available at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejconline.com

Cancer screening: Evidence and practice in Europe 2008

Matti Hakama[a], *Michel P. Coleman*[b], *Delia-Marina Alexe*[b], *Anssi Auvinen*[a,c,*]

[a]*Tampere School of Public Health, FI-33014 University of Tampere, Finland*
[b]*Non-communicable Disease Epidemiology Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK*
[c]*Finnish Cancer Institute, Liisankatu 21 B, FI-00170 Helsinki, Finland*

ARTICLE INFO

Article history:
Received 21 January 2008
Received in revised form 6 February 2008
Accepted 12 February 2008
Available online 17 March 2008

Keywords:
Screening
Neoplasms
Randomised controlled trial
Programme evaluation

ABSTRACT

We examine the theoretical basis of screening, followed by an evaluation of screening initiatives from a population health perspective and a discussion of the organisation of mass screening programmes. Evidence for the effectiveness of screening by primary site from both randomised trials and evaluation of service screening is summarised and the existing cancer screening programmes in the European Union are described. Sufficient evidence from several randomised trials to demonstrate mortality reduction exists for breast cancer and colorectal cancer screening. At least one trial has shown efficacy with a mortality endpoint in screening for hepatocellular carcinoma and oral cancer. Randomised trials have demonstrated a lack of mortality effect in lung cancer screening based on chest X-ray and sputum cytology. Despite the lack of randomised trials, population screening for cervical cancer with cytological smears has been convincingly shown to reduce cervical cancer incidence and mortality.

1. Introduction

Screening can be defined as the application of a simple test to identify early, asymptomatic disease. The primary purpose of cancer screening is to reduce cancer mortality. In addition, screening has other important consequences, including the use of economic resources (health expenditure) and potential quality of life effects, whether positive or negative.

Since cancer is always a potentially lethal disease, the primary goal of cancer screening (and subsequent treatment) is extending lives. The ultimate effectiveness of cancer screening is measured in terms of mortality reduction and this aim motivates initiation, management and evaluation of cancer screening as a public health policy.

Cancer screening requires a detectable pre-clinical phase of the target disease, during which the disease can be treated so that its progression to overt disease is stopped.[1] The duration of the detectable pre-clinical phase is referred to as the 'sojourn time'.[2] Its duration varies, depending on the natural course of the disease, the screening test and diagnostic assessment.

Ideally, a screening programme should reduce the burden of disease in terms of death and morbidity and/or improve the quality of life. The prognosis of screen-detected cancer cases should be better than for those detected clinically, due to earlier detection and treatment. Yet, screening always has also adverse effects. Screen-detected cases often include indolent lesions, some of which would not progress even if untreated.[3–5] Such lesions may fulfil the histological criteria for malignancy, but nevertheless behave clinically in an indolent fashion. Any screening programme will detect such abnormalities; therefore one of the adverse effects of screening is overdiagnosis and over-treatment, i.e. detection and management of disease that would not have been diagnosed during the lifetime of the subject in the absence of screening.

The natural history of disease and treatment outcomes define the limits of screening. If a disease is fully curable

* *Corresponding author:* Address: Tampere School of Public Health, FI-33014 University of Tampere, Finland. Tel.: +358 3 3551 6883; fax: +358 3 3551 6057.
E-mail address: anssi.auvinen@uta.fi (A. Auvinen).
0959-8049/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ejca.2008.02.013

after detection at a symptomatic phase, there is no need for screening. Screening should not be applied to diseases for which no effective treatment is available.

2. Evaluating the effectiveness of screening

A screening programme should have high sensitivity and specificity. Sensitivity is the capacity to detect cases in the pre-clinical detectable phase amongst those screened.[6] Sensitivity sets the limits for achievable effectiveness, since, reduction in disease burden is achieved only through screen-detected disease. Specificity is the ability to correctly identify subjects without the disease. Specificity is important because it is essential to minimise the number of 'false positives' or healthy persons incorrectly identified as having the disease. These and other measures of performance depend on several factors, including screening attendance, the reproducibility of the screening test, the diagnostic procedures used to confirm positive results and the interval between successive screening tests.

Effects of screening can be described in terms of process and outcome measures. In mammography screening for breast cancer, process indicators include coverage of the target population, identification of pre-clinical breast cancer and improvement in the stage distribution of screen-detected disease compared with disease detected in the absence of screening. An evaluation based on process indicators alone is inadequate, however, because while they are necessary requirements for effectiveness, they are not in themselves sufficient.

Case detection by screening and a favourable stage distribution may simply indicate over-diagnosis or length-biased sampling.[7] Over-diagnosis is common in screening for pre-invasive lesions of cervix uteri and for prostate cancer, because of the high prevalence of pre-invasive or indolent lesions during the detectable pre-clinical phase. Initial results from spiral computerised tomography (CT) to screen for lung cancer also suggest over-diagnosis.

Screening detects a disproportionate number of slow-growing cancers compared with normal clinical diagnosis, because slow progression implies a long sojourn time. Thus, screen-detected cancers tend to have longer survival time from diagnosis than clinically detected disease. This effect is referred to as length bias.[7] Therefore, prognostic indicators such as stage are subject to bias as end-points in the evaluation of screening and can be meaningfully assessed only in studies designed to eliminate, as far as possible, the consequences of length bias.

Lead time[8] is the amount of time by which the diagnosis of disease is brought forward in time compared with diagnosis in the absence of screening. By definition, an effective screening programme gives some lead time, because earlier diagnosis is a requirement for achieving the goals of screening. However, lead time alone does not indicate effectiveness, because survival from the time of diagnosis is longer for a screen-detected case than for a clinically detected case, even if screening does not postpone death. Comparison of survival between screen-detected and symptom-detected patients is therefore biased, unless it is corrected for lead time. Such correction methods tend to be crude and, in general, survival is not a valid indicator of the effectiveness of screening.

Because process indicators cannot provide an estimate of effectiveness, evaluation of cancer screening should be based on cancer mortality. However, screening programmes also affect morbidity and quality of life. Even though not commonplace, such effects should also be considered in screening decisions, to supplement knowledge about effectiveness.

