, no significant differences ($P = .62$) in the distributions of cancer stage were observed between the 2 tests. For both tests, the test sensitivities for stage 0–I and stage II–IV CRCs were estimated to be 62% (95% CI, 60%–64%) and 91% (95% CI, 90%–92%), respectively.

Regarding the location of CRC in the overall population, the proportions of proximally located CRC were 23.4%, 27.2%, and 23.8% for non–screen-detected cancer, screen-detected cancer, and interval cancer, respectively. Regarding the 2 FITs and the location of screen-detected cancer, a slightly higher percentage of proximally located CRC was observed for OC-Sensor as compared with HM-Jack (28.1% vs 23.4%; $P = .06$). Concerning the 2 FITs and the location of interval cancer, a significantly higher percentage

Table 5. Comparisons of the Test Performance Between 2 Quantitative Fecal Immunochemical Tests Using the Poisson Regression Models

Model	Relative risk	95% CI
Positive predictive value for advanced adenoma detection		
Model 1[a]		
FIT[b] 1 vs FIT 2[c]	1.14	1.05–1.23[d]
Model 2[e]		
FIT 1 vs FIT 2	1.13	1.03–1.24[d]
Age 60–69 vs 50–59 years	1.09	1.02–1.17[d]
Male vs female	1.91	1.79–2.04[d]
Mean ambient temperature,°C	1.00	0.99–1.01
Medical center/regional hospital vs local hospital/clinic	1.00	0.91–1.11
Advanced adenoma detection rate		
Model 1		
FIT 1 vs FIT 2	1.06	0.98–1.15
Model 2		
FIT 1 vs FIT 2	0.99	0.85–1.15
Age 60–69 vs 50–59 years	1.07	0.96–1.19
Male vs female	2.04	1.84–2.27[d]
Mean ambient temperature,°C	0.99	0.97–1.00
Medical center/regional hospital vs local hospital/clinic	0.93	0.81–1.08
Positive predictive value for cancer detection		
Model 1		
FIT 1 vs FIT 2	1.30	1.15–1.46[d]
Model 2		
FIT 1 vs FIT 2	1.29	1.14–1.46[d]
Age 60–69 vs 50–59 years	1.45	1.31–1.60[d]
Male vs female	1.24	1.14–1.35[d]
Mean ambient temperature,°C	1.02	1.00–1.03[d]
Medical center/regional hospital vs local hospital/clinic	1.40	1.26–1.55[d]
Cancer detection rate		
Model 1		
FIT 1 vs FIT 2	1.20	1.07–1.35[d]
Model 2		
FIT 1 vs FIT 2	1.02	0.84–1.24
Age 60–69 vs 50–59 years	1.44	1.26–1.66[d]
Male vs female	1.51	1.34–1.69[d]
Mean ambient temperature,°C	1.00	0.98–1.02
Medical center/regional hospital vs local hospital/clinic	1.32	1.12–1.56[d]
Interval cancer rate		
Model 1		
FIT 1 vs FIT 2	0.76	0.62–0.92[d]
Model 2		
FIT 1 vs FIT 2	0.75	0.62–0.92[d]
Age 60–69 vs 50–59 years	2.39	2.02–2.81[d]
Male vs female	1.47	1.25–1.73[d]
Mean ambient temperature,°C	1.00	0.97–1.03

[a]Model 1: the crude Poisson regression model; model 2: the multivariate Poisson regression model adjusted for the city/county clustering, age, and sex distributions, the monthly mean ambient temperature (a quantitative variable) for the positive predictive value and detection rate.
[b]FIT 1, OC-Sensor.
[c]FIT 2, HM-Jack.
[d]$P < .05$.
[e]Model 2: the multivariate Poisson regression model adjusted for the hospital levels (a dichotomous predictor to represent the colonoscopy quality) for the positive predictive value and detection rate.

