HOFF, G., 2004, CRC screening review of the evidence and suggestions on when and how to move on from randomized trials to screening programmes

This is an electronic version of an article published in *Scandinavian journal of gastroenterology*, Vol. 39, No. 2, 2004, p. 99-103, available online at:
http://www.informaworld.com/smpp/content~content=a713799771

CRC Screening: Review of the Evidence and Suggestions on When and How to Move on from Randomized Trials to Screening Programmes

Colorectal cancer (CRC) is the number one incident cancer in Europe for men and women collectively and the second most common cause of cancer deaths in Europe and in the USA (1, 2). Early diagnosis remains crucial for the outcome of surgery, which is the only real option for cure. Norway has recently become one of the highest incidence countries in the world, having surpassed all the other Nordic countries, the USA and even the three European countries where national 'gold standard' colonoscopy screening has already been introduced (Italy, Germany and Poland) (1, 3, 4). Colonoscopy screening has been recommended for some time in the USA (2). The situation for CRC in the European Union is now considered so disturbing that the EU Commissioner on Health has recommended screening programmes for faecal occult blood (FOBT) to be considered in all its member states (5)—a recommendation that is not shared by the World Health Organization (WHO) (6). At the request of the Secretary of Health and Social Affairs in Norway, the Norwegian Centre for Medical Technology Assessment held a conference in 2001 with mostly Cochrane people invited. A report emerging from this meeting concluded that national CRC screening could not be recommended at the time (7). The report listed a number of problems to be looked into, and recommended awaiting the results of ongoing randomized trials on flexible sigmoidoscopy expected to be available in 2004–2006. In view of the imminent development within the European Union it may be of general interest to look into the obvious differences in attitudes to implementation of screening recognized in several countries, but highlighted by Norway emerging as a prosperous non-screening high-risk country where screening is affordable and where traditionally there has been a high attendance rate for other screening services (8).

New screening modalities are being developed, primarily virtual colonoscopy and the search for suitable molecular markers in faeces, but presently there are three CRC screening modalities to be considered: FOBT, flexible sigmoidoscopy and colonoscopy.

FOBT

This is the only CRC screening modality that has been subjected to adequately sized randomized controlled trials (RCTs) with long-term follow-up results (9–11) (Table I). Although sensitivity for strictly asymptomatic CRC is less than 30% for a single screening round (12), programme sensitivity has been estimated to be more than 60% (13). In two studies randomizing from population registries (Denmark and UK), biennial screening with non-rehydrated slides showed a CRC mortality reduction of 15%–18% after approximately 10 years (10, 11). Collectively and individually for each of the three RCTs on FOBT screening, the CRC mortality reduction was significant, and there is no longer any doubt that FOBT does reduce deaths from CRC. It is also clear that there is no evidence of an excess of non-CRC mortality in the screening arm when expressed as relative risk. As a statistician, I would be perfectly happy with that, but reading Table I as a potential screenee I would be disturbed by the excess number of non-CRC deaths (+198, mainly cardiovascular deaths) in the screening groups. This number more than outweighs the benefit expressed by a reduced number of CRC deaths (–143). As a screenee, I would therefore not be very enthusiastic about the apparent prospects and not exactly 'dying' (literarily) to get an easy, premature, non-cancer exit from life. It is not a question of avoiding a miserable CRC death and saving it for an easier cardiovascular exit at a ripe old age. These deaths are all within the same 8–13 years' range of post-screening follow-up. Numbers needed to harm (NNH) to save one life from CRC do not look favourable in a 10-year perspective (14, 15). It is no consolation that these differences will even out with time post-screening, inevitably reaching 100% all-cause mortality in both groups as life itself is 100% lethal. With an absolute life-time risk of about 5% of getting CRC, a 50% mortality gives a 2.5% life-time risk of dying from it. A 16% CRC mortality reduction through an FOBT screening programme will reduce this risk to 2.1%, i.e. a miserable 0.4% risk reduction for the whole screening group (intention-to-screen analysis), possibly up to 0.6% for those attending. So, the contents of the scientific basis for FOBT screening will obviously be hard to sell despite being the only modality proven, through RCTs, to reduce deaths from our most common cancer.