A randomised controlled trial (RCT), with mortality as its end-point, is the optimal and often the only valid means of evaluating the effectiveness of a screening programme. Cohort and case–control studies are often used in evaluating screening programmes. Most evidence on the effectiveness of screening programmes stems from the comparisons of time trends and geographical differences between populations that were subjected to screening of variable intensity. Non-experimental studies do not provide equally valid estimates of effectiveness as an RCT, however, and do not provide a solid basis for decision-making.

Effectiveness trials also provide efficacy estimates, when corrected for non-response and selection by attendance.[9] For instance, screening for colorectal cancer with faecal occult blood test was evaluated in randomised trials in Denmark and in England. The estimates for effectiveness were 18% in Denmark[10] and 12% in England.[11] After correction for attendance in the first round and selection (mortality difference between non-attenders and controls), the efficacy estimates were 24% in Denmark and 32% in England.

Intervention studies without control groups (also called demonstration projects or misleadingly 'single-arm trials') and other non-experimental designs (cohort and case–control studies) have been proposed for the evaluation of mass screening programmes. Biases are inherent in each of these approaches. A randomised approach must always be considered the gold standard.

In some cases, a screening programme can be introduced as a public health policy in an experimental fashion, with a comparison of randomly allocated screened and unscreened groups. When the programme is introduced, it may not immediately cover the entire population, because of limited resources. Under such circumstances, provision of screening may be restricted to a randomly allocated segment of the population, instead of an arbitrarily selected group. As long as the resources allow only a proportion of the population to be covered, it is ethically acceptable to carry out a randomised trial, because the trial does not withhold screening from anybody, but gives *a priori* an equal chance to be included in the screening programme to everyone in the target population. In this context, the equipoise (lack of firm evidence for or against an intervention), which is an ethical requirement for conducting a randomised trial, gradually disappears as evidence is accrued within the programme. For those planning public health services, this approach will provide the optimal knowledge base for accepting or withholding new activities in a specific context.

3. Organising a screening programme

Screening is a chain of activities that starts with defining the target population and extends to the treatment and follow-up of the screen-detected patients. A screening programme consists of several interlinked activities. Various cancer screening programmes may consist of different components (Table 1).

Screening can be opportunistic (spontaneous and unorganised) or organised (mass screening and screening programmes). The major differences lie in the level of organisation and planning and the systematic nature and scope of the activity. The components described in Table 1 are characteristics of organised screening, as opposed to opportunistic screening.

Major organisational considerations in any screening programme are the age range to be covered and the screening interval (see also Martin-Moreno et al., 2008).[12] For example, in western populations with a similar risk of disease and available resources, cervical cancer screening policies range from annual smears from the start of sexual activity to a cervical smear every 5 years in the ages 30–55 years. Hence, the difference in the cumulative life-time number of tests is 10-fold.

Selective screening involves applying the screening test to only a segment of the population at above-average risk for disease. The purpose of screening only a high-risk group is to reduce the resources required for the programme or to limit the adverse effects of the screening test. A selective screening programme should detect a substantial proportion of the disease in the entire target population, i.e. the majority of the entire disease burden should occur in the screened group.[13] All screening programmes are, of course, selective to some degree, according to age and sex, although the term is more often applied to selection on other risk-defining characteristics, such as carriers of the hepatitis B virus infection in liver cancer screening.

So far, selective screening based on high-risk populations defined by aetiological risk factors has failed in cervical and breast cancer screening. The programme sensitivity has been low and a substantial proportion of the disease in the total target population has occurred in the low-risk group not covered by the screening programme. Existing methods of selective screening, based on risk factors, are not likely to be sufficiently effective to be incorporated in public health policy, except in countries with very low resources, with the alternative of not to screen at all. We now examine the evidence in relation to specific cancers.

Table 1 – Components of cancer screening programmes

1. Definition of target population 2. Identification of individuals 3. Measures to achieve sufficient coverage and attendance, such as personal letter of invitation	Population component
4. Test facilities for collection and analysis of the screen material 5. Quality control programme for both obtaining screen material and its analysis	Test execution
6. Facilities for diagnosis, treatment and follow-up of patients with screen-detected disease	Clinical component
7. A referral system linking the screen, screening unit and clinical facility (responsible for diagnostic examinations following an abnormal screening test and management of screen-detected abnormalities) 8. Monitoring, quality control and evaluation of the programme: follow-up of incidence and mortality in the entire target population, and for both attenders and non-attenders	Coordination

4. Screening for cervical cancer

4.1. *Effectiveness of cervical cancer screening*

The objective of cervical cancer screening is to reduce both cervical cancer incidence and mortality. A successful screening programme detects early, pre-invasive lesions during the pre-clinical detectable phase and is able to reduce deaths by preventing the occurrence of invasive cancer. Diagnostic assessment requires colposcopy examination, with the evaluation of morphological features of the cervix as well as histological assessment.

The value of the Papanicolaou (Pap) smear in reducing the risk of invasive cancer and mortality has been firmly established, and it is estimated that regular screening reduces the risk of cancer by 80%.[14,15] Organised screening programmes for cervical cancer using Pap smears have been shown to be more effective than opportunistic or non-organised screening, because opportunistic screening tends to miss the women at the greatest risk.[16]

Most screening programmes start at between 18 and 30 years of age and are discontinued after 60–70 years of age. In some programmes, the frequency of screening varies according to the woman's initial screening result, either starting with annual screening and increasing the interval after negative results or conversely, initially offering a longer, 3–5 year interval that is shortened if there is any abnormality.[21]

The effectiveness of cytological smears in cervical cancer screening has never been established with the current, methodologically rigorous evaluation criteria. However, there is extensive and consistent evidence showing reductions in both the incidence of and mortality from invasive carcinoma. This requires a well-organised screening programme.

In Finland, the population-based cervical cancer screening programme, which began in 1963, achieved a reduction in the incidence of cancer of 60% at 10 years.[17] In Norway, 2 years after the introduction of a population-based nationwide cervical screening programme in 1995, the incidence of invasive cancer was 22% lower than previously.[18] From 1991–1993 to 1998–2000, the incidence of cervical cancer in women aged 20–69 years decreased by 33% in the UK, while mortality fell by 36%.[19] Conversely, in an area where organised screening had been discontinued, the incidence of invasive cancer increased.[20]

Evaluation of the effectiveness of screening programmes has varied widely, from 0% to 60% in terms of relative risk

reduction. As a measure of absolute effect, the 'number needed to screen' (NNS) indicates the number of subjects that need to be screened in order to prevent one cancer death. The NNS is likely to vary between 600 and 2500 at 10 years, depending on the underlying mortality (baseline risk) and effect of the screening programme. These figures apply to the impact of introducing a screening programme to a previously unscreened population and illustrate the effect of a programme with repeated tests. For a single screening test (an individual attending a screen), the NNS is likely to be an order of magnitude higher.

