International Journal of Surgery (2005) 3, 107–112

INTERNATIONAL JOURNAL OF SURGERY

www.int-journal-surgery.com

EDITORIAL

Screening for disease – The good – The bad and the thoughtful

Wilson and Jungner[1] laid down 10 principles of screening in 1968:

1. The disease should be an important public health problem.
2. There should be an accepted treatment for patients with recognized disease.
3. Facilities for diagnosis and treatment should be available.
4. There should be a recognizable latent or early symptomatic stage.
5. There should be a suitable test or examination.
6. The test should be acceptable to the population.
7. The natural history of the disease, including development from latent to declared disease, should be adequately understood.
8. There should be an agreed policy on whom to treat as patients.
9. The cost of case-finding, including diagnosis and treatment of patients diagnosed should be economically balanced in relation to possible expenditure on medical care as a whole.
10. Case-finding should be a continuing process and not a ''once and for all'' project.

No test fulfils all these requisites. The minimum requirements could be, therefore, said to be the following. An ideal screening test should be looking for a disease that has a great impact that could be demonstrably averted by earlier treatment, usually by reducing mortality. It should be highly sensitive and specific. It should be cheap and easily administered. At the very least, it should prevent the devastating effects of advanced disease. In any case, an informed consent should be mandatory before an individual begins screening.

Examples of such tests are rare. Heel-prick blood test for the metabolic disorders such as phenylketonurea and congenital hypothyroidism has probably prevented many cases of mental retardations and cretinism. A more surgical example is ultrasonographic screening for aortic aneurysms. The natural history of abdominal aortic aneurysm (AAA) above a certain size could almost be determined by physical laws alone, making rupture is almost inevitable. So prevention of such a dramatic and painful death is an obvious target. A recent meta-analysis by the US task force[2] suggests that screening reduces AAA-related mortality. Intervention is necessary in less than 5% of the screened group and the number needed to screen to save one life ranges from 350 to 700. An estimate of the financial cost of saving a life-year (after 10 years) comes to a modest £8000/-,[3] less than the cost of a good car! The human cost to someone undergoing screening is complex. If the ultrasonography reveals a large aneurysm, the person needs to decide to either accept the intervention – open or even endoluminal repair with its complication risks, which are not insignificant, or take the much higher risk of rupture. So in theory, the consent for screening would be straightforward and Chancellor of the Exchequer willing, this could be a good example of success of screening.[4]

Screening for cancer

However, screening for a malignant tumour is not so straightforward and the rest of this paper will address the issue of screening for solid tumours with some illustrative examples.

1743-9191/$ - see front matter
doi:10.1016/j.ijsu.2005.06.001

Screening would be effective only for a cancer if its natural history can be significantly modified by treatment. So in theory any cancer that is curable when treated early, usually by surgery, and incurable later, would be suitable for screening. However, we do not fully understand the process of carcinogenesis and metastasis, nor do we have efficient (as opposed to 'reasonably effective') treatment for most common malignant tumours. The other 'problem' is the large difference in biological behaviour of a cancer in people — from rapid progression to spontaneous regression. Let us discuss the success stories first.

The good

Screening for cervical cancer with a PAP smear is the epitome of a screening test. It is cheap, easy to administer, and sensitive. It certainly has resulted in down-staging of cervical cancer and may have even reduced mortality.[5,6] Of course, there is overdiagnosis of cervical intraepithelial neoplasia, but there is little social taboo associated with its diagnosis, its treatment is not mutilating and usually completed in an outpatient setting. Hence, the human cost of this overdiagnosis is not high, if one discounts the high profile and expensive lawsuits from 'missed diagnosis'. But that is the subject of a different and ethical/moral debate.

Screening for colorectal cancer with faecal occult blood (FOB) testing followed by colonoscopy reduces mortality by 11%.[7] The FOB test is simple to administer and only about 5% of those screened needed to undergo colonoscopy. In the Scandinavian study, those who attended all 9 biennial screening rounds had a 43% reduction in mortality and importantly, the survival of patients diagnosed between screens (interval cancers) had better survival than controls. The incidence of colorectal carcinoma did not increase in the screened group, suggesting that there was little, if any overdiagnosis. The success of screening for colorectal cancer suggests that surgery is curative in a significant proportion of colorectal cancers and early surgery has a significantly higher impact than later surgery. The fact that colorectal cancer does not usually recur after first few years of primary treatment corroborates this contention.

Screening for oral cancer by looking in the mouth is an elegant demonstration of common sense! A recent study from India[8] revealed that this simple and easily taught manoeuvre has high sensitivity, high specificity and appears to save lives. When this is implemented it will be a yet another successful battle before the war on the tobacco industry is won and the tobacco plant goes the way of small pox virus (Table 1).