Flexible sigmoidoscopy

Using 'any adenoma' at flexible sigmoidoscopy (FS) screening as criterion for a positive test and threshold for work-up, FS has been estimated to have a 70% sensitivity for advanced neoplasia (i.e. cancer, adenoma $\geq$10 mm, severe dysplasia or villous components) (16). There are three large-scale RCTs on FS screening having presented their baseline findings (17–19) (Table II), one of which is the Norwegian

Table I. Colorectal cancer (CRC) and non-CRC deaths at 8–13 years follow-up in FOBT screening studies

	Screening frequency	No. of persons included	No. of CRC deaths (Ctr. group minus screening group)*	No. of non-CRC deaths (Ctr. group minus screening group)*
Minnesota (9)	†Annual	15570	82 (−39)	3279 (+60)
	Control	15394	121	3219
Nottingham (10)	Biennial	75253	360 (−60)	12264 (+169)
	Control	74998	420	12095
Funen (11)	Biennial	30967	205 (−44)	6023 (−31)
	Control	30966	249	6054
Total	Screening groups	121790	647 (−143)	21566 (+198)
	Control groups	121358	790	21368

†An additional arm with biennial screening omitted as no effect was shown on CRC mortality at this stage of follow-up.
*(−xxx) indicates 'no. of lives saved through screening'. (+xxx) indicates 'no. of excess deaths in screening groups'.

Table II. Baseline findings in three flexible sigmoidoscopy (FS) screening studies in Norway, UK and Italy with population coverage 65%, 39% and 10%, respectively. Findings in age groups 55–64 years only (percent)

				National CRC incidence 2000, ASR (world)	
	Any adenoma	*High-risk adenoma	Carcinoma	Men	Women
NORCCAP (Norway) (19)	1631 (19)	423 (5.0)	31 (0.4)	40.0	33.8
FlexiScope (UK) (17)	4931 (12)	1905 (4.7)	131 (0.3)	35.4	25.3
SCORE (Italy) (18)	1070 (11)	120 (1.2)	47 (0.5)	35.3	24.0

*Defined as adenoma ≥10 mm diameter, with villous components or severe dysplasia.

Colorectal Cancer Prevention (NORCCAP) trial (20). Five-year follow-up results from these studies are expected in 2004–2006. In fact, we already have long-term follow-up results from one small FS screening study—the Telemark Polyp Study no. I (TPS-I) (21). This study showed an 80% reduction in accumulated CRC incidence rate after 13 years follow-up (intention-to-screen analysis). TPS-I was too small for analysis on CRC mortality outcome, but it showed the same trend as the FOBT studies on all-cause mortality: an excess of non-CRC deaths in the screening arm, this time reaching significance level (Fig. 1) (22). Apart from its small size, there are other weaknesses in this study which have been dealt with in a separate publication (22), but it formed a hypothesis that screening may have an unfavourable influence on lifestyle and lifestyle-related disease, the 'health certificate effect' which has been described in other contexts (23). This is now being looked into in the 'lifestyle sub-study' of the NORCCAP trial, results expected in 2005. The TPS-I obtained an impressive 81% attendance rate (24). Similar to the British FS study (25) there were no untoward psychological effects of screening in TPS-I as measured by the HADS questionnaire and GHQ-26 registrations (26). In fact, a beneficial effect on general well-being seemed to last for at least 1 year post-screening. This may fit well with a 'health certificate effect', a feeling of being invincible and a trend towards being less receptive to smoke cessation campaigns and lifestyle advice as was observed at 13 years follow-up in TPS-I (22).

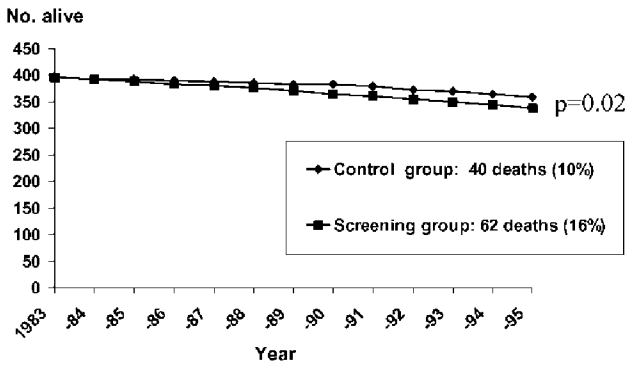

Fig. 1. Survival curve for Telemark Polyp Study no. I (TPS-I).

