Table 3. Factor loadings and Cronbach alpha coefficients.

		Average	SD	Factor1	Factor2	Factor3	Factor4	Factor5	Cronbach alpha coefficients
1	Perceived barrier $\alpha = .916$								0.829
	It is difficult to undergo an FOBT because the time that the medical institution is available is inconvenient	2.36	1.07	0.98					
	Having to wait for a long time makes it difficult to undergo an FOBT	2.42	1.06	0.93					
	Time and effort required for an appointment making makes it difficult to undergo an FOBT	2.37	1.01	0.89					
	Inconvenient transportation to medical institutions makes it difficult to undergo an FOBT	2.20	0.97	0.80					
	Not knowing when and where to do an FOBT	2.55	1.27	0.59					
	It costs money to do an FOBT	2.92	1.18	0.51					
2	Subjective norms $\alpha = .929$								
	I have been recommended to do an FOBT by family members (brother, sister, son, and daughter)	1.96	1.03		0.96				
	I have been recommended to do an FOBT by friends and acquaintances	1.91	0.99		0.90				
	I have been recommended to do an FOBT by my superior at work	1.90	1.01		0.89				
	I have been recommended to do an FOBT by a partner	2.10	1.11		0.79				
	I have been recommended to do an FOBT by my primary care doctor	2.07	10.7		0.74				
3	Low importance $\alpha = .826$								
	Undergoing a FOBT is not as important as dealing with other health issues	2.23	0.92			0.78			
	There is no need to do an FOBT for a few years after undergoing FOBT once	2.20	0.95			0.77			
	There is no need to do an FOBT because I have been consulting my primary care doctor	1.99	0.86			0.64			
	I can self-check my health status without doing an FOBT	2.23	0.89			0.61			
4	Descriptive norms $\alpha = .858$								
	I think women of my age have done FOBT	2.85	0.98				0.95		
	I think men of my age have done FOBT	2.95	1.01				0.93		
	I think the acquaintances in the workplace or close friends have done FOBT	2.64	1.13				0.52		
5	Uncertainty $\alpha = .611$								
	A positive result of the FOBT does not always indicate that there is a cancer	2.86	0.86					0.70	
	I think there is still a possibility that the colon cancer is overlooked even if I undergo an FOBT	3.17	0.98					0.63	

Table 4. Logistic regression of FOBT screening behavior ($n = 592$).

Variables	Range or category	% or mean (SD)	Bivariate analyses Crude OR (95% CL)	*p* value	Multivariate analyses Adjusted[a] OR (95% CL)	*p* value
FOBT screening rate[b]	Yes	44.9				
Perceived risk	2–106	30.2(19.5)	1.01(1.00–1.015)	0.115	–	–
Perceived severity	2–10	5.4(1.8)	0.92 (0.84–1.01)	0.07	–	–
Belief of FOBT						
Barrier	5–30	14.8(5.5)	0.83(0.80–0.86)	<0.001	0.87(0.84–0.92)	<0.001
Subjective norms	5–25	8.7(2.9)	1.07(1.03–1.11)	<0.001	–	–
Low importance	5–20	6.0(1.6)	0.75(0.70–0.80)	<0.001	0.91(0.84–0.99)	0.03
Objective Norms	5–15	9.9(4.6)	1.31(1.22–1.41)	<0.001	1.18(1.09–1.28)	<0.001
Uncertainty	5–10	8.4(2.8)	0.9(0.82–1.00)	0.058	–	–
CRC worry	4–20	20.1(3.4)	1.00(0.95–1.05)	0.936	–	–
Self-efficacy of health	8–40	21.8(8.8)	0.98(0.95–1.01)	0.189	–	–
Knowledge of CRC	0–13	5.9(3.4)	1.06 (1.01–1.11)	0.015	–	–

[a]Rate of participants undergoing FOBT in the past year.
[b]Adjusted for age, gender, cancer insurance, drinking habits, exercising habits, exposed CRC information.

which many previous studies indicated as obstructive factors for intention for or compliance with CRC screening (James, Campbell, & Hudson, 2002; Jones, Woolf, et al., 2010; Kiviniemi, Bennett, Zaiter, & Marshall, 2011; Power, Miles, von Wagner, Robb, & Wardle, 2009). Moreover, we consider this indicative of an association between *perceived barrier* and FOBT screening that is reasonable given that prior Japanese cancer screening studies for gastric and breast cancer have also shown this to be a factor with a negative correlation (Seki et al., 2011; Tsubono et al., 1993). The questions included in *Low importance* related to ideas that FOBT is less important than the screening of other illnesses, and that regular visits to the primary care doctor and self-examination are sufficient for prevention of CRC. For these items related to the importance of FOBT, attitudinal barriers such as 'I am in good health' and 'I do not undergo FOBT because there is no health problem (especially in the stomach)' were considered as a part of *Perceived barrier* or *Cons* (Beeker, Kraft, Southwell, & Jorgensen, 2000; Jones, Devers, et al., 2010; Klabunde et al., 2005; Liang et al., 2006; Matsuda et al., 2012). Kandula, Wen, Jacobs, and Lauderdale (2006) investigated cultural influences on cancer screening behavior based on the fact that Asian Americans have lower rates of CRC screening compared to non-Hispanic whites. He observed a trend of thinking in cancer screening that affects the low FOBT screening rate among Asian Americans, which considers cancer screening as a reaction to certain perceivable symptoms and not as a proactive measure to prevent cancer when the subjects are asymptomatic. Moreover, recent research in Japan revealed that 'not having time to get screened' (Cabinet Office, Government of Japan, 2013) is the most common reason for not receiving cancer screening.

In our study, only *Descriptive norms* was identified as a promoting factor of CRC screening. Honda and Kagawa-Singer (2006) conducted research on the associated factors of CRC screening adherence, targeting Japanese people living in the USA. The result of this research discussed that

Subjective norms (family or friends) are the strongest predictive factors because sharing of value and attitude is enhanced by the influence of Japanese culture. However, this research measured only the Subjective norms and not the Objective norms. The three previously conducted meta-analyses on Social norms (Armitage & Conner, 2001; Manning, 2009; Rivis & Sheeran, 2003) indicated that the Descriptive norms are strongly associated with behavior and intention, and it revealed that Subjective norms and Injunctive norms (the perception that others will approve or disapprove of what one dose) are less influential than Descriptive norms. Therefore, it is reasonable for Descriptive norms to remain in the present study, which employed both Subjective and Descriptive variables. Our study indicates that Descriptive norms influence health risk behaviors more strongly than health-promoting behaviors. Since cancer screening is considered to be secondary prevention (risk behavior), the result of Descriptive norms being significant is reasonable. However, among the studies on Normative factors in cancer screening, some studies have not shown an association between Descriptive norms and intention in CRC screening (Smith-McLallen & Fishbein, 2008). In addition, studies acknowledging the significant association between Subjective norms and intention report a possibility that Subjective norms become a predictable factor in certain subgroups (Sieverding et al., 2010). For example, subjective norms may function as a trigger of a behavior in a culture that strongly recognizes the value of cancer screening, such as that in the USA (Schwartz, Woloshin, Fowler, & Welch, 2004) where Honda conducted research. On the other hand, the Descriptive norms represent a behavior in line with others rather than the value of cancer screening, which could be caused by the influences of collectivism in Japanese culture. Caution is necessary before interpreting cultural differences because there have been researchers that deny the influence of cultural differences such as collectivism and individualism on human behavior (Takano & Sogon, 2009). However, we consider that future studies are necessary, which bear in mind these cultural differences and Social norms.

Based on these results, we believe that providing the public with messages and information about cancer screening, particularly stressing the following concepts, will be effective in order to improve the CRC screening rate: (1) everyone should be regularly screened for cancer, (2) ease of obtaining FOBT, and (3) the importance of FOBT. Furthermore, knowledge of CRC screening is perceived to be a precognition in an effort to effectively intervene in behavioral change (Kiviniemi et al., 2011). The knowledge of CRC and CRC screening has been identified as a predictor of CRC screening (Ng et al., 2007; Sieverding et al., 2010; Takano & Sogon, 2009). However, an association between knowledge and behavior has yet to be identified (Liang et al., 2006; Ng et al., 2007; Subramanian et al., 2004; Weinberg et al., 2009). This study revealed an unawareness of both the CRC incidence rate increase in Japan and the link between CRC and lifestyle. We would also like to make reference to the necessity for dissemination of appropriate knowledge regarding the results of this study.

In conclusion, we would d like to discuss the limitations of this study and the challenges for the future. First, we determined whether a patient completed FOBT test or not while depending on the self-report submitted by a patient. Therefore, a research in cooperation with an institution which conducts the screening will be required in the future studies as we have not confirmed who actually underwent FOBT this time. Second, we should consider the possibility of sample bias. The FOBT screening rate in the year preceding the survey was 44.9%, which was significantly higher than the FOBT screening rate previously reported in Japan (Tomotaka et al., 2005). Behind these differences lies the fact that the FOBT is performed not only in medical checkups by municipalities targeting their citizens, but also in the complete medical checkups conducted by companies for their employees. However, because this study was conducted through the internet, there is inherent sampling bias, which indicates the interests of participants in cancer screening. We believe that the study participants readily undergo FOBT screening. As for our metrics, we modified the items used in a separate study of the beliefs of breast cancer

screening in Japan. For this reason, the items cited as perceived barriers such as a feeling of discomfort towards collecting stool samples and embarrassment (Cabinet Office, Government of Japan, 2013; Jones, Devers, et al., 2010; Kandula et al., 2006) were not included in the items relating to a specific belief regarding FOBT. We consider that we can extract barriers that are more influential to CRC screening in the future by adding the items specific to CRC screening.

Despite these limitations, this study is important in terms of revealing the psychological factors influencing CRC screening behavior, particularly in a Japanese population. Based on the results of this study, cancer screening rates of people between the ages of 40 and 60 can be increased by eliminating the barriers to CRC screening, acknowledging the importance of CRC screening, and becoming aware that CRC screening is very common among healthy people. Development of specific messages reflecting the factors outlined in this study and an intervention study testing whether the actual interventions changes screening behavior are conceivable for the future.

5. Informed consent

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Acknowledgment

We would like to express the deepest appreciation to Ms Aiko Seki, Mr Takuya Narao, and Dr Kayo Tajime and all of our coworkers whose support and discussions have helped throughout this study.

Disclosure statement

Author T Taniguchi, Author K Hirai, Author K Harada, Author Y Ishikawa, Author M Nagatsuka, Author J Fukuyoshi, Author H Arai, Author H Saito, Author Y Mizota, Author S Yamamoto, and Author D Shibuya declare that they have no conflict of interest.

Funding

This work was supported in part by National Cancer Research and Development Fund [23-A-5].

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*, 179–211.

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *The British Journal of Social Psychology*, *40*(4), 471–499. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11795063

Bandura, A. (1998). Health promotion from the perspective of social cognitive theory. *Psychology & Health*, *13*(4), 623–649.