The bad

Neuroblastoma is the second most common childhood solid tumour and is a classic example of natural regression that was unearthed because of screening. Screening for neuroblastoma is easy — a urine test at 6 months of age for vanillylmandelic acid and homovanillic acid (metabolites of the noradrenaline and dopamine). Contrary to all expectations, screening for neuroblastoma in infants only increases the incidence of early-stage neuroblastoma,[9,10] without a concurrent reduction of advanced-stage disease, or of any disease in older children. The cases identified by screening almost exclusively have biologically favourable properties. This leads to unnecessary diagnostic and therapeutic procedures with consequent physical and psychological morbidity, including death from treatment complications. When natural regression was so clearly demonstrated, screening for neuroblastoma was abandoned.

Screening for lung cancer with spiral-CT[11] scans in Japanese populations uncovered many small lung cancers at a rate much higher than population incidence of clinical cancer. Many of these were of course curable, but was the diagnosis-and-cure unnecessary? Only half of these cancers were among smokers compared with 95% of clinical tumours. This discrepancy is too large to ignore[12] and suggests that most of these sub-clinical cancers would never have progressed to clinical disease. It is estimated that even in the highest risk group with the most effective spiral-CT, screening would not be cost effective,[13] using a threshold of about £30000/- per life-year gained. Screening for lung cancer is really an irony when 95% of it could be prevented by tobacco eradication!

Thinking maketh it so

The case of breast and prostate cancers is different and for different reasons. Both diseases are common and share hormonal aetiology in respective sexes, although the exact cause is not known and treatment can have devastating physical and psychological effects.

The story of prostate cancer appears to be much clearer today than ever before. It is the only major cancer in which the effectiveness of surgery itself has been tested. Recent results from the Swedish randomised trial[14–16] of watchful waiting vs. radical prostatectomy proved that surgery reduced

Table 1 Summary of effectiveness of screening for various diseases

	Natural history	Diagnosis of fatal disease at curable stage?	Screening tool	Curative intervention	Human cost
Neuroblastoma	Spontaneous regression or aggressive behaviour	No	Simple urine test	Aggressive and potentially hazardous	Higher than the benefit
Cervical cancer	Slow growth, with some natural regression	Yes	Relatively non-intrusive smear test	Relatively harmless	Lower than benefit
Lung cancer	Aggressive growth	Unlikely	Expensive and possibly hazardous	Aggressive	Very high financial costs
Colorectal cancer	Relatively slow growth and orderly spread	Yes	Simple FOB test	Relatively non-aggressive	Lower than the benefit
Oral cancer	Recognisable precancerous conditions and mainly local spread	Yes	Simple oral inspection	Relatively non-aggressive	Lower than the benefit
Breast cancer	Chaotic growth and spread	Yes, but unpredictable	Complex and subjective	Relatively non-aggressive	Very high and an informed decision to screen or not should be left to those being screened
Prostate cancer	Slow growth	Yes	Simple blood test	Aggressive with many complications	Very high and an informed decision to screen or not should be left to those being screened

disease-specific mortality as well as overall mortality at 8-year follow-up. There is balance of costs: erectile dysfunction (80% vs. 45%) and urinary leakage (49% vs. 21%) after radical prostatectomy, vs. urinary obstruction (e.g., 44% vs. 28% for weak urinary stream) and death (14.4% vs. 8.7%) with observation. It, therefore, should follow that screening for prostate cancer with PSA would work. We need to wait another 7 years until the results of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial are available. In this trial, at the end of the first screening round, 7.5% of men had a positive digital rectal examination and 7.9% had a PSA level higher than 4 ng/m. A large proportion of these men underwent further investigations and 1.4% were diagnosed to have prostate cancer, most usually a localised one. The Quebec-1988 screening trial[17] found that screening reduced mortality from 5 per 1000 to 1.3 per 1000 at 8-year follow-up, a 62% relative risk reduction. The human price to pay, of course, is possibly overdiagnosis and morbidity from radical local treatment, but it appears that, at least in theory, a well informed choice could be available to men.

Breast cancer is the most researched cancer and benefits from a long history of evolution of its management. It is a common cause of death among women aged 45—65, second only to the diseases caused by smoking. Advanced breast cancer is almost always fatal and earlier diagnosis appears to confer a better prognosis. Mechanistic principles would dictate that detection and removal of breast cancer early enough would make it curable, or at least, its treatment less morbid.

Unfortunately, breast cancer does not behave in a mechanistic fashion and its natural history is still an enigma.[18] A large proportion of normal women harbour cancers in their breasts — up to a third by the time they reach 55 years of age — as revealed

by autopsy studies.[19] Only a proportion of these give rise to clinical disease and uncovering the others with more sensitive imaging would, therefore, be unnecessary at best or harmful at worst.