Table 6. Comparisons of Cancer Stage and Location Between 2 Quantitative Fecal Immunochemical Tests

	Screen-detected cancer, n (%)	Interval cancer, n (%)
Cancer stage		
FIT 1[a]		
0	146 (12.2)	39 (10.2)
I	423 (35.3)	101 (26.5)
II	262 (21.9)	71 (18.6)
III	278 (23.2)	107 (28.1)
IV	88 (7.4)	63 (16.6)
Total	1197 (100)	381 (100)
FIT 2[b]		
0	38 (13.4)	10 (9.9)
I	93 (32.7)	21 (20.8)
II	73 (25.7)	20 (19.8)
III	61 (21.5)	35 (34.7)
IV	19 (6.7)	15 (14.8)
Total	284 (100)	101 (100)
Cancer location		
FIT 1		
Proximal colon	435 (28.1)	100 (22.0)
Distal colon	1111 (71.9)	355 (78.0)
Total	1546 (100)	455 (100)
FIT 2		
Proximal colon	84 (23.4)	35 (31.0)
Distal colon	275 (76.6)	78 (69.0)
Total	359 (100)	113 (100)

NOTE. In the FIT 1, 349 and 79 cases did not have the information of cancer stages of the screen-detected and interval cancers, respectively, and in the FIT 2, the respective numbers were 75 and 27. Regarding the cancer location, 5 and 15 interval cancer cases did not have information for the FIT 1 and FIT 2, respectively.
[a]FIT 1, OC-Sensor.
[b]FIT 2, HM-Jack.

of proximally located interval cancers was observed for HM-Jack as compared with OC-Sensor (31% vs 22%; $P = .044$).

Additionally, test sensitivities were estimated according to proximal and distal CRC. For OC-Sensor, the test sensitivities were 81% (95% CI, 72%–90%) and 81% (95% CI, 76%–85%) for proximal and distal CRC, respectively ($P = .99$), and for HM-Jack, the test sensitivities were 56% (95% CI, 44%–71%) and 79% (95% CI, 70%–90%), respectively ($P = .006$). When the 2 FITs were compared, a significant difference in the test sensitivity between the 2 tests was observed for proximal cancer ($P = .003$), but not for distal cancer ($P = .69$).

Discussion

In the present study, a single quantitative threshold for FIT, even when calculated as the mass of feces collected in relation to the buffer volume, was not found to function identically across products for detection of CRC. In addition, the specific epitopes of hemoglobin detected by different tests are likely to have contributed substantially to test performance. Although important differences in short-term indicators were identified, no significant difference in subsequent CRC mortality was observed between the 2 quantitative FITs mostly commonly used in Taiwan.

Features and findings of population-based screening studies based on quantitative FITs are summarized in Supplementary Table 6.[18–31] Among different brands of FIT, manufacturer cutoff concentrations range from 8 to 176 ng hemoglobin/mL buffer; however, after transformation to the proposed standardized unit, this range narrows to 15–67 μg hemoglobin/g feces. This transformation supports, in part, the use of the proposed standardized unit because the cutoff concentration of FITs is usually designed to fit the screening capacity of endoscopists, a capacity that is globally constrained. Based on OC-Sensor with the cutoff concentration of 20 μg hemoglobin/g feces, the reported positive rate, positive predictive value for cancer detection, and cancer detection rate have been found to range from 3.8% to 6.1%, from 5.8% to 10.2%, and from 2.1% to 3.3%, respectively. Such variations may reflect the observation that, without a randomized allocation, performance indicators are affected by differences in baseline characteristics.[32,33] Nonetheless, the advantage of a quantitative FIT can be found by comparing the findings of Faivre et al[26] with those of Quintero et al[28]; adjustment of the cutoff concentration from 30 to 15 μg hemoglobin/g feces yielded a higher positive rate but a lower positive predictive value. Regarding different FITs with different manufacturer cutoff concentrations, comparisons would prove difficult in the absence of an experimental design and sophisticated analysis.[27]

In the present study, test sensitivity was established to be the most objective indicator for comparison as this indicator is much less affected by the age and sex of the screened population. In a study involving Italian subjects, test sensitivities ranging from 73.2% to 82.1% were reported using different generations of FITs from the same manufacturer (OC-Hemodia or OC-Sensor-micro) with the same cutoff concentration (20 μg hemoglobin/g feces).[19–21] In the present study, in which the cutoff concentration was also 20 μg hemoglobin/g feces, a substantial difference in test sensitivities (68% vs 80%) was observed between FITs from 2 different manufacturers. This difference became especially apparent in the present study because a nationwide cohort composed of nearly 1 million CRC-screened subjects was utilized.