Bandura, A. (2004). Health promotion by social cognitive means. *Health Education & Behavior*, *31*(2), 143–164.

Beeker, C., Kraft, J. M., Southwell, B. G., & Jorgensen, C. M. (2000). Colorectal cancer screening in older men and women: Qualitative research findings and implications for intervention. *Journal of Community Health*, *25*(3), 263–278. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10868818

de Bosset, V., Atashili, J., Miller, W., & Pignone, M. (2008). Health insurance-related disparities in colorectal cancer screening in Virginia. *Cancer Epidemiology Biomarkers Prevention*, *17*(4), 834–837.

Cabinet Office, Government of Japan. (2013). *Public opinion survey the cancer prevention measures* (Japanese Only). http://www8.cao.go.jp/survey/h24/h24-gantaisaku/2–2.html (2013.9.23)

Center, M. M., Jemal, A., & Ward, E. (2009). International trends in colorectal cancer incidence rates. *Cancer Epidemiology Biomarkers Prevention*, *18*, 1688–1694.

Galvin, J. E., Fu, Q., Nguyen, J. T., Glasheen, C., & Scharff, D. P. (2008). Psychosocial determinants of intention to screen for Alzheimer's disease. *Alzheimer's & Dementia*, *4*(5), 353–360.

Honda, K., & Kagawa-Singer, M. (2006). Cognitive mediators linking social support networks to colorectal cancer screening adherence. *Journal of Behavioral Medicine*, *29*(5), 449–460.

James, A. S., Campbell, M. K., & Hudson, M. A. (2002). Perceived barriers and benefits to colon cancer screening among African Americans in North Carolina: How does perception relate to screening behavior? *Cancer Epidemiology Biomarkers Prevention*, *11*(6), 529–534. Retrieved from http://cebp.aacrjournals.org/content/11/6/529.short

Janz, N. K., & Becker, M. H. (1984). The health belief model: A decade later. *Health Education & Behavior*, *11*(1), 1–47.

Jones, R. M., Devers, K. J., Kuzel, A. J., & Woolf, S. H. (2010). Patient-reported barriers to colorectal cancer screening: A mixed-methods analysis. *American Journal of Preventive Medicine*, *38*(5), 508–516. doi:10.1016/j.amepre.2010.01.021

Jones, R. M., Woolf, S. H., Cunningham, T. D., Johnson, R. E., Krist, A. H., Rothemich, S. F., & Vernon, S. W. (2010). The relative importance of patient-reported barriers to colorectal cancer screening. *American Journal of Preventive Medicine*, *38*(5), 499–507. doi:10.1016/j.amepre.2010.01.020

Kandula, N. R., Wen, M., Jacobs, E. A., & Lauderdale, D. S. (2006). Low rates of colorectal, cervical, and breast cancer screening in Asian Americans compared with non-Hispanic whites: Cultural influences or access to care? *Cancer*, *107*(1), 184–192. doi:10.1002/cncr.21968

Kiviniemi, M. T., Bennett, A., Zaiter, M., & Marshall, J. R. (2011). Individual-level factors in colorectal cancer screening: A review of the literature on the relation of individual-level health behavior constructs and screening behavior. *Psycho-Oncology*, *20*(10), 1023–1033. doi:10.1002/pon.1865

Klabunde, C. N., Vernon, S. W., Nadel, M. R., Breen, N., Seeff, L. C., & Brown, M. L. (2005). Barriers to colorectal cancer screening: A comparison of reports from primary care physicians and average-risk adults. *Medical Care*, *43*(9), 939–944.

Lee, K. J., Inoue, M., Otani, T., Iwasaki, M., Sasazuki, S., & Tsugane, S. (2007). Colorectal cancer screening using fecal occult blood test and subsequent risk of colorectal cancer: A prospective cohort study in Japan. *Cancer Detection and Prevention*, *31*(1), 3–11.

Lerman, C., Trock, B., Rimer, B. K., Jepson, C., Brody, D., & Boyce, A. (1991). Psychological side effects of breast cancer screening. *Health Psychology*, *10*(4), 259–267.

Levin, B., Lieberman, D. A., McFarland, B., Smith, R. A., Brooks, D., Andrews, K. S., ... Winawer, S. J. (n.d.). Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: A joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology. *A Cancer Journal for Clinicians*, *58*, 130–160. doi:10.3322/CA.2007.0018

Liang, S. Y., Phillips, K. A., Nagamine, M., Ladabaum, U., & Haas, J. S. (2006). Rates and predictors of colorectal cancer screening. *Preventing Chronic Disease*, *3*(4), 1–13.

Manne, S., Markowitz, A., Winawer, S., Meropol, N. J., Haller, D., Rakowski, W., & Babb, J. (2002). Correlates of colorectal cancer screening compliance and stage of adoption among siblings of individuals with early onset colorectal cancer. *Health Psychology*, *21*(1), 3–15.

Manning, M. (2009). The effects of subjective norms on behaviour in the theory of planned behaviour: A meta-analysis. *The British Journal of Social Psychology*, *48*(4), 649–705. doi:10.1348/014466608X393136

Matsuda, T., Marugame, T., Kamo, K. I., Katanoda, K., Ajiki, W., & Sobue, T. (2012). The Japan cancer surveillance research group. Cancer incidence and incidence rates in Japan in 2006: Based on data from 15 population-based cancer registries in the monitoring of cancer incidence in Japan (MCIJ) project. *Japenese Journal of Clinical Oncology*, *42*, 139–147.

Nagatsuka, M., Arai, H., & Hirai, K. (2009). Stage of change and decisional balance in medical check-ups. *Japanese Journal of Behavioral Medicine*, *15*(2), 61–68.

Nancy, N. B. et al. (2009). Association of colonoscopy and death from colorectal cancer. *Annals of Internal Medicine*, *150*(1), 1–8.

Ng, E. S., Tan, C. H., Teo, D. C., Seah, C. Y., & Phua, K. H. (2007). Knowledge and perceptions regarding colorectal cancer screening among Chinese: A community-based survey in Singapore. *Preventive Medicine*, *45*(5), 332–335. doi:10.1016/j.ypmed.2007.06.021

Peterson, N. B., Dwyer, K. A., Mulvaney, S. A., Dietrich, M. S., & Rothman, R. L. (2007). The influence of health literacy on colorectal cancer screening knowledge, beliefs and behavior. *Journal of the National Medical Association*, *99*(10), 1105–1112.

Power, E., Miles, A., von Wagner, C., Robb, K., & Wardle, J. (2009). Uptake of colorectal cancer screening: System, provider and individual factors and strategies to improve participation. *Future Oncology*, *5*(9), 1371–1388. doi:10.2217/fon.09.134

Prochaska, J. O., & Velicer, W. F. (1997). The transtheoretical model of health behavior change. *American Journal of Health Promotion*, *12*(1), 38–48.

Rawl, S., Menon, U., Champion, V., May, F., Loehrer, P., Hunter, C., & Azzouz, F. (2005). Do benefits and barriers differ by stage of adoption for colorectal cancer screening? *Health Education Research*, *20*(2), 137–148.

Rex, D. K., Johnson, D. A., Lieberman, D. A., Burt, R. Y., & Sonnenberg, A. (2000). Colorectal cancer prevention: Screening recommendations of the American College of Gastroenterology. *Americian Journal of Gastroenterology*, *95*, 868–877.

Rivis, A., & Sheeran, P. (2003). Descriptive norms as an additional predictor in the theory of planned behaviour: A meta-analysis. *Current Psychology*, *22*(3), 218–233. doi:10.1007/s12144-003-1018-2

Saito, H., Soma, Y., Koeda, J., Wada, T., Kawaguchi, H., Sobue, T., & Aisawa, T. (1995). Reduction in risk of mortality from colorectal cancer by fecal occult blood screening with immunochemical hemagglutination test. A case-control study. *International Journal of Cancer*, *61*(4), 465–469.

Salz, T., Brewer, N. T., Sandler, R. S., Weiner, B. J., Martin, C. F., & Weinberger, M. (2009). Association of health beliefs and colonoscopy use among survivors of colorectal cancer. *Journal of cancer survivorship*, *3*(4), 193–201.

Schwartz, L. M., Woloshin, S., Fowler, F. J., & Welch, H. G. (2004). Enthusiasm for cancer screening in the United States. *JAMA: The Journal of the American Medical Association*, *291*(1), 71–78. doi:10.1001/jama.291.1.71

Seki, A., Hirai, K., Nagatsuka, M., Harada, K., Arai, H., Hazama, R., Hamashima, C., Saito, H., & Shibuya, D. (2011). Measurement of attitudes toward breast cancer screening. *Journal of Health and Welfare Statistics*, *58*(2), 14–20.

Sieverding, M., Matterne, U., & Ciccarello, L. (2010). What role do social norms play in the context of men's cancer screening intention and behavior? Application of an extended theory of planned behavior. *Health Psychology*, *29*, 72–81. doi:10.1037/a0016941

Smith-McLallen, A., & Fishbein, M. (2008). Predictors of intentions to perform six cancer-related behaviours: Roles for injunctive and descriptive norms. *Psychology, Health & Medicine*, *13*(4), 389–401. doi:10.1080/13548500701842933

Subramanian, S., Klosterman, M., Amonkar, M., & Hunt, T. (2004). Adherence with colorectal cancer screening guidelines: A review. *Preventive Medicine*, *38*, 536–550.

Takano, Y., & Sogon, S. (2009). Are Japanese more collectivistic than Americans?: Examining conformity in in-groups and the reference-group effect. *Journal of Cross-Cultural Psychology*, *39*(3), 237–250. doi:10.1177/0022022107313902

Togari, T., Yamazaki, Y., Koide, S., & Miyata, A. (2006). Reliability and validity of the modified perceived health competence scale (PHCS) Japanese version. *Japanese Journal of Public Health*, *53*(1), 51–57.

Tomotaka, S., Chisato, H., Hiroshi, S., Takenobu, S., Kazuo, M., & Hiroshi, N. (2005). Japanese Guideline for Colorectal Cancer Screening. *Japanese Journal of Cancer Chemotheraphy*, *32*(6), 901–915.

Tsubono, Y., Fukao, A., Hisamichi, S., Sugawara, N., & Hosokawa, T. (1993). Health belief model and attendance at screenings for gastric cancer in a population in Miyagi, Japan. *Japanese Journal of Public Health*, *40*(4), 255–264.

Vernon, S. W., Myers, R. E., & Tilley, B. C. (1997). Development and validation of an instrument to measure factors related to colorectal cancer screening adherence. *Cancer Epidemiology, Biomarkers & Prevention*, *6*(10), 825–832.

Watanabe, R. (2003). An analysis of participation in cancer screening in Japan. *Journal of Health Care and Society*, *13*(2), 113–132.

