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