In the present study, the positive predictive value for either advanced adenoma or CRC differed between the 2 FITs regardless of the similar test positivity rates. This finding indicated that some analytical factor other than the mass of feces and volume of buffer may have affected the transferability between different FITs. Both FITs apply the turbidimetric immunoassay based on anti-human hemoglobin polyclonal antibodies, and manufacturers provide users with validated calibrators and reagents. These antibodies may display 100% reactivity with intact hemoglobin (calibrator); however, heterogeneous forms of hemoglobin are found in stools; both intact and partially denatured forms are observed. The degree to which available antibodies react with denatured hemoglobin has not been established. Furthermore, immunized antibodies may cross-react to some extent with human protein contaminants, with each manufacturer providing its own procedure for absorbing the nonspecific antibodies reacting with these contaminants. It therefore appears reasonable to speculate that, because they employ different antibodies, the 2 FITs examined in the present study detect different spectra of hemoglobin breakdown products. In addition, the different capacities of these FITs to detect partially degraded globin moieties are likely to result in different sensitivities for proximal CRC, where degradation of hemoglobin exceeds that of distal cancers.[34] This speculation is supported by the findings[35–37] that reduced sensitivity of FITs for proximal colon lesions is related to hemoglobin breakdown during transit with loss of detectable epitopes.

Undoubtedly, the transferability of quantitative results between different FITs can be improved through use of a standardized reporting unit system; however, findings of the present study reveal that current systems are not adequate for this purpose. In particular, antibodies provided by manufacturers of FITs are likely to differ considerably. To address this problem, the World Endoscopy Organization has proposed that an independent calibration process of analytical performance is needed, in which the system under investigation is compared with an internationally accepted hemoglobin standard (eg, artificial stool material).[38,39] Findings of the present report support this proposal.

Strengths of the present study include the large sample size, long follow-up time, execution on a nationwide scale, and registry of cancer incidence and mortality, such that both short-term and long-term indicators could be evaluated. In addition to highlighting the need to improve the capacity of FITs to detect proximal CRC, findings of the present study support the findings of others[40] that hemoglobin concentrations fall at higher ambient temperatures; the latter indicates the need to improve the stability of hemoglobin molecules present in fecal samples before conducting measurements. However, certain limitations of the present study should be noted. First, this study was not a randomized trial; the higher adherence rate of subjects receiving HM-Jack for diagnostic examination may have attenuated the differences in the advanced adenoma detection rate and cancer detection rate between this group and those receiving OC-Sensor. In addition, their shorter follow-up time, which was related to the later marketing and selling of HM-Jack in Taiwan, may have led to an underestimation of the difference in test sensitivity between the 2 FITs. Although regression analysis was employed in an attempt to address the baseline difference between the 2 groups, the absolute differences in test performance were small and residual confounding from measured or

unmeasured factors cannot be excluded. Second, given the quantitative nature of this study, the possibility that some laboratories have adjusted the cutoff concentrations for both tests according to local screening capacities cannot be excluded. However, results in the conventional ranges of 50–100 ng hemoglobin/mL buffer for OC-Sensor and 8–12 ng hemoglobin/mL buffer for HM-Jack accounted for only 3% of both measures in the present study, and almost all interval cancers were below the defined cutoff concentrations and unlikely to alter the findings. Third, during this extended study period, the manufacturers may have modified the composition of test reagents without informing users of the potential effects of such modifications on analytical performance.[39] Additionally, the findings of this report may not apply to updated products (eg, HM-Jack$_{arc}$, launched in 2011 with different system, collection device, and analytical range). In a recent Italian study inviting subjects to receive both HM-Jack and OC-Sensor tests,[41] the same cutoff concentration of HM-Jack was associated with a higher test positivity rate than that associated with OC-Sensor (6.2% vs 3.5%). This observation is consistent with the findings of the present study that, even though a standardized reporting unit system was selected, identical hemoglobin thresholds performed differently between products and product performance depended on the specific mechanics of the test. Finally, although both FITs were found to be associated with reduced CRC mortality, the significant difference in test sensitivities observed between them should theoretically have been associated with different CRC mortalities. However, no difference in CRC mortality was observed. Because both tests were able to detect significant proportions (approximately 50%) of early-stage CRC and because the prognosis for advanced cancer is improved by advances in cancer treatment, it is conceivable that the follow-up time may not have been adequate for evaluation of this indicator; additional observation is needed.

In conclusion, a discrepancy in FIT performance between laboratory and population levels was observed. Different brands of FIT, which claimed the same cutoff concentration of hemoglobin in feces, performed differently in mass screening. In addition to the measurements of fecal mass collected/volume of buffer in the collection bottle, the capacities of different antibodies to detect different epitopes of degraded hemoglobin may decrease the transferability of the standardized reporting unit system. A transparent verification of the quantitative findings from use of existing FITs is therefore anticipated. For an ongoing mass screening program, the present study lends support to continued efforts to monitor test sensitivity in order to improve the effectiveness of FIT screening and thereby decrease the occurrence of interval cancer.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at http://dx.doi.org/10.1053/j.gastro.2014.08.043.