Weinberg, D. S., Miller, S., Rodoletz, M., Egleston, B., Fleisher, L., Buzaglo, J., ... Bieber, E. (2009). Colorectal cancer knowledge is not associated with screening compliance or intention. *Journal of Cancer Education*, *24*(3), 225–232. doi:10.1080/08858190902924815

Zauber, A. G., Winawer, S. J., O'Brien, M. J., Lansdorp-Vogelaar, I., Ballegooijen, M., Hankey, B. F., ... Waye, J. D. (2012). Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *New England Journal of Medicine*, *366*, 687–696.

JJCO Japanese Journal of Clinical Oncology

Japanese Journal of Clinical Oncology, 2016, 46(2) 116–125
doi: 10.1093/jjco/hyv186
Advance Access Publication Date: 18 December 2015
Original Article

Original Article

Optimal use of colonoscopy and fecal immunochemical test for population-based colorectal cancer screening: a cost-effectiveness analysis using Japanese data

Masau Sekiguchi[1], Ataru Igarashi[2], Takahisa Matsuda[1,3,*], Minori Matsumoto[1], Taku Sakamoto[1], Takeshi Nakajima[1], Yasuo Kakugawa[1], Seiichiro Yamamoto[4], Hiroshi Saito[5], and Yutaka Saito[1]

[1]Endoscopy Division, National Cancer Center Hospital, Tokyo, [2]Graduate School of Pharmaceutical Science, The University of Tokyo, Tokyo, [3]Cancer Screening Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, [4]Public Health Policy Research Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, and [5]Screening Assessment and Management Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan

*For reprints and all correspondence: Takahisa Matsuda, Cancer Screening Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan; Endoscopy Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo, 104-0045, Japan. E-mail: tamatsud@ncc.go.jp

Received 28 July 2015; Accepted 9 November 2015

Abstract

Objective: There have been few cost-effectiveness analyses of population-based colorectal cancer screening in Japan, and there is no consensus on the optimal use of total colonoscopy and the fecal immunochemical test for colorectal cancer screening with regard to cost-effectiveness and total colonoscopy workload. The present study aimed to examine the cost-effectiveness of colorectal cancer screening using Japanese data to identify the optimal use of total colonoscopy and fecal immunochemical test.

Methods: We developed a Markov model to assess the cost-effectiveness of colorectal cancer screening offered to an average-risk population aged 40 years or over. The cost, quality-adjusted life-years and number of total colonoscopy procedures required were evaluated for three screening strategies: (i) a fecal immunochemical test-based strategy; (ii) a total colonoscopy-based strategy; (iii) a strategy of adding population-wide total colonoscopy at 50 years to a fecal immunochemical test-based strategy.

Results: All three strategies dominated no screening. Among the three, Strategy 1 was dominated by Strategy 3, and the incremental cost per quality-adjusted life-years gained for Strategy 2 against Strategies 1 and 3 were JPY 293 616 and JPY 781 342, respectively. Within the Japanese threshold (JPY 5–6 million per QALY gained), Strategy 2 was the most cost-effective, followed by Strategy 3; however, Strategy 2 required more than double the number of total colonoscopy procedures than the other strategies.

Conclusions: The total colonoscopy-based strategy could be the most cost-effective for population-based colorectal cancer screening in Japan. However, it requires more total colonoscopyprocedures than the other strategies. Depending on total colonoscopy capacity, the strategy of adding total

colonoscopy for individuals at a specified age to a fecal immunochemical test-based screening may be an optimal solution.

Key words: colorectal cancer screening, cost-effectiveness analysis, fecal immunochemical test, total colonoscopy

Introduction

Colorectal cancer (CRC) has markedly increased and is now the second most commonly diagnosed cancer and the third leading cause of cancer-related mortality in Japan (1). For the secondary prevention of CRC, a Japanese population-based CRC screening system has used the 2-day fecal immunochemical test (FIT) as a primary screening procedure on the basis of the evidence regarding its effectiveness for CRC screening (2). The effectiveness of the fecal occult blood test (FOBT) for reducing CRC-associated mortality has been clearly shown in several randomized controlled trials (3–7), whereas other case–control or cohort studies have shown the effectiveness of FIT for CRC screening and the superior sensitivity of FIT for CRC compared with that of FOBT (8–14). Japanese population-based CRC screening is offered to the entire population aged 40 years and over, and total colonoscopy (TCS) is performed for those with a positive FIT result. Recently, however, it has been reported that TCS-based CRC screening, in which TCS is performed as a primary screening procedure, is effective for reducing CRC incidence and mortality, based on long-term follow-up data in cohort studies (15,16). In this context, an analysis of the optimal combination of TCS and FIT for population-based CRC screening is required because there is yet no consensus regarding the issue.

Cost-effectiveness analysis is an essential part of the evaluation of screening strategies. Several cost-effectiveness analyses of CRC screening have been reported from the USA and several other countries (17–24). In Japan, however, there have been only limited analyses (25,26). Recently, by analyzing the TCS screening database of our institution's cancer screening division and the Japanese nationwide survey data of CRC screening, we reported that not only FIT but also TCS might be cost-effective for primary screening (27). However, the study retrospectively evaluated only the cost of identifying a CRC patient; further study using a Markov model analysis is necessary to evaluate the true cost-effectiveness of Japanese CRC screening.

In the present study, we aimed to identify the optimal combination of TCS and FIT for population-based CRC screening in the Japanese setting from the perspective of cost-effectiveness. To evaluate cost-effectiveness, we performed a Markov model analysis using Japanese clinical and cost data. To determine the optimal screening strategy, we also considered the number of TCS procedures required.

Patients and methods

Decision analytic model

We developed a state-transition Markov model that simulated the natural history of CRC development, and the actual cost-effectiveness was analyzed by Monte Carlo simulation using Tree Age Pro 2014 (TreeAge Software Inc., Williamstown, MA, USA) (28). In a Markov model, clinical situations are described in terms of discrete health states, 'Markov states,' that individuals can be in; an individual is always in one of these states, and all events of interest are modeled as transitions from one state to another. In this study, the natural history of CRC development was simulated as a transition from normal epithelium to low-risk adenomatous polyps sized 1–4 and 5–9 mm, to high-risk polyps, to CRC (from Dukes' A to Dukes' D), and ultimately to death from CRC, with reference to previous studies (17–24). Therefore, the Markov states were set as shown in Fig. 1. In addition, the detection status of colorectal polyps and CRC ('detected' or 'undetected') was considered, with CRC screening affecting the transition from 'undetected' to 'detected.' CRC was defined, according to the international classification, as a malignant epithelial tumor originating in the large bowel with invasion beyond the muscularis mucosae (29). High-risk polyps included intramucosal cancers and adenomas with a diameter ≥10 mm, with high-grade dysplasia, or with villous histology (≥25%) (30). The study setting was Japan and the initial population comprised 100 000 individuals aged 40 years who were at an average risk of CRC. The screening and analysis continued through the lifetime of the cohort. The time frame of the analysis was divided into 1 year, during which individuals were in the same health state before having the opportunity to transition to another state. The transition was governed by transition probability values mostly estimated from Japanese literature as described later. Japanese data for age-specific CRC incidence rates was the basis for determining the number of individuals in the population would develop CRC without any screening or intervention (1).

The validity of the model was assessed by comparing the lifetime cumulative risks for CRC incidence and mortality for the 40-year-old Japanese population estimated from the model of this study with those estimated from Japan's Cancer Registry and Statistics (http://gdb.ganjoho.jp/graph_db/gdb1?smTypes=67, Cancer Information Service, National Cancer Center, Japan) (1). When estimating these risks using the model, CRC screening with FIT (primary screening) and TCS (for those with a positive FIT) were considered with uptake rates set at 37 and 55% for FIT and TCS, respectively, based on the data of current Japanese uptake rates (31,32).

CRC screening strategies

To evaluate the optimal use of TCS and FIT for CRC screening, a total of three CRC screening strategies with TCS and/or FIT, including a

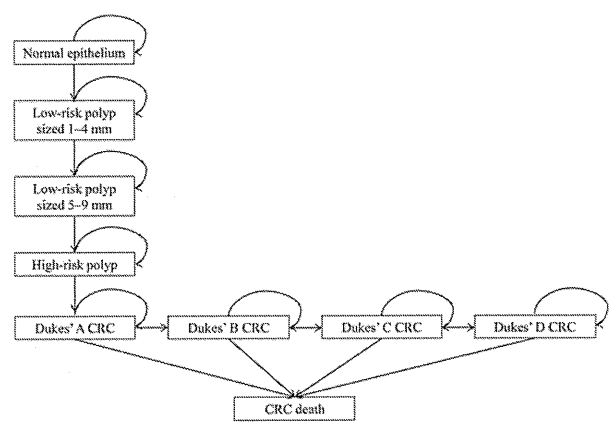

Figure 1. The natural history model of colorectal cancer. CRC, colorectal cancer.

FIT-based strategy which mostly corresponded to the current strategy of Japanese population-based CRC screening and other two strategies which used TCS more actively than the current strategy, were examined in this study (Fig. 2).

Strategy 1: a FIT-based screening strategy
The population is offered FIT at the age of 40 years. When the test is negative, it is repeated annually. Individuals with a positive FIT result are invited for TCS examination; any polyps found are removed and surveillance TCS is repeated every 3 years until no more polyps are found. When the results on TCS are normal, FIT is resumed 5 years after the TCS (Fig. 2a).

Strategy 2: a TCS-based screening strategy
The population is offered TCS as primary screening at the age of 40 years. When the test is negative, TCS is repeated 10 years later. If polyps are found, they are removed and surveillance TCS is repeated every 3 years until no more polyps are found. When the TCS results are normal, TCS is resumed 10 years later (Fig. 2b).

Strategy 3: a strategy of adding population-wide TCS for 50-year-old individuals to a FIT-based screening
This screening strategy is the same as Strategy 1 for individuals aged 40–49 years. The difference is that at the age of 50 years the whole population undergoes TCS, apart from those who underwent TCS in their 40s. After TCS, the screening continues according to the TCS results as with Strategy 1 (Fig. 2c).

Model parameters

Model parameters, including transition probabilities, test characteristics and cost, are summarized in Table 1. Most data were based on Japanese data (1,33–38), except for some data that were only available from foreign studies (20,39). The disease progression parameters from normal epithelium to colorectal polyps and cancer were calculated on the basis of the CRC incidence data from a study of 25 population-based cancer registries for the Monitoring of Cancer Incidence in Japan project (1), and the polyp prevalence data at Cancer Screening Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan (33). The possibility of new polyps developing after endoscopic removal of polyps was estimated with reference to the data from the Japan polyp study (34). The references for the data regarding other transition probabilities are provided in Table 1 (20,38,39).

With regard to the parameters of test characteristics, the sensitivities and specificities of FIT for colorectal polyps and cancer were set on the basis of data from detailed previous studies by Morikawa et al. (35,36). The sensitivities and specificities of TCS for colorectal polyps and cancer were set according to the data from the Japan polyp study (34). The possibility of complication (perforation and bleeding) following TCS were estimated from the nationwide report from the Japan Gastroenterological Endoscopy Society (37).