Other screening methods include direct visualisation of the cervix, liquid-based cytology and screening for the human papilloma virus (HPV). HPV testing may have higher sensitivity.[22–24] Liquid-based cytology does not seem to offer real improvement in accuracy.[25] Visual inspection was recently shown to be an effective method in reducing the risk of invasive disease and death in developing countries.[26]

4.2. *Status of cervical screening in the European Union*

Almost all EU countries have a policy on screening for cervical cancer. There are however, major variations in screening organisation, type of screening activities, age range targeted and screening interval, as well as payment strategies. A review of the situation in 2004 (http://www.cancer.org/downloads/AA/CancerAtlas22.pdf) showed that a national screening programme was in place in the Nordic countries, the UK, Latvia, Slovenia, the Netherlands and Hungary. Sub-national screening programmes were operational in Spain, Portugal, Italy, Romania, Czech Republic, Austria and Belgium. Pilot programmes existed in France, Greece, Ireland and Estonia. No population-based screening programme was in place in Germany, although a screening policy does exist.

In many cases, inadequacies exist in the population targeted, or in registration, evaluation or monitoring, as well as in the choice of screening interval. Thus although the recommended screening interval ranges between 3 and 5 years in most European Union countries for which information is available, in some countries or regions an excessive number of smears is recommended, with consequent potential for over-diagnosis and over-treatment. Similarly, the population covered by the screening programmes varied between 30% in Slovenia and 100% in the Nordic countries and Italy.[16]

According to the European recommendations, cervical cancer screening should be offered on a population basis in organised screening programmes. Pap smear screening for cervical abnormalities should start by the age of 30 at the latest and definitely not before the age of 20.[27]

European guidelines on quality assurance of screening programmes have been developed (http://www.cancernetwork.de/cervical/cerv_guidelines.htm). Centralised data systems are essential for monitoring and evaluating the effectiveness of such programmes.

4.3. *Human papilloma virus (HPV) and cervical cancer*

As noted above, the human papilloma virus (HPV) is the principal cause of cervical cancer. There are several types of HPV, of which at least 20 are regarded as oncogenic (cancer-causing). Commercially available tests based on nucleic acid hybridisation can identify more than 10 different types. No trials comparing the effectiveness of HPV testing with cytological smears have been completed, but preliminary findings indicate that HPV screening is likely to be at least as effective as screening based on Pap smears, although it is also likely to have more adverse effects, due to lower specificity.

The development of a vaccine against the most important HPV strains is likely to have major public health implications. The HPV vaccine has the potential to influence the conditions in which screening operates, possibly reducing the demand for cervical cancer screening by decreasing the risk of disease. This may, however, take a considerable amount of time, perhaps a generation.

5. Screening for breast cancer

In breast cancer screening, the primary target lesion is early invasive cancer, but ductal carcinoma *in situ* is also detected with a frequency that is up to a fifth of that for invasive cancer.

Mammography involves radiological imaging of the breast with one or two views, read by one or two radiologists. A screen-positive finding is a lesion that is suspicious for breast cancer. Two views are likely to increase the sensitivity by approximately 20%, with the greatest incremental benefit for the detection of small cancers amongst women with dense breast tissue. In some screening programmes, two views are used only at the first screening, with only one view (mediolateral oblique) in the subsequent screens. Similarly, double reading appears to increase both the recall rate and the detection of breast cancer by some 10%. Diagnostic assessment requires an initial needle biopsy or excision (open surgical) biopsy.

5.1. *Effectiveness of breast cancer screening*

The effectiveness of mammography screening has been evaluated in 12 screening trials (Table 2). These have shown consistent mortality reductions of 20–35% amongst women in the age range 50–69 years. Such NNS estimates for mammography screening have been of the order of magnitude of 1000. However, in the interpretation of these figures it should be noted that the NNS estimates from randomised trials are based on refined mortality, i.e. subjects at entry are free from cancer and therefore, the mortality is lower than population rates.

In Sweden, for example, the reduction in breast cancer mortality after 15–20 years of follow-up ranged from 12% (in the Stockholm trial) to 18% (Kopparberg and Malmö trials). The Edinburgh trial in Scotland reported that, after 14 years of follow-up, the relative reduction in breast cancer mortality in the intervention group was 21%.

Existing randomised trials have, however, been criticised for methodological weaknesses,[28,29] on the grounds that both the randomisation and exclusions after randomisation were inadequate, leading to a lack of comparability between the trial arms. In a systematic review that excluded studies with possible shortcomings, only two trials were finally evaluated and they showed no benefit of breast cancer screening.[29] It was also argued that breast cancer mortality is not a valid

Table 2 – Screening trials evaluating effects of mammography-based screening on breast cancer mortality

Reference (setting)	Sample size	Age range	Follow-up (years)	Mortality[a] (10^{-5})
Shapiro et al. (1982)[39] (Greater NY)	60,995	40–64	18	23/29
Andersson and Janzon (1997)[40] (Malmö)	42,283	45–70	19	45/55
(Malmö)	17,793	43–49	9	26/38
Hakama et al. (1997)[41] (Finland)	158,755	50–64	4	16/21
Tabár et al. (2000)[42] (Kopparberg)	56,448	40–74	20	27/33
Alexander et al. (1999)[43] (Edinburgh)	52,654	45–64	13	34/42
Miller et al. (2000)[44] (Canada)	39,405	50–59	13	50/49
Miller et al. (2002)[45] (Canada)	50,430	40–49	13	37/38
Nyström et al. (2002)[46] (Östergötland)	76,617	40–74	17	30/33
Nyström et al. (2002)[46] (Stockholm)	60,117	40–64	15	15/17
Bjurstam et al. (2003)[47] (Gothenburg)	51,611	39–59	13	23/30
Moss et al. (2006)[48] (UK)	160,921	40–49	10	17/20

a Mortality in the screening arm versus control arm.

end-point for screening trials. These criticisms would be a serious challenge to the evidence base of mammography screening if they were accepted, but their validity has been rebutted by several investigators. The critics' dismissal of all positive randomised trials, essentially based on a mechanistic evaluation of technical criteria that are of questionable relevance to the results, is generally considered to be inappropriate.