References

1. Shaukat A, Mongin SJ, Geisser MS, et al. Long-term mortality after screening for colorectal cancer. N Engl J Med 2013;369:1106–1114.
2. Imperiale TF. Noninvasive screening tests for colorectal cancer. Dig Dis 2012;30(Suppl 2):16–26.
3. **Chiang TH, Lee YC**, Tu CH, et al. Performance of the immunochemical fecal occult blood test in predicting lesions in the lower gastrointestinal tract. CMAJ 2011;183:1474–1481.
4. Lee JK, Liles EG, Bent S, et al. Accuracy of fecal immunochemical tests for colorectal cancer. Ann Intern Med 2014;160:171–181.
5. Evolution of Cancer Screening in Taiwan. Health Promotion Administration, Ministry of Health and Welfare, Taiwan (R.O.C.). Available at: http://www.hpa.gov.tw/English/index.aspx. Accessed January 8, 2014.
6. Fraser CG, Allison JE, Halloran SP, et al. A proposal to standardize reporting units for fecal immunochemical tests for hemoglobin. J Natl Cancer Inst 2012;104:810–814.
7. Benson VS, Atkin WS, Green J, et al. International Colorectal Cancer Screening Network. Toward standardizing and reporting colorectal cancer screening indicators on an international level: The International Colorectal Cancer Screening Network. Int J Cancer 2012;130:2961–2973.
8. Bosman FT, Carneiro F, Hruban RH, et al, eds. WHO Classification of Tumours of the Digestive System. IARC: Lyon, 2010.
9. Chiang CJ, Chen YC, Chen CJ, et al. Taiwan Cancer Registry Task Force. Cancer trends in Taiwan. Jpn J Clin Oncol 2010;40:897–904.
10. Day NE. Estimating the sensitivity of a screening test. J Epidemiol Community Health 1985;39:364–366.
11. Bray F, Ren JS, Masuyer E, et al. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. Int J Cancer 2013;132:1133–1145.
12. Brenner H, Hoffmeister M, Stegmaier C, et al. Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840,149 screening colonoscopies. Gut 2007;56:1585–1589.
13. Cheng TJ, Lee FC, Lin SJ, et al. Improper cause-of-death statements by specialty of certifying physician: a cross-sectional study in two medical centres in Taiwan. BMJ Open 2012;2:e001229.
14. Duffy SW, Cuzick J, Tabar L, et al. Correcting for non-compliance bias in case-control studies to evaluate cancer screening programmes. J R Stat Soc Ser C (Appl Stat) 2002;51:235–243.
15. Steele RJ, McClements P, Watling C, et al. Interval cancers in a FOBT-based colorectal cancer population screening programme: implications for stage, gender and tumour site. Gut 2012;61:576–581.
16. Edge SB, Byrd DR, Carducci MA, et al., eds. American Joint Committee on Cancer (AJCC) Cancer Staging Manual. 7th ed. New York: Springer, 2009.
17. Hakama M, Auvinen A, Day NE, et al. Sensitivity in cancer screening. J Med Screen 2007;14:174–177.
18. Zappa M, Castiglione G, Paci E, et al. Measuring interval cancers in population-based screening using different