Colonoscopy

This is the gold standard method for the work-up of screen positives and patients with symptoms. It has an estimated sensitivity for CRC of >90% (27). There are no RCTs on screening colonoscopy with mortality or CRC incidence as end-points. The US National Polyp Study testing the effect of colonoscopy and polypectomy on CRC incidence is not an RCT, but the investigators claim to have demonstrated a CRC reduction comparing their data with two control populations; one was historical SEER data, the other control group was not only historical but from another continent (UK) (28). With significant differences in time trends, between countries and even within countries, basing screening recommendations on these kinds of data is bound to create heated mammography-

like debates on CRC screening once such programmes start to come out with their (inevitably) uncontrolled results. It has been suggested that an RCT on colonoscopy should use 'best clinical practice' as a control group, i.e. an arm with annual or biennial FOBT screening. There are two obvious disadvantages with this. First, any possibly increased risk of non-CRC deaths caused by screening will not be detectable. Second, the number of inclusions needed in the colonoscopy arm will have to be considerably larger than when using a no-screening control group. In addition, it is highly questionable if any screening modality showing any number of excess deaths in the intervention arm deserves the designation 'best clinical practice'.

Discussion

Screenees attend primarily to be reassured that they have no lesions and nothing to worry about, in which case only colonoscopy will do. Providers of a screening service, on the other hand, want to pick up the highest possible proportion of prevalent early cases in a population, in which case attendance is crucial and colonoscopy requiring extensive bowel preparation may not be the best choice. It has been said that the best screening method is the one that is being done, but results from RCTs are only available for FOBT screening and are not very encouraging looking at the unaccounted trend of increased cardiovascular death-risk supported by data from TPS-I (22).

Symptoms of CRC are unspecific and they often appear at an advanced stage of the disease. Still, greater awareness may contribute to improvement on 'patients delay', 'doctors delay' and 'hospital delay' (29, 30). Awareness campaigns may also make people pay more attention to and seek advice on possible risk factors, particularly familial predisposition that may qualify for surveillance.

Many of us truly believe that screening is presently the best method for improving the outlook for CRC in countries where sporadic CRC contributes more than 80% of the total burden of CRC, but at present there is not enough good quality data to recommend this for average risk individuals. Reports that FS and colonoscopy have a higher sensitivity for advanced colorectal lesions than FOBT is not sufficient evidence for recommending endoscopy screening. This kind of extrapolation from the only RCT results we have on CRC screening (FOBT) is at least unusual and possibly unacceptable from a scientific point of view. Drawing a parallel with the pharmaceutical industry it would be similar to a situation allowing, for example, proton-pump inhibitors (PPI) to be introduced on to the market without RCTs, knowing that PPIs were well tolerated and more potent than histamine-2 receptor blockers. This would never have happened. Why should the requirements for documentation be less for introducing screening methods than for marketing drugs? We are even talking about addressing whole populations above a certain age and not only dyspeptic patients needing acid suppression.

Research on screening outcome is different from research on response to drugs or therapeutic procedures. The outcome of a study on screening in one population may be less representative for another population than the results from therapeutic trials, the reason being that screening outcome is dependent on attendance rates, people's knowledge of risk factors and how they respond to this knowledge, the reputation of work-up colonoscopy and cancer treatment in the community in addition to (and linked to) cultural and socioeconomic differences. One possible scenario of attendance-dependent self-selection is shown in Fig. 2. The broken diagonal line represents an ideal situation where 100% attendance gives a 100% pick-up rate of recognizable lesions by means of an ideal (so far non-existent) screening tool. But we may try to approach this ideal curve by adding one or more inferior screening modalities into a multi-modality package. The more complex the package, the higher the risk of compromising the attendance rate so desperately needed for success. Even if we had the ideal screening tool, the diagnostic yield according to attendance rate is more likely to follow the drawn, curved line than the interrupted, diagonal line. The first to attend will usually be those strongly motivated through knowledge of familial predisposition, thus giving a high pick-up rate for advanced lesions in programmes with a low population coverage, such as the FS in Italy (10% population coverage) (18). At the other end of the attendance rate scale, you are likely to find another high-risk group, i.e. the reluctant compliers who are high-risk by effect of lifestyle. In the NORCCAP FS trial in Norway (65% population coverage) the drop in attendance caused by adding FOBT to FS gave a diagnostic loss (intention-to-diagnose analysis) that did not justify the addition of FOBT at this particular range of attendance (19). This might have come out differently in another country and with another attendance rate (Fig. 2). Any country considering itself economically capable of establishing national screening programmes can also afford their own RCTs and management studies. With an increasing political and public pressure 'to do something