In addition to the trials noted above, several studies have evaluated the service screening programmes. In The Netherlands, a statistically significant reduction in breast cancer mortality was reported following the introduction of mammography screening amongst women aged 50–69 years: compared with the period before the introduction of screening, the reduction was 19% at 11 years of follow-up.[32]

An evaluation of mammography screening in seven Swedish counties, begun between 1978 and 1990 and targeted mostly at women aged 40–69 years, found a significant 32% reduction in breast cancer mortality in those counties with a 10-year history of screening, and an 18% reduction in counties where screening had been in place for a shorter time.[33]

The effectiveness of the breast screening programme in England and Wales was assessed by comparing mortality from breast cancer after the introduction of the programme with that expected in the absence of screening, predicted using an age cohort model.[34] The screening interval was 3 years and women aged 50–69 years were invited. Breast cancer mortality fell by 21% after the introduction of screening, but most of the decline was attributed to improvements in treatment. The estimated reduction in breast cancer mortality gained by screening was 6%.

In Denmark, a significant 25% reduction in non-refined breast cancer mortality was shown within 10 years after the introduction of screening in Copenhagen compared with earlier rates and control areas.

These assessments are based on non-randomised studies, which are more prone to bias than randomised trials. The results are, however, generally consistent with the mortality reduction observed in screening trials, suggesting that the results of the trials are not atypical.

Other screening tests include digital mammography, which has been adopted recently; magnetic resonance imaging, clinical breast examination and breast self-examination. No studies have evaluated the effect of digital mammography on breast cancer mortality, and there are no randomised trials comparing the performance of magnetic resonance imaging with mammography. Likewise, no randomised trials have evaluated the effectiveness of clinical breast examination alone, but it was included in the intervention arm of some trials. It may increase the sensitivity of screening, if used as an ancillary test in a mammography screening programme. Lastly, no reduction in breast cancer mortality has been reported in the two trials that estimated the effectiveness of breast self-examination.[30,31]

5.2. *Status of mammography screening in the European Union*

The Council of Europe recommends population-based, organised mammography screening for breast cancer in women aged 50–69 years, and that screening programmes comply with European guidelines on quality assurance.[27]

Screening programmes are organised either regionally or nationally, incorporating quality assurance mechanisms for both radiology and pathology services. Most programmes target women in the age group 50–69 years, with a 2-year interval between screening tests. In several Northern European countries, participation of around 80% has been achieved, with recall rates of 1–8%.

6. Screening for colorectal cancer

6.1. *Effectiveness of colorectal cancer screening*

Several screening methods are available for colorectal cancer screening, including faecal occult blood testing, sigmoidoscopy, colonoscopy and double-contrast barium enema.

Faecal occult blood (FOB) testing is based on the detection of haemoglobin in stools using guaiac-impregnated patches, where an oxidative reaction results in a colour change that is detectable on inspection. The most commonly used test, Hemoccult II®, is not specific for human blood and so may yield false positives in people who have recently eaten undercooked meat. Other tests are available that detect human haemoglobin immunologically, but they are also more expensive. Rehydration (adding water to the specimen) can be used

to increase the detection rate, but this also leads to increased number of false positive results. For screening, two specimens are usually obtained on three consecutive days.

The effectiveness of FOB screening has been evaluated in several randomised trials. Both 1- and 2-year screening intervals have been used and most studies have targeted age groups 45–75 years. Three randomised trials evaluating incidence and mortality have been completed (Table 3). They consistently find a 6–18% reduction in mortality with biennial screening. A recent meta-analysis of the trials estimated the pooled reduction in mortality to be 15% for biennial screening, with a 25% effect amongst screening attenders.[35] It was estimated that the NNS ranges between 215 and 1250.

In the Nottingham trial, for instance, using biennial FOB (Hemoccult II) tests and three to six rounds of screening, 2.1% of individuals were screen-positive at the first and 2.7% at subsequent rounds.[36] Adenomas were identified in 0.8–1% of those screened, with the corresponding figure 0.2–0.5% for cancer. There was, however, no reduction in colorectal cancer incidence in the screened group (151 versus 153 per 100,000). In contrast, those screened experienced a significant 19% reduction in mortality (70 versus 81 per 100,000) after a median 12 years of follow-up.

In the Danish Fynen study, using a biennial FOB (Hemoccult II) testing protocol over a total of nine screening rounds, 1% of subjects were screen-positive at the first screen and, on average, 1.2% in subsequent screens.[10,37] After a mean follow-up of 14 years, the mortality reduction was 11% (99 versus 110 per 100,000, including deaths from screening-related interventions). No decrease in colorectal cancer incidence was observed (206 versus 202 per 100,000).

Service screening has recently been piloted or launched in several European countries including Finland, France, Italy and the UK. In Burgundy, France, mortality reduction by 16% was achieved at 11 years compared with the neighbouring areas, when FOBT screening was offered to a population of 90,000 subjects.[38] Incidence of colorectal cancer was unaffected.

Other screening tests include flexible sigmoidoscopy, screening colonoscopy and the recently introduced faecal DNA analysis. Compliance with screening sigmoidoscopy has been 50% or lower, while the detection rate is higher than in FOB testing, suggesting higher sensitivity. Several case–control studies have found that having a sigmoidoscopy significantly reduces the risk of mortality from colorectal cancer by about 60% to 80%. These studies do not provide as strong evidence as randomised trials, however, because selection bias and other systematic errors may affect the results. Therefore, the reduction in mortality achievable with sigmoidoscopy remains unclear. A population-based randomised trial is under way in Norway, comparing one sigmoidoscopy with no intervention in 20,000 subjects aged 50–64. This will provide important new information on this issue.