assays of fecal occult blood testing: the District of Florence experience. Int J Cancer 2001;92:151–154.
19. Castiglione G, Visioli CB, Ciatto S, et al. Sensitivity of latex agglutination faecal occult blood test in the Florence District population-based colorectal cancer screening programme. Br J Cancer 2007;96:1750–1754.
20. Zorzi M, Fedato C, Grazzini G, et al. High sensitivity of five colorectal screening programmes with faecal immunochemical test in the Veneto Region, Italy. Gut 2011;60:944–949.
21. Campari C, Sassatelli R, Paterlini L, et al. Test and programme sensitivities of screening for colorectal cancer in Reggio Emilia. Epidemiol Prev 2011;35:118–124 [Article in Italian].
22. van Rossum LG, van Rijn AF, Laheij RJ, et al. Random comparison of guaiac and immunochemical fecal occult blood tests for colorectal cancer in a screening population. Gastroenterology 2008;135:82–90.
23. Hol L, van Leerdam ME, van Ballegooijen M, et al. Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy. Gut 2010;59:62–68.
24. Launoy GD, Bertrand HJ, Berchi C, et al. Evaluation of an immunochemical fecal occult blood test with automated reading in screening for colorectal cancer in a general average-risk population. Int J Cancer 2005;115:493–496.
25. Guittet L, Bouvier V, Mariotte N, et al. Comparison of a guaiac based and an immunochemical faecal occult blood test in screening for colorectal cancer in a general average risk population. Gut 2007;56:210–214.
26. Faivre J, Dancourt V, Denis B, et al. Comparison between a guaiac and three immunochemical faecal occult blood tests in screening for colorectal cancer. Eur J Cancer 2012;48:2969–2976.
27. Raginel T, Puvinel J, Ferrand O, et al. A population-based comparison of immunochemical fecal occult blood tests for colorectal cancer screening. Gastroenterology 2013;144:918–925.
28. **Quintero E, Castells A**, Bujanda L, et al, COLONPREV Study Investigators. Colonoscopy versus fecal immunochemical testing in colorectal-cancer screening. N Engl J Med 2012;366:697–706.
29. Zubero MB, Arana-Arri E, Pijoan JI, et al. Population-based colorectal cancer screening: comparison of two fecal occult blood test. Front Pharmacol 2014;4:175.
30. Major D, Bryant H, Delaney M, et al. Colorectal cancer screening in Canada: results from the first round of screening for five provincial programs. Curr Oncol 2013;20:252–257.
31. Chen LS, Liao CS, Chang SH, et al. Cost-effectiveness analysis for determining optimal cut-off of immunochemical faecal occult blood test for population-based colorectal cancer screening (KCIS 16). J Med Screen 2007;14:191–199.
32. Guittet L, Bailly L, Bouvier V, et al. Indirect comparison of two quantitative immunochemical faecal occult blood tests in a population with average colorectal cancer risk. J Med Screen 2011;18:76–81.
33. Brenner H, Haug U, Hundt S. Sex differences in performance of fecal occult blood testing. Am J Gastroenterol 2010;105:2457–2464.
34. Young GP, St John DJB. Selecting an occult blood test for use as a screening tool for large bowel cancer. Front Gastrointest Res 1991;18:135–156.
35. Morikawa T, Kato J, Yamaji Y. A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. Gastroenterology 2005;129:422–428.
36. Levi Z, Rozen P, Hazazi R. A quantitative immunochemical fecal occult blood test for colorectal neoplasia. Ann Intern Med 2007;146:244–255.
37. Chiu HM, Lee YC, Tu CH, et al. Association between early stage colon neoplasms and false-negative results from the fecal immunochemical test. Clin Gastroenterol Hepatol 2013;11:832–838.
38. Allison JE, Fraser CG, Halloran SP, et al. Population screening for colorectal cancer means getting FIT: the past, present, and future of colorectal cancer screening using the fecal immunochemical test for hemoglobin (FIT). Gut Liver 2014;8:117–130.
39. Rapi S, Rubeca T, Fraser CG. How to improve the performances of Fecal Immunological Tests (FIT): need for standardization of the sampling and pre-analytical phases and revision of the procedures for comparison of methods. Int J Biol Markers 2014 May 22 Epub ahead of print.
40. Grazzini G, Ventura L, Zappa M, et al. Influence of seasonal variations in ambient temperatures on performance of immunochemical faecal occult blood test for colorectal cancer screening: observational study from the Florence district. Gut 2010;59:1511–1515.
41. Rubeca T, Peruzzi B, Confortini M, et al. Overall evaluation of an immunological latex agglutination system for fecal occult blood testing in the colorectal cancer screening program of Florence. Int J Biol Markers 2012;27:e195–e3202.

Author names in bold designate shared co-first authorship.

Received April 24, 2014. Accepted August 27, 2014.

Reprint requests
Address requests for reprints to: Yi-Chia Lee, MD, PhD, Department of Internal Medicine, National Taiwan University Hospital, No. 7, Chung-Shan South Road, Taipei 10002, Taiwan. e-mail: yichialee@ntu.edu.tw; fax: +886-2-23412775.

Conflicts of interest
The authors disclose no conflicts.

Funding
This study was supported by the Health Promotion Administration, Ministry of Health and Welfare (A1011119 and A1021227). The funding source had no role in study design, data collection, analysis, or interpretation, report writing or the decision to submit this paper for publication.