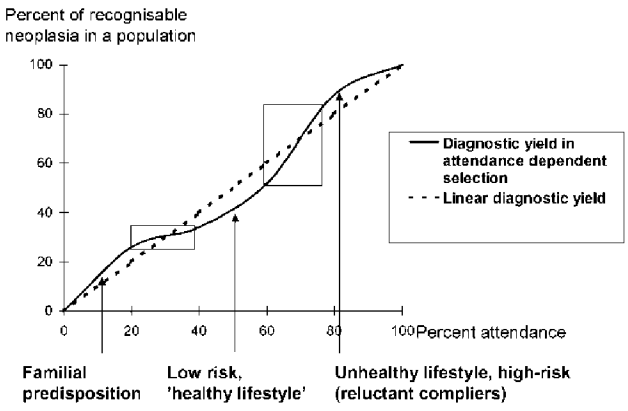

Fig. 2. Scenario of attendance-dependent self-selection (drawn line) showing a greater diagnostic loss by any given attendance reduction in the 60%–80% range than in the 20%–40% range.

about this terrible disease' many feel that time is running out and give up doing a proper RCT. Accepting that there is not sufficient capacity to go countrywide overnight, a stepwise, randomized introduction of screening may be the solution, e.g. first randomize an adequate sample from the population registry for FOBT screening and wait 5 years before expanding to a larger sample, etc. Then you will at least have the 5-year survival data for the population you are addressing before moving on. This type of approach is now being planned in some countries.

In a world where euthanasia is accepted in some countries and suicide may be understandable, we must accept that people have the right to choose not to participate in national screening programmes that are emerging. The aim should therefore be to render our citizens capable of declaring informed consent on participation or non-participation, which requires giving the whole panorama of pros and cons of screening to the public; what are the indisputable facts, what are the 'facts' that we do not believe in, what are just expert opinions and what do we simply not know. A paternalistic attitude towards the public, suppressing uncertainties for the good cause of obtaining high attendance and omitting undesirable observations with poor level of evidence either way, may bounce back after a few years. The epidemiological data on CRC has raised awareness for action in many countries. We should not be so trigger-happy and impatient that we skip the RCT phase that should precede the introduction of any new treatment if at all possible. Also, sustaining a satisfactory screening programme over time requires meticulous attention to delivery of services needed right through to colonoscopic work-up and treatment. This includes building up high quality performance all along and securing a good reputation in the community for the services given. Although many cost-benefit analyses have concluded that CRC screening is probably at least as effective as already established screening programmes like mammography, we have to do the job better than our mammography predecessors. We must avoid building screening 'castles' on scientifically speaking 'slippery slopes'. There is nothing more persuasive to the public, professional critics and health care providers than *good, robust data referable to the target population.*