Screening colonoscopy has the advantage of visualising the entire colon, but the procedure is expensive, involves substantial discomfort, and has a risk of complications such as perforation of the bowel (reported in 1–2 patients per 10,000). No trials have evaluated the effectiveness of screening colonoscopy, but demonstration projects, which lacked control groups, have reported detection rates of 5–10% for advanced neoplasia (carcinoma or large, dysplastic or villous adenoma), which is approximately one-third higher than in examinations covering only the distal colon.

Faecal DNA analysis has been introduced as a new option for colorectal cancer screening. Early results have shown good sensitivity and acceptable specificity, but no studies assessing effectiveness in reducing mortality have been conducted.

6.2. *Status of screening for colorectal cancer in the European Union*

The Council of Europe recommends faecal occult blood screening for colorectal cancer in men and women aged 50–74.[27] Guidelines on quality assurance of screening are being development by a consortium of experts supported by the European Commission, using similar methods to those employed previously for breast and cervical cancer. At present, a national screening programme exists in Finland. The programme is expanding gradually by randomisation. In 2007, approximately a third of the Finnish population was covered. Regional initiatives have been implemented in several other European Union countries, including France, Italy, Poland, the Netherlands and the United Kingdom.

In conclusion, faecal occult blood testing has been shown to reduce mortality from colorectal cancer in several randomised trials. It appears to be an under-utilised opportunity for cancer control. Other screening modalities are also available, but evidence for their effectiveness is very limited.

7. Screening for prostate cancer

Screening for prostate cancer is mainly based on serum prostate-specific antigen (PSA), which is a serine protease (enzyme) secreted by the prostate gland. It is usually found in low concentrations in serum, with levels increased by prostate diseases such as benign prostatic hyperplasia, prostatitis or prostate cancer.

Two large randomised trials are being carried out, one in Europe and the other in the USA. The European Randomised trial of Screening for Prostate Cancer (ERSPC, http://www.

Table 3 – Randomised trials evaluating mortality effects of colorectal cancer screening based on faecal occult blood testing

Reference (setting)	Sample size	Age range	Length of FU	Mortality (screening/control)
Mandel et al. (1999)[49] (Minnesota, USA)	46,551	50–80	15	5.9 versus 8.3 versus 8.8[a]
Scholefield et al. (2002)[36] (Nottingham, UK)	152,303	45–74	11	7.0 versus 8.1
Jørgensen et al. (2002)[10] (Fynen, Denmark)	61,933	45–75	11	8.3 versus 9.7

a Three arms: annual and biennial screening with control (mortality rates shown respectively).

erspc.org/) includes eight centres in the Netherlands, Finland, Sweden, Italy, Belgium, Spain, Switzerland and France.[50] So far, more than 200,000 men aged 50–74 years have been recruited, and the first mortality analysis is planned in 2010.

In the USA, the Prostate, Lung, Colorectal and Ovary screening trial (PLCO) recruited 76,705 men aged 55–74 years in the prostate screening component between 1993 and 2001.[51] Both serum PSA and digital rectal examination are used as screening tests. No mortality results are available yet.

A trial in Quebec, Canada with 46,500 men aged 45–80 years was commenced in 1988.[52] At 10 years, intention-to-screen analysis revealed no reduction in prostate cancer mortality (54.4 versus 50.4 per 100,000, relative risk = 1.1, 95% confidence interval 0.8–1.4), which may reflect the low compliance with annual PSA testing (24%).

Several ecological studies and time series analyses have been published correlating the frequency of PSA testing (or the incidence of prostate cancer, as a surrogate for PSA testing) with prostate cancer mortality. The results have been inconsistent. The short-comings inherent in these approaches preclude firm conclusions.

The other main screening test for prostate cancer is digital rectal examination. Its impact as a screening test on death from prostate cancer has been evaluated in five case–control studies. These have not yielded consistent results: two indicated a 30–50% reduction in risk, while the other three failed to show a benefit. The lack of a clear effect is thought to be due to the fact that digital rectal examination only detects cancers that are large enough to be palpable. By this stage, however, the cancer has frequently spread beyond the prostate capsule and is no longer curable. Hence, the chief limitation of digital rectal examination is its low sensitivity for the detection of early disease.

In summary, mortality effect of prostate cancer screening by serum PSA or any of the test has not been properly evaluated. The use of serum PSA levels as a screening test should be restricted to randomised trials. Such trials are on-going and should provide important evidence in due course.

8. Screening for other cancers

8.1. *Lung cancer*

The target lesion for lung cancer screening is early, resectable (stage 1) carcinoma. Conclusive diagnosis of early lung cancer is based on biopsy, usually obtained by bronchoscopy for central tumours and excision biopsy for peripheral tumours.

The available screening protocols for lung cancer include screening with chest X-rays, with or without sputum cytology, and spiral low-dose CT. Several trials have been conducted to assess the mortality effect of chest X-rays and sputum cytology compared with chest X-rays alone (Table 4). They have not shown mortality reduction, which may be due to low sensitivity for the detection of early tumours. Several randomised trials are underway to evaluate effectiveness of helical CT in lung cancer screening both in Europe and in the USA. The largest one is the U.S. National Lung Cancer Screening Trial as sub-study of PLCO, followed by the Dutch-Belgian NELSON trial.[53] Smaller studies are on-going in Denmark and Italy. The control arm receives no intervention in the European trials, while chest X-ray is offered in the US.

Table 4 – Randomised trials of chest X-rays with or without sputum cytology in screening for lung cancer

Reference (setting)	Sample size	Length of FU	Mortality (screening/control)
Melamed et al. (1984)[57] (Sloane-Kettering, USA)	10,040	8	2.7 versus 2.7
Levin et al. (1982)[58] (Johns Hopkins, USA)	10,386	8	3.4 versus 3.8
Marcus et al. (2000)[59] (Mayo Clinic, USA)	9211	20	4.4 versus 3.9

Table 5 – Summary of evidence for cancer screening

Primary site	Screening method	Efficacy		Effectiveness	
		Non-randomised	Randomised	Randomised trial	Service screening
Cervical cancer	Pap smear		NA	NA	0–80%
	Visual inspection			NA	NA
	HPV testing		NA	NA	NA
Breast cancer	Mammography		35%	15–25%	6–20%
Colorectal cancer	Faecal occult blood		24%	15%	NA
	Sigmoidoscopy		NA	NA	NA
	Colonoscopy		NA	NA	NA
Lung cancer	Chest X-ray ± sputum cytology		None		NA
	Helical CT		NA	NA	NA
Prostate cancer	Serum PSA		NA	NA	NA
	Digital rectal exam	None	NA	NA	NA
Oral cancer	Visual inspection		NA	20%	NA
Liver cancer	Serum AFP ± ultrasound		NA	20–33%	NA
Ovarian cancer	Ultrasound + CA-125		NA	NA	NA

All trials focus on smokers. The effectiveness of screening based on spiral CT remains unclear.