The cost included the screening-related cost and CRC treatment-related cost. The screening-related cost was set on the basis of Japanese national reimbursement tables. The CRC treatment-related cost was calculated from the cost of the treatment procedure, hospitalization, adjuvant chemotherapy and follow-up care on the basis of Japanese national reimbursement tables and expert discussion.

The uptake rate of each test (FIT and TCS) was also built into this analysis. The CRC screening uptake rate in Japan has been increasing, but the current rate (~30–40%) is lower than the Japanese government's target values (50%) and the cut-off value for the desirable level of the uptake rate (65%) provided in the European guidelines (31,40). These guidelines based their evidence on performance indicators for FIT on data with a FIT uptake rate of 61.5% (41). From this, it ideally appears that an uptake rate of at least 60% is required for population-based CRC screening. Thus, in the present study, all uptake rates were first set at 60% in the base case analysis and then changed in the sensitivity analyses.

Cost-effectiveness analysis

The cost-effectiveness analysis was performed from a healthcare payer's perspective. The effectiveness of screening was measured in terms of the quality-adjusted life-years (QALYs) gained. Costs and QALYs were discounted at an annual rate of 3% (42). Strategies that were more costly and less effective than other strategies were ruled out by simple dominance. Among the remaining strategies, the incremental cost-effectiveness ratio (ICER) was evaluated. ICER was determined for a strategy by comparing the additional cost and effectiveness of the strategy with those of a less costly and less effective strategy; ICER was calculated as the difference in costs divided by the difference in effectiveness.

To compare the demand for endoscopic resources between different screening strategies, the number of TCS procedures performed in each strategy was also calculated.

Sensitivity analyses

In addition to the base case analysis, scenario analyses were performed with regard to the uptake rates (10% and 100%), the initial age of screening (50 years), and the age for population-wide TCS in Strategy 3 (40–60 years). A probabilistic sensitivity analysis was performed for the parameters of transition probabilities, costs, test characteristics, uptake rates and quality of life scales. In a probabilistic sensitivity analysis, these multiple parameters were varied simultaneously. We used β distributions for the parameters for which we could acquire raw data (the denominator and numerator of parameters), including the sensitivities of FIT and TCS, the probability of perforation after TCS, and that of new polyps developing after polyp resection, and gamma distributions for the other variables with a range of ±25%. A cost-effectiveness acceptability curve was drawn to show the correlation between the probability of being chosen as the most cost-effective scenario for each strategy and the willingness-to-pay (WTP) values for one additional QALY gained. The WTP value is the maximum cost that an individual is willing to pay to gain one additional QALY, and the value varies according to country; the Japanese threshold is reported to be JPY 5–6 million per QALY gained (43).

Results

Validity of the model

The cumulative risks for CRC incidence and mortality for the Japanese 40-year-old population estimated from the Cancer Registry and Statistics and those estimated from the model are shown in Fig. 3. The risks estimated from the model generally matched those from the Cancer Registry and Statistics, particularly ≤65 years of age. After the age of 65 years, the risks estimated from the model were slightly lower than those estimated from the Cancer Registry and Statistics.

Base case analysis

The outcomes for the three screening strategies and for no screening in the base case analysis are summarized in Table 2. Without any

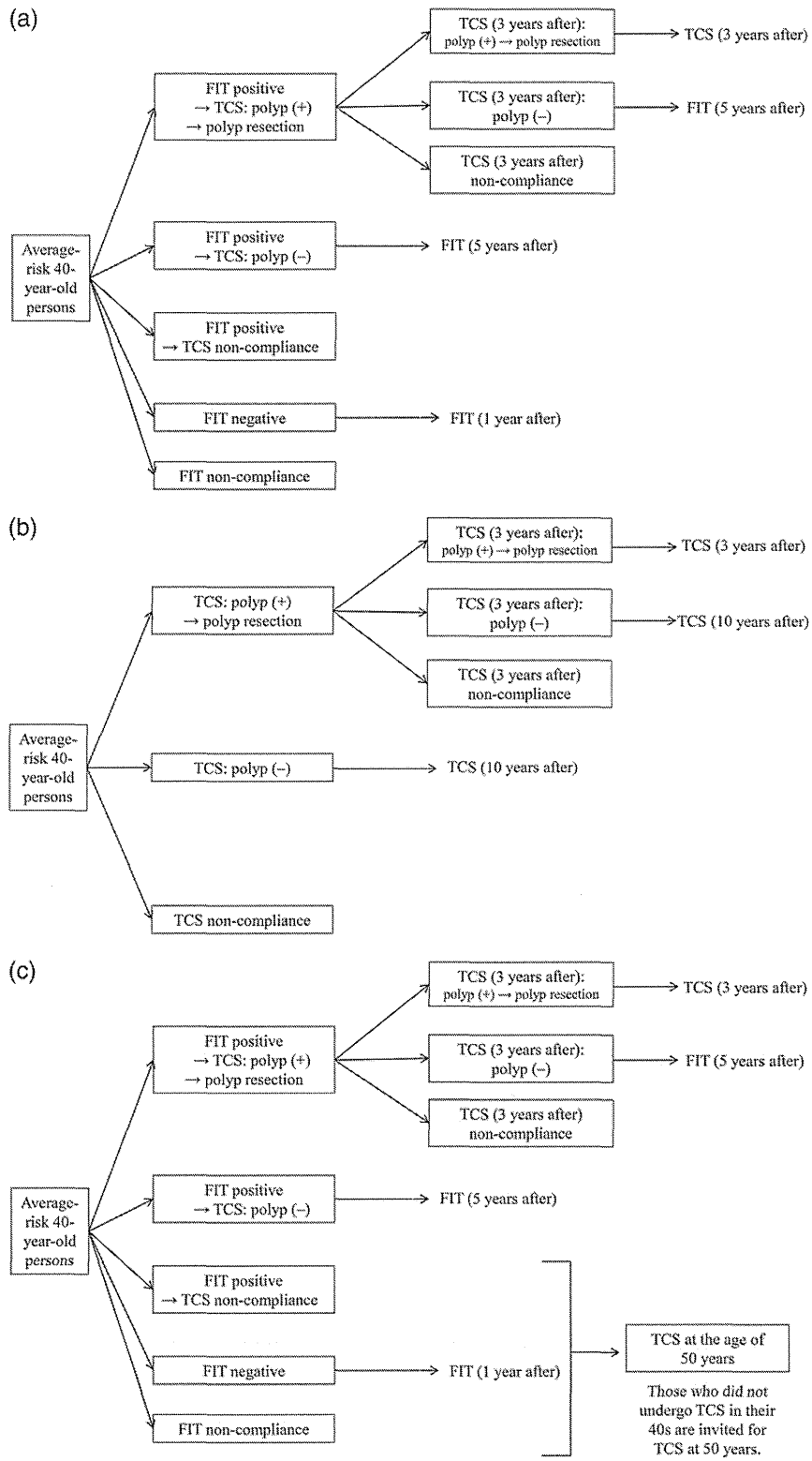

Figure 2. Three screening strategies analyzed in this study. (a) Strategy 1: A fecal immunochemical test-based screening strategy. FIT, fecal immunochemical test; TCS, total colonoscopy. (b) Strategy 2: A total colonoscopy-based screening strategy. (c) Strategy 3: A strategy of adding population-wide total colonoscopy for 50-year-old individuals to a fecal immunochemical test-based screening. During the first 10 years (40–49 years), individuals follow Strategy 1. All of those who did not undergo total colonoscopy during the first 10 years undergo total colonoscopy at the age of 50 years.

Table 1. Model parameters in the cost-effectiveness analysis

Model parameters	Baseline value	References
Transition probabilities (per year)		
Probability of progression to CRC		
From normal epithelium to 1–4 mm sized low-risk polyp	3.4–6.6% (different by age)	33
From 1–4 mm low-risk polyp to 5–9 mm low-risk polyp	1.4–5.6% (different by age)	33
From 5–9 mm low-risk polyp to high-risk polyp	1.3–5.6% (different by age)	33
From high-risk polyp to Dukes' A CRC	3.4%	20, 39
From Dukes' A CRC to Dukes' B CRC	58.3%	20, 39
From Dukes' B CRC to Dukes' C CRC	65.6%	20, 39
From Dukes' C CRC to Dukes' D CRC	86.5%	20, 39
Probability of death from CRC		
Dukes' A	1.7%	38
Dukes' B	3.2%	38
Dukes' C	7.2%	38
Dukes' D	28.4%	38
Probability of symptomatic presentation of CRC		
Dukes' A	6.5%	20, 39
Dukes' B	26.0%	20, 39
Dukes' C	46.0%	20, 39
Dukes' D	92.0%	20, 39
Probability of developing polyps following endoscopic polyp resection		
Developing low-risk polyp (1–4 mm) after endoscopic polyp resection	10.0%	34
Developing low-risk polyp (5–9 mm) after endoscopic polyp resection	5.3%	34
Developing high-risk polyp after endoscopic polyp resection	0.7%	34
Probability of recurrence after treatment of colorectal cancer		
Dukes' A	0.8%	38
Dukes' B	2.8%	38
Dukes' C	7.1%	38
Test characteristics		
FIT		
Sensitivity for 1–4 mm low-risk polyp	6.3%	35, 36
Sensitivity for 5–9 mm low-risk polyp	7.9%	35, 36
Sensitivity for high-risk polyp	26.5%	35, 36
Sensitivity for Dukes' A CRC	52.8%	35, 36
Sensitivity for Dukes' B CRC	70.0%	35, 36
Sensitivity for Dukes' C and D CRC	78.3%	35, 36
Specificity for colorectal polyp and CRC	94.6%	35, 36
TCS		
Sensitivity for 1–4 mm low-risk polyp	74.1%	34
Sensitivity for 5–9 mm low-risk polyp	86.5%	34
Sensitivity for high-risk polyp	97.6%	34
Sensitivity for CRC (Dukes' A–D)	99.9%	34
Specificity for colorectal polyp and CRC	100.0%	34
Probability of perforation after TCS without endoscopic polyp resection	0.01%	37
Probability of perforation after TCS with endoscopic polyp resection	0.06%	37
Probability of death following perforation	6.7%	37
Probability of bleeding after TCS with endoscopic polyp resection	0.5%	37
Cost (JPY)		Japanese national reimbursement tables
FIT	1600	
TCS	15 500	
Endoscopic resection of low-risk polyp	50 000	
Endoscopic resection of high-risk polyp	157 114	
Annual cost of CRC management by Dukes classification		
Dukes' A (1 year)	1 319 816	
Dukes' A (2–5 years)	35 570	
Dukes' B (1 year)	1 399 034	
Dukes' B (2–5 years)	35 570	
Dukes' C (1 year)	2 340 416	
Dukes' C (2–5 years)	44 972	
Dukes' D (1 year)	2 687 125	
Dukes' D (2–5 years)	2 544 972	

CRC, colorectal cancer; FIT, fecal immunochemical test; TCS, total colonoscopy.

screening, there would be 9541 CRC cases among the cohort of 100 000 individuals, and the calculated QALYs and total cost per person were 22.8 and JPY 156 125, respectively. Compared with no screening, all three screening strategies (Strategies 1, 2 and 3) experienced fewer CRC cases, gained more QALYs, and were less costly; i.e. all three strategies dominated no screening.