References

1. http://www-dep.iarc.fr/globocan/globocan.html.
2. Levin B, Smith RA, Feldman GE, Colditz GA, Fletcher RH, Nadel M, et al. Promoting early detection tests for colorectal carcinoma and adenomatous polyps. Cancer 2002;95:1618–25.
3. Svensson E, Grotmol T, Hoff G, Langmark F, Norstein J, Tretli S. Trends in colorectal cancer incidence in Norway by gender and anatomic site: an age-period-cohort analysis. Eur J Cancer Prev 2002;11:489–95.
4. Butruk E, Regula J, Polkowski M, et al. National colorectal cancer screening programme in Poland. Endoscopy 2002;34: 939–40.
5. http://europa.eu.int/comm/health/ph_determinants/genetics/keydo_genetics_en.htm.
6. WHO Executive Summary: National Cancer Control Programmes, Policies and managerial guidelines, 2nd ed. WHO; 2002.
7. Norwegian Centre for Medical Technology Assessment, Screening for colorectal cancer. SMM report no. 3/2001. In Norwegian.
8. Wang H, Kåresen R, Hervik A, et al. Mammography screening in Norway: results from the first screening round in four counties and cost-effectiveness of a modeled nationwide screening. Cancer Causes Control 2001;12:39–45.
9. Mandel JS, Bond J, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. N Engl J Med 1993;328:1365–71.
10. Hardcastle JD, Chamberlain JO, Robinson MHE, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet 1996;348:1472–7.
11. Kronborg O, Fenger C, Olsen J, et al. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet 1996;348:1467–71.
12. Ahlquist DA, Shuber AP. Stool screening for colorectal cancer: evolution from occult blood to molecular markers. Clin Chim Acta 2002;315:157–68.
13. Kronborg O. Screening for colorectal cancer in the average-risk population. Sem Colon Rect Surg 2002;13:16–30.
14. Osiri M, Suarez-Almazor ME, Wells GA, Robinson V, Tugwell P. Number needed to treat (NNT): implication in rheumatology clinical practice. Ann Rheum Dis 2003;62:316–21.
15. Robinson MHE, Hardcastle JD, Moss SM, Amar SS, Chamberlain JO, Armitage NCM, et al. The risks of screening: data from the Nottingham randomized controlled trial of faecal occult blood screening for colorectal cancer. Gut 1999;45:588–92.
16. Liebermann DA, Weiss DG. One-time screening for colorectal cancer with combined fecal occult-blood testing and examination of the distal colon. N Engl J Med 2001;345:555–60.
17. UK Flexible Sigmoidoscopy Screening Trial Investigators, Single flexible sigmoidoscopy screening to prevent colorectal cancer: baseline findings of a UK multicentre randomised trial. Lancet 2002;359:1291–300.
18. Segnan N, Senore C, Andreoni B, et al. Baseline findings in the Italien multicenter randomised controlled trial of 'once-only sigmoidoscopy'–SCORE. JNCI 2002;94:1763–72.
19. Gondal G, Grotmol T, Hofstad B, Bretthauer M, Eide TJ, Hoff G. The Norwegian Colorectal Cancer Prevention (NORCCAP) screening study: baseline findings and implementations for clinical work-up in age groups 50–64 years. Scand J Gastroenterol 2003;38:635–42.
20. Bretthauer M, Gondal G, Larsen IK, Carlsen E, Eide TJ, Grotmol T, et al. Design, organisation and management of a controlled population screening study for detection of colorectal neoplasia. Attendance rates in the NORCCAP study (Norwegian Colorectal Cancer Prevention). Scand J Gastroenterol 2002;37:568–73.
21. Thiis-Evensen E, Hoff GS, Sauar J, Langmark F, Majak BM, Vatn MH. Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I. Scand J Gastroenterol 1999;34:414–20.
22. Hoff G, Thiis-Evensen E, Grotmol T, Sauar J, Vatn MH, Moen IE. Do undesirable effects of screening affect all-cause mortality in flexible sigmoidoscopy programmes? Experience from the Telemark Polyp Study 1983–1996. Eur J Cancer Prev 2001;10: 131–7.
23. Tymstra T, Bieleman B. The psychosocial impact of mass screening for cardiovascular risk factors. Fam Pract 1987;4:287–90.

24. Hoff G, Vatn M, Gjone E, Larsen S, Sauar J. Epidemiology of polyps in the rectum and sigmoid colon. Design of a population screening study. Scand J Gastroenterol 1985;20:351–5.
25. Wardle J, Taylor T, Sutton S, Atkin W. Does publicity about cancer screening raise fear of cancer? Randomised trial of the psychological effect of information about cancer screening. Br Med J 1999;319:1037–8.
26. Thiis-Evensen E, Wilhelmsen I, Hoff GS, Blomhoff S, Sauar J. The psychologic effect of attending a screening program for colorectal polyps. Scand J Gastroenterol 1999;34:103–9.
27. Hosokawa O, Shirasaki S, Kaizaki Y, Hayashi H, Douden K, Hattori M. Invasive colorectal cancer detected up to 3 years after a colonoscopy negative for cancer. Endoscopy 2003;35:506–10.
28. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. N Engl J Med 1993;329: 1977–81.
29. Carter S, Winslet M. Delay in the presentation of colorectal carcinoma: a review of causation. Int J Colorectal Dis 1998;13: 27–31.
30. Langenbach MR, Schmidt J, Neumann J, et al. Delay in treatment of colorectal cancer: multifactorial problem. World J Surg 2003; 27:304–8.