8.2. *Liver cancer*

Serum alpha-fetoprotein (AFP) levels and ultrasound have been used as a combined screening test for hepatocellular cancer. Two randomised trials in high-risk subjects have been carried out in China amongst chronic carriers of hepatitis B virus, who are at greatly increased risk of liver cancer. The smaller study, amongst 5500 chronic carriers of hepatitis B virus in Qidong County, found a non-significant 17% reduction (1138 versus 1114 per 100,000) with six-monthly AFP tests.[54] The larger trial, including 18,000 people with hepatitis B virus infection from Shanghai, used six-monthly AFP tests and ultrasonography and achieved a 37% reduction in 5-year mortality from liver cancer (83 versus 132 per 100,000).[55]

8.3. *Oral cancer*

A cluster-randomised trial of visual inspection for oral cancer has been conducted in India with more than 190,000 subjects in 13 clusters.[56] After three screening rounds with 3-year intervals (91% participation), a 20% reduction in oral cancer mortality (16 versus 21 per 100,000) was shown. The difference was mainly attributable to high-risk subjects.

8.4. *Ovarian cancer*

The natural history of ovarian carcinoma is not well understood, in particular the relative frequency of cancers developing from benign or borderline lesions or *de novo*. The duration of any detectable pre-clinical phase is also unknown.

Screening tests include transvaginal or transabdominal ultrasound for imaging, and serum CA-125 as a biochemical marker. There is no evidence that ovarian cancer screening can reduce mortality. Preliminary results from non-randomised studies are not encouraging: the sensitivity is low (too many cases missed) and false positive findings are common (too many healthy women identified as having disease). A randomised trial is currently in progress in the USA as part of the PLCO study.

8.5. *Cutaneous melanoma*

Melanoma survival is favourable if detected at an early stage. A substantial proportion of melanomas (approximately a fifth) arise from atypical naevi. Visual inspection can be used to identify early melanoma (or pre-malignant lesions), while diagnostic assessment requires a skin biopsy. No randomised trials have been conducted to evaluate the effect of screening on melanoma mortality.

8.6. *Neuroblastoma*

Screening for neuroblastoma, an uncommon childhood tumour, is based on a urine test for the catecholamine metabolites homovanillic acid and vanillylmandelic acid, which are secreted by most (60–80%) of the tumours. Screened and unscreened cohorts have been compared in Germany, Canada, the UK and Japan to evaluate the effects of screening. Screening has been associated with a 2- to 6-fold increase in the incidence rate, with cases being diagnosed at earlier ages. Unfortunately, this has not been counterbalanced by a reduction in incidence at older ages, and no reduction in mortality or in the occurrence of advanced disease has been demonstrated. Neuroblastoma screening in Japan has been stopped as a result.

8.7. *Gastric carcinoma*

Fluoroscopic imaging (photofluorography) and endoscopy have been used to screen for stomach cancer. Several case–control studies and two cohort studies have evaluated the effect of gastroscopy, but have not provided consistent results. No randomised trials have been reported, so there is insufficient evidence of effectiveness.

9. Conclusion

In summary, establishing the benefits of screening requires evidence on mortality effects from large randomised trials. A summary of the current evidence for cancer screening is provided in Table 5. Screening tests are available for many primary cancer sites, but either their effectiveness has not been evaluated adequately or a lack of effectiveness has been demonstrated. Even when efficacy trials (typically conducted in specialist centres with volunteer subjects) have been successful, mass screening requires pilot studies to demonstrate its feasibility and, once an organised screening programme is in place, continued evaluation is required to ensure that the benefits are maintained. Ideally, this is achieved by incorporating a randomised design, comparing the outcome in subjects who were randomly allocated to early entry to the screening programme with the outcome in those who were included in the programme later. These estimates are highly context-specific and not directly applicable in other European countries with different cancer profiles (see Karim-Kos et al., 2008).[60]

Our assessment is largely consistent with the European Code Against Cancer (www.cancercode.org), which, however,

Table 6 – Number needed to screen in order to prevent one cancer death over 10 years for different cancer sites

Site	RRR[a] (%)	Mortality[b]	Number needed to screen
Cervix	20–40	20–40	600–2500
Breast	20	5–75	700–1000
Colorectal	15	30–60	1100–2200

The estimates apply to starting a screening programme in a previously unscreened population and reflect the impact of a programme with multiple screening rounds.
a Relative risk reduction; estimates represent typical values in randomised trials for breast and colorectal cancer, while for cervix cancer RRR is based on effectiveness estimates for service screening.
b Mortality per 100,000, the range represents rates in unscreened populations; (age-specific rates for 25–60 years for cervical cancer and 50–70 years for breast and colorectal cancer).

deals mainly with the implementation of screening. Therefore, it addresses only screening with established effectiveness. After its latest revision in 2003, new evidence has emerged concerning oral cancer and liver cancer screening, as well as visual inspection in cervical cancer screening. These are, however, not a high priority in Europe.

The effectiveness of service screening ranges from zero to similar to that achieved in randomised trials. This is reflected in estimates of the 'number needed to screen' (NNS), which indicate the number of subjects that need to be screened in order to prevent one cancer death. Given the mortality rates from cancer in Europe, the NNS for cervical cancer may vary from approximately 500 in high-risk populations to 2500 in low-risk countries (Table 6). For breast and colorectal cancers, the NNS is likely to be also within that range, with less variability in both risk and effectiveness estimates (based on randomised trials for the two latter cancers). In priority setting, cervical cancer screening should probably be first considered in most countries, where it has not yet been implemented, in particular Eastern European countries with high rates of the disease (see also Zatoński and Didkowska, 2008).[61]

It has been estimated that approximately 5% of cancer deaths in the Nordic countries could be prevented by breast, cervix and colorectal cancer screening.[62] In terms of life-years gained, about 25,000 life-years (or one year per 1000 persons) could be saved annually in the Nordic countries.