Among the three strategies, simple dominance of Strategy 3 over Strategy 1 was observed: Strategy 3 resulted in more QALYs and less cost than Strategy 1. Compared with Strategies 1 and 3, Strategy 2 yielded more QALYs, but involved greater cost. The ICERs per QALY gained for Strategy 2 against Strategies 1 and 3 were JPY 293 616 and JPY 781 342, respectively.

With regard to the number of TCS procedures, Strategy 2 required the most procedures (294 322 procedures per 100 000 population), followed by Strategy 3 (126 171 procedures per 100 000), and Strategy 1 (100 740 procedures per 100 000).

Scenario analyses

When the uptake rates decreased to 10%, Strategy 2 showed simple dominance over no screening and the other two screening strategies, whereas the ICER per QALY gained for no screening against Strategy 3 was JPY 218 464 (Table 3). When the uptake rates increased to 100%, all three screening strategies showed simple dominance over no screening, and the ICERs per QALY gained for Strategy 2 against Strategies 1 and 3 were JPY 126 810 and JPY 19 475, respectively (Table 3).

When the initial age of screening changed to 50 years, all three screening strategies dominated no screening, and the ICERs were JPY 87 804 and JPY 125 953 per QALY gained for Strategy 2 against Strategies 1 and 3 (Table 3).

The results for QALYs, costs and required number of TCS procedures when the age for population-wide TCS in Strategy 3 was changed between 40 and 60 years are shown in Table 4. Compared with the base case scenario of Strategy 3 with TCS at 50 years, the strategy with population-wide TCS at the age of 40 years resulted in fewer QALYs and higher cost. In contrast, when the population-wide TCS was performed at 55 years, more QALYs were gained with lower cost than when the TCS was performed at 50 years. The ICER per QALY gained for the strategy with TCS at 55 years against the strategy with TCS at 60 years was JPY 206 113. Against the strategy with TCS at 55 years, the ICER per QALY gained for the strategy with TCS at 45 years was JPY 782 013. The strategy with TCS at 45 years yielded more QALYs and was less costly than Strategy 2, and the ICER per QALY gained for this strategy against Strategy 1 was JPY 151 856. The required number of TCS procedures decreased as the age for population-wide TCS increased.

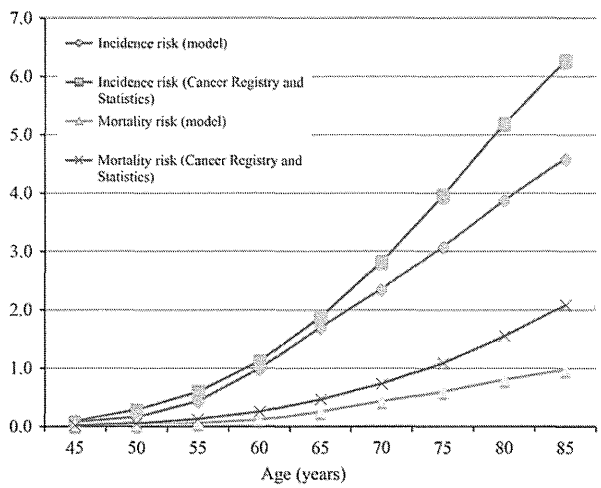

Figure 3. Comparison between cumulative risks for colorectal cancer incidence and mortality estimated from the study model and those estimated from the Cancer Registry and Statistics.

Probabilistic sensitivity analysis

The probabilistic sensitivity analysis performed for no screening and the three strategies (Strategies 1, 2 and 3) and the cost-effectiveness acceptability curve showed a correlation between the probability of being chosen as the most cost-effective scenario for each strategy and the WTP values (Fig. 4). In the figure, the horizontal axis represents the WTP value for one additional QALY, with a range of JPY 0–10 000 000, and the vertical axis represents the probability of being chosen as the most cost-effective scenario for each strategy. When the WTP value was set at JPY 5 000 000, the probability of being chosen as the most cost-effective scenario was 2.2% for no screening, 21.0% for Strategy 1, 48.7% for Strategy 2 and 28.1% for Strategy 3. When the age for population-wide TCS was changed to 45 years in Strategy 3, the probability resulted in 2.4% for no screening, 21.8% for Strategy 1, 53.2% for Strategy 2, and 22.6% for Strategy 3.

Table 2. Results of the base case analysis

	No screening	Strategy 1	Strategy 2	Strategy 3
Cost (per person, JPY)	156 125	94 733	99 930	93 523
QALYs (per person)	22.7986	23.0001	23.0178	23.0096
CRC cases (per 100 000 persons)	9541	3926	2989	3625
TCS procedures (per 100 000 persons)	—	100 740	294 322	126 171
Incremental cost per QALY gained (JPY)				
vs. No screening	—	Dominates[a]	Dominates	Dominates[a]
vs. Strategy 1	Dominated[b]	—	293 616	Dominates[a]
vs. Strategy 2	Dominated[b]	see Strategy 2 vs. 1	—	see Strategy 2 vs. 3
vs. Strategy 3	Dominated[b]	Dominated[b]	781 342	—

[a]'Dominates' denotes a strategy (column) that is less costly and more effective than its comparator (row).
[b]'Dominated' denotes a strategy (column) that is more costly and less effective than its comparator (row).
QALY, quality-adjusted life-years.

Table 3. Results of the scenario analyses on the uptake rates and initial age of screening

	No screening	Strategy 1	Strategy 2	Strategy 3
Uptake rates: 100%				
Cost (per person, JPY)	154 694	99 382	104 961	103 789
QALYs (per person)	22.8026	23.0770	23.1210	23.0608
Incremental cost per QALY gained (JPY)				
vs. No screening	—	Dominates[a]	Dominates[a]	Dominates[a]
vs. Strategy 1	Dominated[b]	—	126 810	Dominated[b]
vs. Strategy 2	Dominated[b]	see Strategy 2 vs. 1	—	see Strategy 2 vs. 3
vs. Strategy 3	Dominated[b]	Dominates[a]	19 475	—
Uptake rates: 10%				
Cost (per person, JPY)	153 653	152 928	137 289	151 710
QALYs (per person)	22.8209	22.8278	22.8753	22.8120
Incremental cost per QALY gained (JPY)				
vs. No screening	—	Dominates[a]	Dominates[a]	See No screening vs. 3
vs. Strategy 1	Dominated[b]	—	Dominates[a]	see Strategy 1 vs. 3
vs. Strategy 2	Dominated[b]	Dominated[b]	—	Dominated[b]
vs. Strategy 3	218 464	77 010	Dominates[a]	—
Starting age: 50 years				
Cost (per person, JPY)	154 107	99 793	104 069	99 043
QALYs (per person)	22.8194	23.0845	23.1332	23.0933
Incremental cost per QALY gained (JPY)				
vs. No screening	—	Dominates[a]	Dominates[a]	Dominates[a]
vs. Strategy 1	Dominated[b]	—	87 804	Dominates[a]
vs. Strategy 2	Dominated[b]	see Strategy 2 vs. 1	—	see Strategy 2 vs. 3
vs. Strategy 3	Dominated[b]	Dominated[b]	125 953	—

[a]'Dominates' denotes a strategy (column) that is less costly and more effective than its comparator (row).
[b]'Dominated' denotes a strategy (column) that is more costly and less effective than its comparator (row).

Table 4. Results of the scenario analyses on the age for population-wide total colonoscopy in Strategy 3

	40 years	45 years	50 years	55 years	60 years
Cost (per person, JPY)	99 602	97 679	93 523	92 049	91 142
QALYs (per person)	22.9979	23.0195	23.0096	23.0123	23.0079
TCS procedures (per 100 000 persons)	138 687	133 193	126 171	123 659	123 106
Incremental cost per QALY gained (JPY)					
vs. 40 years	—	Dominates[a]	Dominates[a]	Dominates[a]	Dominates[a]
vs. 45 years	Dominated[b]	—	see 45 years vs. 50 years	see 45 years vs. 55 years	see 45 years vs. 60 years
vs. 50 years	Dominated[b]	420 284	—	Dominates[a]	see 50 years vs. 60 years
vs. 55 years	Dominated[b]	782 013	Dominated[b]	—	see 55 years vs. 60 years
vs. 60 years	Dominated[b]	564 055	1.400 462	206 113	—

[a]'Dominates' denotes a strategy (column) that is less costly and more effective than its comparator (row).
[b]'Dominated' denotes a strategy (column) that is more costly and less effective than its comparator (row).

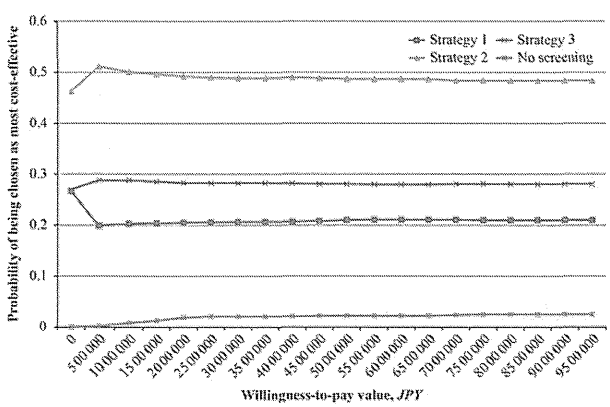

Figure 4. Probabilistic sensitivity analysis performed for the three strategies (1, 2 and 3) and no screening.

Discussion

This study examined in detail the cost-effectiveness of CRC screening with FIT and/or TCS in the Japanese settings by performing a simulation model analysis. For this analysis, we constructed a model of CRC using Japanese clinical data. The validity of the model was indicated by the finding that the cumulative risks for CRC incidence and mortality estimated from the model and the Cancer Registry and Statistics matched mostly, particularly for people ≤65 years of age. Although these risk estimates differed slightly after the age of 65 years, we believe that it does not matter in this study. On the contrary, the difference strengthens the evidence for the favorable cost-effectiveness of CRC screening indicated by the model analysis because the lower CRC incidence and mortality estimated from the model means that it may be more difficult to prove the (cost-)effectiveness of screening using the model than with real-life data.

Our results indicate that CRC screening with FIT and/or TCS was superior to no screening from the perspective of cost-effectiveness in most cases. This finding agrees with previous foreign cost-effectiveness studies on CRC screening (17–24). However, when the uptake rates decreased to 10%, the ICER per QALY gained for no screening against Strategy 3 was well below JPY 5-6 million. Considering that this amount is the upper limit of the WTP value for one additional QALY in Japan (43), it is postulated that the superiority of CRC screening to no screening in terms of cost-effectiveness will be more difficult to maintain when uptakes rates are low. To maintain the superior cost-effectiveness of CRC screening, it will be essential to achieve high screening uptake rates.