Large randomised trials have shown efficacy of screening in reducing mortality from breast cancer with mammography and colorectal cancer with the FOB test. Screening as a public health policy is justified by sufficient evidence of a reduced risk of invasive cervical cancer with the Pap test or visual inspection and reduced mortality from breast cancer with mammography.

Some evidence from screening trials suggests mortality reduction from liver cancer with the AFP test and ultrasound, and from oral cancer with visual inspection. Both have shown effect mainly in high-risk populations. There is no evidence for the time being that these results are applicable to effective mass screening.

There is sufficient evidence based on randomised trials that screening for lung cancer with sputum cytology and X-ray will not reduce the mortality from the disease. Evaluation of screened and unscreened cohorts for deaths from neuroblastoma showed lack of effect.

No randomised trials have been conducted for any other type of cancer or other screening tests.

Conflict of interest statement

None declared.

Acknowledgements

The article is partly based on 'Cancer Screening' chapter in Encyclopedia of Public Health (Elsevier 2008) by Matti Hakama and Anssi Auvinen. We thank prof. Jan-Willem Coebergh for comments and suggestions.

REFERENCES

1. Cole P, Morrison AS. Basic issues in cancer screening. In: Miller AB, editor. *Screening in cancer. UICC Technical Report Series*, vol. 40. Geneva: UICC; 1978. p. 7–39.
2. Day NE, Walter SD. Simplified models for screening: estimation procedures from mass screening programmes. *Biometrics* 1984;**40**:1–14.
3. Breast Cancer Screening. *IARC handbooks of cancer prevention*, vol. 7. Lyon: IARC Press; 2002.
4. Cervix Cancer Screening. *IARC handbooks of cancer prevention*, vol. 10. Lyon: IARC Press; 2005.
5. Hugosson J, Aus G, Becker C, et al. Would prostate cancer detected by screening with prostate specific antigen develop into clinical cancer if left undiagnosed. *BJU Int* 2000;**85**:1978–84.
6. Hakama M, Auvinen A, Day NE, Miller AB. Sensitivity in cancer screening. *J Med Screen* 2007;**14**:174–7.
7. Feinleib M, Zelen M. Some pitfalls in the evaluation of screening programs. *Arch Environ Health* 1969;**19**:412–5.
8. Hutchison GB, Shapiro S. Lead time gained by diagnostic screening for breast cancer. *J Natl Cancer Inst* 1968;**41**:665–81.
9. Cuzick J, Edwards R, Segnan N. Adjusting for non-compliance and contamination in randomised clinical trials. *Stat Med* 1997;**16**:1017–29.
10. Jørgensen O, Kronborg O, Fenger C. A randomised study on screening for colorectal cancer using faecal occult blood testing: results after 13 years and seven biennial screening rounds. *Gut* 2002;**50**:29–39.
11. Hardcastle J, Chamberlain J, Robinson M, et al. Randomised controlled trial of faecal occult-blood screening for colorectal cancer. *Lancet* 1996;**348**:1472–7.
12. Martin-Moreno JM, Soerjomataram I, Magnusson G. Cancer causes and prevention: A condensed appraisal in Europe in 2008. *Eur J Cancer* 2008;**44**:1390–403.
13. Hakama M. Selective screening by risk groups. In: Prorok PC, Miller AB, editors. *Screening for cancer I – general principles on evaluation of screening for cancer and screening for lung, bladder and oral cancer. UICC Technical Report Series*, vol. 78. Geneva: UICC; 1984. p. 71–80.
14. Screening for cancer of the uterine cervix. Geneva: WHO; 1986.
15. Läärä E, Day NE, Hakama M. Trends in mortality from cervical cancer in the Nordic countries. *Lancet* 1987;**i**:1247–9.
16. Anttila A, Ronco G, Clifford G, et al. Cervical cancer screening programmes and policies in 18 European countries. *Br J Cancer* 2004;**91**:935–41.
17. Nieminen P, Kallio M, Hakama M. The effect of mass screening on incidence and mortality of squamous and adenocarcinoma of cervix uteri. *Obstet Gynecol* 1995;**85**:1017–21.
18. Nygard JF, Skare GB, Thoresen SO. The cervical cancer screening programme in Norway, 1992–2000: changes in Pap smear coverage and incidence of cervical cancer. *J Med Screen* 2002;**9**:86–91.
19. Canfell K, Sitas F, Beral V. Cervical cancer in Australia and the United Kingdom: comparison of screening policy and uptake, and cancer incidence and mortality. *Med J Aust* 2006;**185**:482–6.
20. Lynge E, Clausen LB, Guignard R, Poll P. What happens when organization of cervical cancer screening is delayed or stopped? *J Med Screen* 2006;**13**:41–6.
21. Olesen F. Detecting cervical cancer: the European experience. *Hong Kong Med J* 1999;**5**:272–4.
22. Bulkmans NW, Berkhof J, Rozendaal L, et al. Human papillomavirus DNA testing for the detection of cervical

intraepithelial neoplasia grade 3 and cancer. *Lancet* 2007;**370**:1764–72.
23. Mayrand MH, Duarte-Franco N, Rodrigues I, et al. Human papilloma DNA versus Papanicolau screening test for cervical cancer. *New Engl J Med* 2007;**357**:1579–88.
24. Naucler P, Ryd W, Törnberg S, et al. Human papillomavirus and Papanicolaou tests to screen for cervical cancer. *New Engl J Med* 2007;**357**:1589–97.
25. Ronco G, Cuzick J, Pierotti P, et al. Accuracy of liquid-based versus conventional cytology. *Br Med J* 2007;**335**:28.
26. Sankaranarayanan R, Esmy P, Rajkumar R, et al. Effect of visual screening on cervical cancer incidence and mortality in Tamil Nadu, India: a cluster-randomised trial. *Lancet* 2007;**370**:398–406.
27. Council Recommendation on Cancer Screening. 2003/0093. Commission of the European Communities Brussels; 2003.
28. Gøtzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* 2000;**355**:129–34.
29. Olsen O, Gøtzsche PC. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2001:CD001877.
30. Thomas DB, Gao DL, Ray RM, et al. Randomized trial of breast self-examination in Shanghai: final results. *J Natl Cancer Inst* 2002;**94**:1445–57.
31. Semiglazov VF, Manikhas AG, Moiseenko VM, et al. Results of a prospective randomized investigation to evaluate the significance of self-examination for the early detection of breast cancer. *Vopr Onkol* 2003;**49**:434–41 [in Russian].
32. Otto SJ, Fracheboud J, Looman CW, et al. Initiation of population-based mammography screening in Dutch municipalities and effect on breast-cancer mortality. *Lancet* 2003;**361**:1411–7.
33. Duffy SW, Tabar L, Chen HH, et al. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. *Cancer* 2002;**95**:458–69.
34. Blanks RG, Moss SM, McGahan CE, Quinn MJ, Babb PJ. Effect of NHS breast screening programme on mortality from breast cancer in England and Wales, 1990–8: comparison of observed with predicted mortality. *Br Med J* 2000;**321**:665–9.
35. Hewitson P, Glasziou P, Irvig L, Towler B, Watson E. Screening for colorectal cancer using the faecal occult blood test, Hemoccult. *Cochrane Database Syst Rev* 2007;**1**:CD001216.
36. Scholefield JH, Moss S, Sufi F, Mangham CM, Hardcastle JD. Effect of faecal occult blood screening on mortality from colorectal cancer: results from a randomised controlled trial. *Gut* 2002;**50**:840–4.
37. Kronborg O, Jörgensen OD, Fenger C, Rasmussen M. Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds. *Scand J Gastroenterol* 2004;**39**:846–51.
38. 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.
39. Shapiro S, Venet W, Strax P, et al. Ten- to fourteen-year effect of screening on breast cancer mortality. *J Natl Cancer Inst* 1982;**69**:349–55.
40. Andersson I, Janzon L. Reduced breast cancer mortality in women under age 50: updated results from the Malmö mammographic screening program. *J Natl Cancer Inst Monogr* 1997;**22**:63–7.
41. Hakama M, Pukkala E, Heikkilä M, Kallio M. Effectiveness of the public health policy for breast cancer screening in Finland: population based cohort study. *Br Med J* 1997;**314**:864–7.
42. Tabár L, Vitak B, Chen HH, et al. The Swedish two-county trial twenty years later. Updated mortality results and new insights from long-term follow-up. *Radiol Clin North Am* 2000;**38**:625–51.
43. Alexander FE, Andersson TJ, Brown HK, et al. 4 years of follow-up from the Edinburgh randomised trial of breast-cancer screening. *Lancet* 1999;**353**:1903–8.
44. Miller AB, To T, Baines CJ, Wall C. Canadian national breast screening study-2: 13-year results of a randomized trial in women aged 50-59 years. *J Natl Cancer Inst* 2000;**92**:1490–9.
45. Miller AB, To T, Baines CJ, Wall C. The Canadian national breast screening study-1: breast cancer mortality after 11 to 16 years of follow-up. *Ann Intern Med* 2002;**137**:305–12.
46. Nyström L, Andersson I, Bjurstam N, et al. Long-term effects of mammography screening: updated overview of the Swedish randomised trials. *Lancet* 2002;**359**:909–19.
47. Bjurstam N, Björneld L, Warwick J, et al. The Gothenburg breast screening trial. *Cancer* 2003;**97**:2387–96.
48. Moss SM, Cuckle H, Evans A, et al. Effect of mammographic screening from age 40 years on breast cancer mortality at 10 years' follow-up. *Lancet* 2006;**368**:2053–60.
49. Mandel JS, Church TR, Ederer F, Bond JH. Colorectal cancer mortality: effectiveness of biennial screening for fecal occult blood. *J Natl Cancer Inst* 1999;**91**:434–7.
50. de Koning HJ, Auvinen A, Sanchez AB, et al. Large-scale randomized prostate cancer screening trials: program performances in the ERSPC and PLCO trials. *Int J Cancer* 2002;**97**:237–44.
51. Andriole GL, Levin DL, Crawford ED, et al. Prostate cancer screening in the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial: findings from the initial screening round of a randomized trial. *J Natl Cancer Inst* 2005;**97**:433–8.
52. Labrie F, Candas B, Cusan L, et al. Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* 2004;**59**:311–8.
53. van Iersel CA, de Koning HJ, Draisma G, et al. Risk-based selection from the general population in a screening trial: selection criteria, recruitment and power for the Dutch-Belgian randomised lung cancer multi-slice CT screening trial (NELSON). *Int J Cancer* 2007;**120**:868–74.
54. Sankaranarayanan R, Ramadas K, Thomas G, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet* 2005;**365**:1927–33.
55. Chen JG, Parkin DM, Chen QG, et al. Screening for liver cancer: results of a randomised, controlled trial in Qidong, China. *J Med Screen* 2003;**10**:204–9.
56. Zhang BH, Yang BH, Tang ZY. Randomized controlled trial of screening for hepatocellular carcinoma. *J Cancer Res Clin Oncol* 2004;**131**:417–22.
57. Melamed MR, Flehinger BJ, Zaman MB, et al. Screening for early lung cancer. *Results of the Memorial Sloan-Kettering study in New York. Chest* 1984;**86**:44–53.
58. Levin ML, Tockman MS, Frost JK, Ball WC. Lung cancer mortality in males screened by chest X-ray and cytologic sputum examination: a preliminary report. *Recent Res Cancer* 1982;**82**:138–46.
59. Marcus PM, Bergstralh EJ, Fagerstrom RM, et al. Lung cancer mortality in the Mayo Lung Project: impact of extended follow-up. *J Natl Cancer Inst* 2000;**92**:1308–16.
60. Karim-Kos HE, de Vries E, Soerjomataram I, et al. Recent trends of cancer in Europe: A combined approach of incidence, survival and mortality for 17 cancer sites since the 1990s. *Eur J cancer* 2008;**44**:1345–89.
61. Zatoński W, Didkowska J. Closing the gap: Cancer in Central and Eastern Europe (CEE). *Eur J Cancer* 2008;**44**:1425–37.
62. Hristova L, Hakama M. Effect of screening for cancer in the Nordic countries on deaths, cost and quality of life up to the year 2017. *Acta Oncol* 1997;**36**(Suppl. 9):1–60.