Despite a number of previous cost-effectiveness studies on CRC screening, there has been no consensus on the optimal use in terms of cost-effectiveness of FIT and TCS for population-based CRC screening (17–26). In the base case analysis of this study, the ICER per QALY gained for Strategy 2 against Strategy 1 was lower than the upper limit of the WTP value in Japan and Strategy 3 showed simple dominance over Strategy 1, which suggests that the strategies that use TCS more actively (Strategies 2 and 3) could be more cost-effective than the FIT-based screening strategy (Strategy 1). Furthermore, the sensitivity analyses showed that the strategies with greater use of TCS (Strategies 2 and 3) could be more cost-effective than the FIT-based screening strategy (Strategy 1) in most cases. This finding may largely be due to the much lower fee per TCS procedure in Japan than in other countries. Comparing cost-effectiveness between Strategies 2 and 3, the base case and sensitivity analyses showed that Strategy 2 was more cost-effective than Strategy 3 in many cases. However, the sensitivity analyses showed that the superiority of Strategy 2 against Strategy 3 with regard to cost-effectiveness was not always the case and that Strategy 3 could be more cost-effective than Strategy 2 under certain sets of model parameters and the age for population-wide TCS in Strategy 3.

If TCS is to be used more actively for population-based CRC screening, its safety and the availability of TCS resources require discussion. First, with regard to the safety of TCS, recent foreign studies have reported that the perforation rate of TCS without polypectomy was 0.01–0.03%, which is a very low rate that indicates the safety of screening TCS (44–48). Similarly, in Japan, the corresponding rate has been reported to be low, as shown in Table 1 (37). Given the safety of screening TCS, it may be possible to use it more actively than the currently performed FIT-based CRC screening. However, the risk of perforation associated with TCS cannot be completely ignored at present, particularly for the elderly population (44–48). Second, the capacity for screening TCS in Japan has not been clarified, with some surveys currently in progress, including the Japan endoscopy database project (UMIN000016093). Nevertheless, it is obvious that TCS capacity is limited in Japan and that we must arrange the CRC screening system to meet this limitation. Considering the limited TCS capacity, the TCS-based screening (Strategy 2), which requires more than double the number of TCS procedures than the other strategies in this study (Strategies 1 and 3), is likely to be the most difficult to implement.

From the cost-effectiveness aspect only, the TCS-based strategy may be the best; however, considering cost-effectiveness, safety, and the TCS capacity issue together, we postulate that the strategy of adding population-wide TCS at a specific age to the FIT-based strategy (Strategy 3) may be an optimal option for population-based CRC screening in Japan. With regard to the optimal age for population-wide TCS in Strategy 3, TCS at 45 years was the most cost-effective under the condition of the upper limit of WTP being JPY 5-6 million, according to the scenario analyses in this study. Considering that it is necessary to set the age for population-wide TCS as a range rather than one specific age to achieve a higher uptake rate, it appears that TCS within the age range 45–55 years would be acceptable from the perspective of cost-effectiveness on the basis of the study results. This would also be expected to improve the safety of the procedure because of the relatively younger age. With regard to the TCS capacity, although more TCS procedures may be required than with the FIT-based strategy, the increase is considered not to be too great; the number of TCS procedures required in Strategy 3 (TCS at 45–55 years) compared with those required in Strategy 1 was 123 659–133 193 vs. 100 740 per 100 000 individuals, whereas Strategy 2 required 294 322 TCS procedures per 100 000 individuals.

This study had several limitations. First, the natural history model of CRC in this study was based on currently available Japanese data; as a result, it was completely based on the concept of the adenoma-carcinoma sequence on which the previously reported cost-effectiveness analyses were based (17–24). However, other CRC pathways, such as the serrated pathway and the *de novo* pathway, have been reported, and it may be necessary to include these in the natural history model of CRC in future analyses, after the collection of a sufficiently large body of data on serrated polyps or *de novo* cancers (49,50). Second, the values of model parameters set in the base case analysis could vary case by case in the real world. However, sensitivity analyses, including probabilistic sensitivity analyses, were performed for the parameters. Third, indirect costs such as productivity loss cost due to CRC treatment were not considered in this study. Because limited data are available on indirect costs in Japan at present, it is currently difficult to include these costs in the cost-effective analysis. However, the cost-effective analyses in this study were performed form the healthcare payer's perspective in Japan, and thus we believe that no inclusion of indirect cost was appropriate for the analyses. For future cost-effectiveness analyses that include other perspectives, inclusion of data on the indirect costs associated with CRC in Japan would be warranted.

In conclusion, the present study examined the cost-effectiveness of population-based CRC screening in Japan. The CRC screening strategies with more active use of TCS could be more cost-effective than the FIT-based screening strategy. The TCS-based screening strategy could be the most cost-effective; however, considering the safety and limited capacity of TCS resources in addition to cost-effectiveness, the strategy of adding population-wide TCS for individuals in the age range 45–55 years to the FIT-based screening may be an optimal solution.

Funding

This work was supported by JSPS KAKENHI, grant number 25871160 and the National Cancer Center Research and Development Fund (26-A-31).

Conflict of interest statement

None declared.

References

1. Matsuda A, Matsuda T, Shibata A, Katanoda K, Sobue T, Nishimoto H; Japan Cancer Surveillance Research Group<. Cancer incidence and

incidence rates in Japan in 2008: a study of 25 population-based cancer registries for the Monitoring of Cancer Incidence in Japan (MCIJ) project. *Jpn J Clin Oncol* 2014;44:388–96.
2. Saito H. Screening for colorectal cancer: current status in Japan. *Dis Colon Rectum* 2000;43:S78–84.
3. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Engl J Med* 1993;328:1365–71.
4. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 1996;348:1472–7.
5. Kronborg O, Fenger C, Olsen J, Jørgensen OD, Søndergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996;348:1467–71.
6. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med* 2000; 343:1603–7.
7. Faivre J, Dancourt V, Lejeune C, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. *Gastroenterology* 2004;126:1674–80.
8. Hiwatashi N, Morimoto T, Fukao A, et al. An evaluation of mass screening using fecal occult blood test for colorectal cancer in Japan: a case-control study. *Jpn J Cancer Res* 1993;84:1110–2.
9. Saito H, Soma Y, Koeda J, et al. Reduction in risk of mortality from colorectal cancer by fecal occult blood screening with immunochemical hemagglutination test. A case-control study. *Int J Cancer* 1995;61: 465–9.
10. Saito H. Screening for colorectal cancer by immunochemical fecal occult blood testing. *Jpn J Cancer Res* 1996;87:1011–24.
11. Zappa M, Castiglione G, Grazzini G, et al. Effect of faecal occult blood testing on colorectal mortality: results of a population-based case-control study in the district of Florence, Italy. *Int J Cancer* 1997; 73:208–10.
12. Saito H, Soma Y, Nakajima M, et al. A case-control study evaluating occult blood screening for colorectal cancer with hemoccult test and an immunochemical hemagglutination test. *Oncol Rep* 2000;7:815–9.
13. Nakajima M, Saito H, Soma Y, Sobue T, Tanaka M, Munakata A. Prevention of advanced colorectal cancer by screening using the immunochemical faecal occult blood test: a case-control study. *Br J Cancer* 2003;89:23–8.
14. Lee KJ, Inoue M, Otani T, Iwasaki M, Sasazuki S, Tsugane S; Japan Public Health Center-based Prospective Study<. Colorectal cancer screening using fecal occult blood test and subsequent risk of colorectal cancer: a prospective cohort study in Japan. *Cancer Detect Prev* 2007;31:3–11.
15. Nishihara R, Wu K, Lochhead P, et al. Long-term colorectal-cancer incidence and mortality after lower endoscopy. *N Engl J Med* 2013;369: 1095–105.
16. Zauber AG, Winawer SJ, O'Brien MJ, et al. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N Engl J Med* 2012;366: 687–96.
17. Frazier AL, Colditz GA, Fuchs CS, Kuntz KM. Cost-effectiveness of screening for colorectal cancer in the general population. *JAMA* 2000;284: 1954–61.
18. Sonnenberg A, Delcò F, Inadomi JM. Cost-effectiveness of colonoscopy in screening for colorectal cancer. *Ann Intern Med* 2000;133:573–84.
19. Pignone M, Saha S, Hoerger T, Mandelblatt J. Cost-effectiveness analyses of colorectal cancer screening: a systematic review for the U.S. Preventive Services Task Force. *Ann Intern Med* 2002;137:96–104.
20. Tappenden P, Chilcott J, Eggington S, Patnick J, Sakai H, Karnon J. Option appraisal of population-based colorectal cancer screening programmes in England. *Gut* 2007;56:677–84.
21. Vijan S, Hwang I, Inadomi J, et al. The cost-effectiveness of CT colonography in screening for colorectal neoplasia. *Am J Gastroenterol* 2007;102: 380–90.
22. Tsoi KK, Ng SS, Leung MC, Sung JJ. Cost-effectiveness analysis on screening for colorectal neoplasm and management of colorectal cancer in Asia. *Aliment Pharmacol Ther* 2008;28:353–63.
23. Lansdorp-Vogelaar I, Knudsen AB, Brenner H. Cost-effectiveness of colorectal cancer screening. *Epidemiol Rev* 2011;33:88–100.
24. Sharaf RN, Ladabaum U. Comparative effectiveness and cost-effectiveness of screening colonoscopy vs. sigmoidoscopy and alternative strategies. *Am J Gastroenterol* 2013;108:120–32.
25. Tsuji I, Fukao A, Shoji T, Kuwajima I, Sugawara N, Hisamichi S. Cost-effectiveness analysis of screening for colorectal cancer in Japan. *Tohoku J Exp Med* 1991;164:269–78.
26. Shimbo T, Glick HA, Eisenberg JM. Cost-effectiveness analysis of strategies for colorectal cancer screening in Japan. *Int J Technol Assess Health Care* 1994;10:359–75.
27. Sekiguchi M, Matsuda T, Tamai N, et al. Cost-effectiveness of total colonoscopy in screening of colorectal cancer in Japan. *Gastroenterol Res Pract* 2012;2012:728454.
28. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;13:322–38.
29. Hamiliton SR, Bosman FT, Boffetta P, et al. Carcinoma of the colon and rectum. In: Bosman FT, Carneiro F, Hruban RH, Theise ND, editors. *WHO Classification of Tumors of the Digestive System*. 4th ed. Lyon: IARC, 2010;134–46.
30. Winawer SJ, Zauber AG. The advanced adenoma as the primary target of screening. *Gastrointest Endosc Clin N Am* 2002;12:1–9. v.
31. Ministry of Health, Labor and Welfare, Japan. Comprehensive Survey of Living Conditions 2013 [in Japanese]. Health, Labor and Welfare Statistics Association, Tokyo, Japan, 2013.
32. Kitagawa S, Miyagawa K, Iriguchi Y, et al. Nationwide survey on gastrointestinal cancer screening 2012. *Journal of Gastrointestinal Cancer Screening* 2015;53:60–101.
33. Hamashima C, Sobue T, Muramatsu Y, Saito H, Moriyama N, Kakizoe T. Comparison of observed and expected numbers of detected cancers in the research center for cancer prevention and screening program. *Jpn J Clin Oncol* 2006;36:301–8.
34. Sano Y, Fujii T, Matsuda T, et al. Study design and patient recruitment for the Japan Polyp Study. *Open Access Journal of Clinical Trials* 2014;6: 37–44.
35. Morikawa T, Kato J, Yamaji Y, Wada R, Mitsushima T, Shiratori Y. A comparison of the immunochemical fecal occult blood test and total colonoscopy in the asymptomatic population. *Gastroenterology* 2005; 129:422–8.
36. Morikawa T, Kato J, Yamaji Y, et al. Sensitivity of immunochemical fecal occult blood test to small colorectal adenomas. *Am J Gastroenterol* 2007;102:2259–64.
37. Yoshino J, Igarashi Y, Ohara H, et al. 5th report of endoscopic complications: results of the Japan Gastroenterological Endoscopy Society. *Gastroenterol Endosc* 2010;52:95–103. (in Japanese).
38. Japanese Society for Cancer of the Colon and Rectum. *JSCCR Guidelines 2014 for the Treatment of Colorectal Cancer*. Tokyo, Japan: Kanehara & Co., Ltd., 2014. (in Japanese).
39. Sweet A, Lee D, Gairy K, Phiri D, Reason T, Lock K. The impact of CT colonography for colorectal cancer screening on the UK NHS: costs, healthcare resources and health outcomes. *Appl Health Econ Health Policy* 2011;9:51–64.
40. Segnan N, Patnick J, von Karsa L, editors. *European guidelines for quality assurance in colorectal cancer screening and diagnosis*, 1st edn. Luxemburg: Publication Office of the EU, 2010.
41. 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;59:62–8.
42. Siegel JE, Torrance GW, Russell LB, Luce BR, Weinstein MC, Gold MR. Guidelines for pharmacoeconomic studies. Recommendations from the panel on cost effectiveness in health and medicine. Panel on cost Effectiveness in Health and Medicine. *Pharmacoeconomics* 1997; 11:159–68.
43. Shiroiwa T, Sung YK, Fukuda T, Lang HC, Bae SC, Tsutani K. International survey on willingness-to-pay (WTP) for one additional QALY

gained: what is the threshold of cost effectiveness? *Health Econ* 2010; 19:422–37.
44. Niv Y, Hazazi R, Levi Z, Fraser G. Screening colonoscopy for colorectal cancer in asymptomatic people: a meta-analysis. *Dig Dis Sci* 2008;53:3049–54.
45. Warren JL, Klabunde CN, Mariotto AB, et al. Adverse events after outpatient colonoscopy in the Medicare population. *Ann Intern Med* 2009; 150:849–57.
46. Bokemeyer B, Bock H, Hüppe D, et al. Screening colonoscopy for colorectal cancer prevention: results from a German online registry on 269000 cases. *Eur J Gastroenterol Hepatol* 2009;21:650–5.
47. Rutter CM, Johnson E, Miglioretti DL, Mandelson MT, Inadomi J, Buist DS. Adverse events after screening and follow-up colonoscopy. *Cancer Causes Control* 2012;23:289–96.
48. Hamdani U, Naeem R, Haider F, et al. Risk factors for colonoscopic perforation: a population-based study of 80118 cases. *World J Gastroenterol* 2013;19:3596–601.
49. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. *Gastroenterology* 2010;138:2088–100.
50. Kashida H, Kudo SE. Early colorectal cancer: concept, diagnosis, and management. *Int J Clin Oncol* 2006;11:1–8.

Commentary

Recommendations for a Step-Wise Comparative Approach to the Evaluation of New Screening Tests for Colorectal Cancer

Graeme P. Young, MD, FRACP, FTSE, AGAF[1]; Carlo Senore, MD, MSc[2]; Jack S. Mandel, PhD, MPH[3]; James E. Allison, MD, FACP, AGAF[4]; Wendy S. Atkin, MPH, PhD[5]; Robert Benamouzig, MD, PhD[6]; Patrick M. M. Bossuyt, PhD[7]; Mahinda De Silva, MB, BS, FRACP[8]; Lydia Guittet, MD, PhD[9]; Stephen P. Halloran, MBE, FRCPath[10]; Ulrike Haug, PhD[11]; Geir Hoff, MB, ChB, PhD[12]; Steven H. Itzkowitz, MD, FACP, FACG, AGAF[13]; Marcis Leja, MD, MBA, PhD, AGAF[14]; Bernard Levin, MB, BCh, FACP[15]; Gerrit A. Meijer, MD, PhD[16]; Colm A. O'Morain, MD[17]; Susan Parry, MbCHB, FRACP18[18]; Linda Rabeneck, MD, MPH, FRCPC[19]; Paul Rozen, MD[20†]; Hiroshi Saito, MD, PhD[21]; Robert E. Schoen, MD, MPH[22]; Helen E. Seaman, BSc, PhD[23]; Robert J. C. Steele, MD, FRCS[24]; Joseph J. Y. Sung, MD, PhD[25]; and Sidney J. Winawer, MD[26]

BACKGROUND: New screening tests for colorectal cancer continue to emerge, but the evidence needed to justify their adoption in screening programs remains uncertain. **METHODS:** A review of the literature and a consensus approach by experts was undertaken to provide practical guidance on how to compare new screening tests with proven screening tests. **RESULTS:** Findings and recommendations from the review included the following: Adoption of a new screening test requires evidence of effectiveness relative to a proven comparator test. Clinical accuracy supported by programmatic population evaluation in the screening context on an intention-to-screen basis, including acceptability, is essential. Cancer-specific mortality is not essential as an endpoint provided that the mortality benefit of the comparator has been demonstrated and that the biologic basis of detection is similar. Effectiveness of the guaiac-based fecal occult blood test provides the *minimum* standard to be achieved by a new test. A 4-phase evaluation is recommended. An initial retrospective evaluation in cancer cases and controls (Phase 1) is followed by a prospective evaluation of performance across the continuum of neoplastic lesions (Phase 2). Phase 3 follows the demonstration of adequate accuracy in these 2 prescreening phases and addresses programmatic outcomes at 1 screening round on an intention-to-screen basis. Phase 4 involves more comprehensive evaluation of ongoing screening over multiple rounds. Key information is provided from the following parameters: the test positivity rate in a screening population, the true-positive and false-positive rates, and the number needed to colonoscope to detect a target lesion. **CONCLUSIONS:** New screening tests can be evaluated efficiently by this stepwise comparative approach. ***Cancer* 2016;122:826-39.** © 2016 The Authors. *Cancer* published by Wiley Periodicals, Inc. on behalf of *American Cancer Society*. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

KEYWORDS: colonoscopy, colorectal cancer, fecal occult blood test, molecular diagnostics, screening test.

Corresponding author: Graeme P. Young, MD, FRACP, Flinders Center for Innovation in Cancer, Flinders University, Level 3, Flinders Medical Center, Bedford Park, Adelaide, SA 5042, Australia; Fax: (011) 61-8-8204-3943; graeme.young@flinders.edu.au

[1]Flinders Center for Innovation in Cancer, Flinders University, Adelaide, South Australia, Australia; [2]Reference Center for Epidemiology and Cancer Prevention, Piedmont Regional Center for Preventive Oncology, City Health and Science University Hospital of Turin, Turin, Italy; [3]Environmental and Occupational Medicine, University of Minnesota, Minneapolis, Minnesota; [4]Division of Gastroenterology, University of California, San Francisco and Kaiser Division of Research, Oakland, California; [5]Gastrointestinal Epidemiology, Imperial College, London, United Kingdom; [6]Gastroenterology Department, Avicenne Hospital, Paris 13 University, Paris, France; [7]Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands; [8]Department of Gastroenterology, Repatriation General Hospital, Adelaide, South Australia, Australia; [9]Unit 1086, French National Institute for Health and Medical Research, Cancers and Preventions Center, Caen University Hospital, Caen, France; [10]Faculty of Health and Medical Sciences, University of Surrey, Guildford, United Kingdom; [11]Department of Clinical Epidemiology, Leibniz Institute for Prevention Research and Epidemiology, Bremen, Germany; [12]Telemark Hospital, Skein Cancer Registry of Norway, University of Oslo, Oslo, Norway; [13]Gastrointestinal Division, Icahn School of Medicine at Mount Sinai, New York, New York; [14]Digestive Diseases Center, GASTRO, Faculty of Medicine, University of Latvia, Riga, Latvia; [15]Division of Cancer Prevention, The University of Texas MD Anderson Cancer Center, Houston, Texas; [16]Netherlands Cancer Institute, Amsterdam, The Netherlands; [17]Faculty of Health Science, Trinity College Dublin, Dublin, Ireland; [18]Ministry of Health Bowel Cancer Program, Auckland Hospital, Auckland, New Zealand; [19]Prevention and Cancer Control, Cancer Care Ontario, and University of Toronto, Toronto, Ontario, Canada; [20]Department of Gastroenterology, Sestopali Fund for Gastrointestinal Cancer Prevention, Tel Aviv, Israel; [21]Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan; [22]Department of Medicine and Epidemiology, University of Pittsburgh, Pittsburgh, Pennsylvania; [23]National Health Service Bowel Cancer Screening Southern Program Hub, Royal Surrey County Hospital, Guildford, United Kingdom; [24]Department of Surgery, University of Dundee, Dundee, Scotland; [25]Office of the Vice Chancellor, The Chinese University of Hong Kong, Shatin, China; [26]Gastroenterology and Nutrition Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, New York, New York

†Deceased.

We are grateful to the secretariat of the World Endoscopy Organization and to Alexandra Whibley, Erin Symonds, and Emma-May Palmer for their assistance in the consensus processes and article preparation.

This article is dedicated to the memory of Professor Paul Rozen, a true pioneer in colorectal cancer screening, who passed away in early 2013.

Members of the original working party responsible for the consensus process underlying these recommendations were convened by the World Endoscopy Organization Colorectal Cancer Screening Committee and were: G.P. Young, J. Mandel, J.J.Y. Sung, J.E. Allison, W. Atkin, R. Benamouzig, G. Hoff, S.H. Itzkowitz, T.R. Levin, E.G. McFarlane, C. O'Morain, S. Parry, L. Rabeneck, P. Rozen, H. Saito, R.E. Schoen, C. Senore, R.J.C. Steele, S.J. Winawer, and B.C.Y. Wong. The listed authors were members of the World Endoscopy Organization/World Gastroenterology Organization New Screening Tests Expert Working Party and met the criteria for authorship of this article.

Additional supporting information may be found in the online version of this article.

DOI: 10.1002/cncr.29865, **Received:** September 2, 2015; **Revised:** November 14, 2015; **Accepted:** November 30, 2015, **Published online** February 1, 2016 in Wiley Online Library (wileyonlinelibrary.com)

INTRODUCTION

New tests to screen for colorectal cancer (CRC) continue to emerge and are based on new biomarkers, new imaging modalities, or variations of existing methods. Efficient evaluation of these options presents a challenge. It has been observed that new *diagnostic* tests frequently enter practice without evidence of improved outcomes.[1] For *screening* tests, the requirement for evidence is more demanding, because more than clinical test accuracy (ie, sensitivity and specificity) is required to justify adoption.[1,2] Safety, public acceptability, and cost effectiveness need to be assessed even more carefully for tests that are to be applied to ostensibly healthy individuals.

The intention of a cancer screening program, or secondary prevention, is to significantly reduce the cancer site-specific mortality and burden of that disease in the target population[2] through programmatic use of a test that detects neoplasia at a stage early enough for treatment to be successful and/or for incidence to be reduced.[3]

It has been demonstrated that certain screening tests reduce cancer site-specific mortality and/or incidence by randomized, population-based evaluation on an intention-to-screen basis,[4-12] thereby limiting biases, such as lead-time, length, and self-selection, that are often present in simpler studies that use surrogate measures of mortality or intermediate endpoints. Evaluation of every new CRC screening test to the endpoint of mortality would be a huge and expensive undertaking and would markedly slow—if not prohibit—the implementation of promising new technologies. Fortunately, simpler studies using surrogate measures or intermediate endpoints can be used to evaluate new tests[1] provided that a carefully validated reference standard is used and biases are minimized. To define what is justifiably required to support the use of a new test for CRC screening, we propose an efficient and rigorous method for how to compare the alternative/new (hereafter "new") with the proven/established screening tests.

METHODS

To establish the guiding principles for comparative evaluation, including the informative endpoints and the appropriate study design, we established a consensus based on the Glaser and Delphi approaches.[13] The *membership* was chosen from experts because of their knowledge or experience in practice or research relevant to screening for CRC. The *problem* was defined by using the consensus process to agree on the goal. To support the consensus process, systematic literature searches were undertaken using Medline and other relevant databases. One search string was optimized for diagnosis and screening with the inclusion of measures like sensitivity, another was optimized for cancer, and a third attempted to identify articles focused

BOX 1: These are the guiding principles that underpin a strategy for comparing screening tests that emerged from the consensus approach and the literature review (a discussion of each is provided in Supporting Table 1; see online supporting information):

Principle 1. Screening aims to reduce the burden of disease in the targeted population, without adversely affecting the health status of those who participate in screening, through early detection and treatment of cancer and/or through detection of precancer lesions, which reduces incidence.

Principle 2. The screening test is just 1 event in a process that includes engagement of the public, testing, validation, communication, and treatment.

Principle 3. Population randomized controlled trials with mortality as the primary outcome set the standard for the evaluation of new tests.

Principle 4. New tests can be assessed in parallel with an existing test all the way through the screening process, from population engagement to population outcomes/measures.

Principle 5. New screening tests might detect a different neoplasia-dependent biology; as a consequence, the value of treatment and benefit to mortality reduction might not be the same.

Principle 6. In 2-step screening, the screening test selects participants who proceed to diagnostic verification by colonoscopy, because a positive test increases the likelihood of neoplasia being present.

Principle 7. It is not ethically justifiable to proceed to study a test in the screening environment, including acceptability to invitees or other screening program outcomes, without studies indicating that the new test is of acceptable accuracy compared with a proven comparator test.

Principle 8. New tests must be clearly defined with provision of adequate technical details, quality-assurance procedures, and performance standards.

TABLE 1. Characteristics of Established Screening Tests Known to Reduce Colorectal Cancer Mortality and the Type of Evidence Supporting Their Value

Detection Goal	Technology	Strongest Evidence for Benefit	Test Objective	Sensitivity Determinant	Specificity Determinants
Fecal blood	Guaiac-based FOBT (gFOBT)	Population RCTs—reduced incidence and mortality	Heme component of hemoglobin	Amount of fecal heme exceeds that needed to generate a positive result (fixed by manufacturer)	Dietary peroxidases; agents interfering with peroxidase reaction; bleeding nonneoplastic lesions; amount of stool in sample.
	Fecal immunochemical test for hemoglobin (FIT)	Case-control and cohort studies—reduced incidence and mortality; comparative screening cohorts (randomized)—higher detection rates and participation compared with gFOBT	Globin component of human hemoglobin	Amount of fecal hemoglobin exceeding selected cutoff concentration (may be fixed by manufacturer or selected by end user)	Bleeding nonneoplastic colonic lesions; amount of stool in sample.
Endoscopic visualization of lesion	Colonoscopy	Case-control and cohort studies—reduced incidence and mortality	Visually apparent lesions (ulcerative, polypoid, or flat/depressed) suspicious of neoplasia	Quality of procedure; ability to negotiate the colonic lumen with adequate views; nature of the lesion	Histopathologic clarification
	Sigmoidoscopy (flexible)	Population RCTs – reduced incidence and mortality	Visually apparent lesions within reach	Quality of procedure; depth of insertion; ability to negotiate the colonic lumen with adequate views; nature of the lesion	Histopathologic clarification

Abbreviations: FOBT, fecal occult blood test; RCTs, randomized controlled trials.
[a] This information is derived from several publications.[5,6,14-19]

on comparison of tests. We also searched for review articles that addressed the evidence supporting screening for CRC.

A series of *specific questions* that focused on the definition of appropriate study designs and outcomes for the comparison of different screening tests were established by agreement. The *answers* to these questions were reached by consensus (requiring 75% agreement) based on dissemination of summaries of the literature searches, detailed examination of methodological articles, a series of semistructured discussions with dissemination of decisions after each critique, followed by consultation with external advisors. On the basis of these processes, progressive drafts of the recommendations were then prepared, circulated, and critiqued.

In this report, we present: 1) the underlying guiding principles that emerged from the consensus; 2) an expert opinion on the methods appropriate for evaluating a new test compared with a proven comparator test (what is needed), 3) practical guidance on how to apply these methods in a 4-step, phased evaluation (how to do it); and 4) examples of published research that exemplify these phases (how it has been done). Therefore, it will guide researchers and enable practitioners to decide whether a new test is suitable for the context in which they practice.

GUIDING PRINCIPLES

The guiding principles that emerged from the consensus approach and the literature review are outlined in Box 1, together with their key consequences for test comparison. A presentation of the reasoning underlying these principles is presented in Supporting Table 1 (see online supporting information).

With regard to Principle 3, which states that "Population randomized controlled trials (RCTs) set the standard for evaluation of new tests," Table 1 outlines the characteristics of major screening tests known to reduce CRC mortality and/or incidence together with the type of evidence supporting their value. Such tests are ideal as a reference point against which to compare a new test.

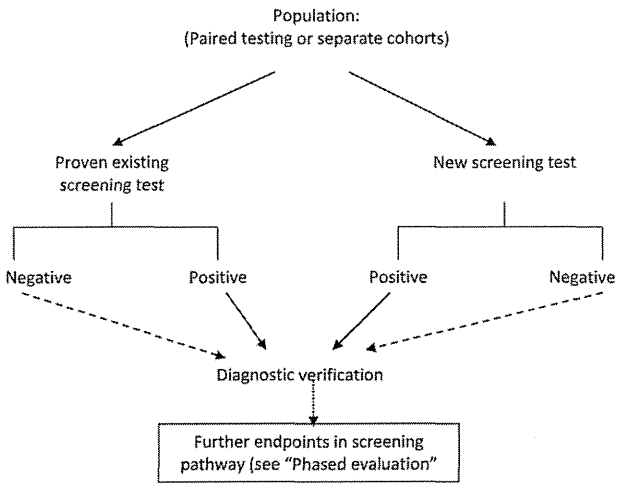

Figure 1. This is a conceptualization of the design for testing a new test relative to an existing (comparator) test. Solid lines represent essential paths in the process, and dashed lines represent discretionary paths that are not essential in some phases of evaluation.

Table 1 also describes the test target (which serves as an informative outcome for comparison), as discussed in Principle 5.

A FRAMEWORK FOR EVALUATING A NEW SCREENING TEST

With these principles in mind, a practical framework for evaluating a "new" test against a proven test can be built. The test of effectiveness for the proven test demands proof at the population level—hence, the context for evaluation must eventually include population outcomes and not just the testing of capacity to detect lesions.

When an RCT establishes that a test is effective in reducing mortality, then a new test does not need to be evaluated with such rigor provided it is compared with the proven test.[1] This is true provided that Principle 5 (Box 1) applies; namely, that the value of treatment and benefit in mortality are not compromised because of potential differences in the biology of detected lesions.

In applying this view, other than effects on CRC mortality and disease stage, there are 3 types of readily determined outcomes that inform the value of a new test: accuracy, acceptability, and impact on other screening program outcomes when applied in a screening context (see Phased Evaluation, below). Such intermediate/surrogate outcomes facilitate the prediction of benefit provided that the new test is directly compared with a test that has been proven to be effective on an intention-to-treat basis, ie, based on an approach that, among other things, takes into account imperfect adherence and overcomes other sources of bias.[1,20,21]

STUDY DESIGN FOR COMPARING TESTS

Accuracy can be assessed through case-control and cohort studies using the framework illustrated in Figure 1. This framework can be adapted to any phase of evaluation, from prescreening assessment to mass population application.

Choice of Comparator Test

The first and well characterized, noninvasive test (in terms of effectiveness) is the guaiac-based fecal occult blood test (gFOBT) Hemoccult (and variants, particularly Hemoccult II; Beckman Coulter Inc, Pasadena, Calif). The screening outcomes achieved with this gFOBT represent the *minimum* that needs to be achieved, because the effect of gFOBT on mortality is modest. The more advanced technology provided by fecal immunochemical tests for hemoglobin (FIT) provides better accuracy, including improved sensitivity for adenomas as well as CRCs and better acceptability when evaluated on an intention-to-screen basis. Population-based and case-control studies support the value of this technology.[22-29] Further studies from the Netherlands[16] confirm the value of FIT in a population RCT when analyzed on an intention-to-screen basis relative to the gFOBT Hemoccult II. This evidence has led to recommendations that FIT replace gFOBT.[15,30] Therefore, a *well studied* FIT sets a new standard against which new tests can be judged.[31] FIT technology tends to have a better capacity to detect adenomas than gFOBT, and repeated testing improves detection.[32,33]

Because population screening trials with flexible sigmoidoscopy (FS) have now been reported,[5] this screening test will serve as a useful comparator for the detection of preinvasive lesions.

The experts concluded that colonoscopy serves to estimate the accuracy of a new test; however, without RCT intention-to-screen evidence of effectiveness, the effectiveness of a new noninvasive test cannot be deduced if it is assessed relative to colonoscopy only. However, as results emerge from the currently underway population screening trials evaluating colonoscopy, we will be able to use colonoscopy as a comparator knowing its benefit to mortality in an unbiased setting.

EVALUATION OF ACCURACY

Clinical accuracy (sensitivity, specificity, and predictive values) is crucial to whether a new test is fully evaluated in