Screening for breast cancer is perhaps the best demonstration of non-generalisability of screening. Breast cancer screening can be performed using X-ray mammography (XRM) and physical examination. The only trial that tested the benefit of one over and above the other[20,21] found that mammography does not add anything to physical breast examination. The well publicised controversy about results of randomised trials of mammographic screening[22,23] stems from two factors: the natural history of breast cancers is not straightforward and mammography is not a good tool, consequently, the benefit from screening is exceedingly modest. It takes about 1500 women to be screened for 10 years to save one life. Assuming the best possible benefit from mammography an elegant model has been recently created.[24] Data from this model suggest that more than half of invasive cancers detected by screening would never become clinically evident,[25] a fact that is never included in the 'information' sheet given to women invited to national screening programme. Mammography is not very sensitive (sensitivity = 65%).[26] So the interval cancer rates are substantial and these cancers typically have a poor prognosis. Wisdom of hindsight can pick an abnormality in a mammogram that was reported as normal and in a blame-and-shame culture can lead to very nasty accusations, expensive law-suits and general discontent. Mammography also has low specificity. The high rate of recall and false alarm mean that about one-third (half in the USA) of women in the UK would be recalled for further tests during their 15 years of screening programme. Women are not informed of these figures before entering the screening programme. So the psychological impact of these false alarms is considerable.[27] The phenomenon of post-event consolation means that these are never considered wrong — because if the alarm was false, there is relief and if the alarm is true then there is the undeniable feeling 'we caught it early'. But the women did not consent for all these stresses! The other fact that remains hidden, more by ignorance than by malice, is that the diagnosis of ductal carcinoma in situ (DCIS) was extremely rare before screening mammography. Less than half of DCIS treated with surgical biopsy progress to invasive carcinoma.[28] Left alone, even fewer may have progressed. Screening mammography does not seem to reduce the incidence of invasive carcinoma even over a 30-year period,[24,25] suggesting that the effect of screening mammography and subsequent treatment on natural history of the disease is small. The diagnosis by screening, of a breast cancer that could well have been inconsequential, can have several repercussions, such as the emotional stress, social implications and finally the financial issues such as difficulty in getting insurance, mortgage and continuing in a well established job. Since we cannot reliably distinguish good from bad screen detected cancers, treatment remains aggressive and sometimes paradoxically so. For example, the rate of mastectomy can be much higher for screen detected cancers than for symptomatic cancers.[29] So screening mammography does not even reduce morbidity of treatment. Furthermore, screening mammography incurs a high opportunity cost. The billions spent on screening could be better spent on finding new drugs to treat breast cancer that may save many more lives.[30,31] Arguably, the principle benefit of screening programmes world over has been to create better organisation of service delivery and increased awareness about breast cancer that may have led to increased investment into research.

So should screening for breast cancer be abandoned? That may be too extreme. Perhaps, we should re-examine our premises and tools. Even with the price of overdiagnosis by an order of 2, X-ray mammography cannot detect cancer until it has lived 90% in its natural history (i.e., 28 of 30 doublings).[32] This is long after the angiogenic switch that could start the metastatic process. Thus, even with a mechanistic theory, the effect of mammographic screening is unlikely to be large. Would newer more sensitive screening tools such as MRI deliver what X-ray mammography did not? In the UK MARIBS study[33] of young high-risk women (BRCA-1/BRCA-2 mutation carriers), MRI detected twice the number of cancers as X-ray mammography. Adding MRI to XRM increased the sensitivity from 40% to 94%. Logic would dictate that this could save lives, but human biology does not always follow logic. Secondly, even MRI needs to wait for increased vascularity to detect a cancer when it may already be too late because the tumour cells would have had access to circulation. The opportunity to test this may not have been lost. A worldwide (1:4, MRI:control) randomised study that includes women from countries at all levels of prosperity could be the only ethical way of rationing scare healthcare resources. That would answer the questions whether 'even earlier' detection by MRI can favourably alter the natural history of the breast cancer and, if so at what cost. Adapting MRI as the standard of care for high-risk women based on sensitivity results alone would be unethical especially if it competes for resources

with new biological treatments such as Trastuzumab, that have been proven to save lives.[34,35]

While women should be well informed about it, one cannot be too critical about overdiagnosis of cancer by screening and view it in the proper perspective. A simple way is to count the numbers needed to treat (NNT). For example, in large sentinel node biopsy studies, 26% of clinically negative axillae harbour involved nodes. If axillary clearance is acceptable treatment for this group, then the NNT would be 4. Similarly, if half of screen detected cancers would have progress to clinical cancers, then the NNT for treating a screen detected cancer (as opposed to screening process itself) is 2. This may be more than symptomatic disease, but one could argue that it is manifold smaller than the NNTs used for several other well accepted interventions such as systemic adjuvant therapy for cancer or even treatment of hypertension, which invariably run into 2 or 3 digit numbers.

The concept of screening is based on the logical axiom ''a stitch in time, saves nine''. It works well for machines, but living organisms have evolved to cope with errors. Many of these errors, therefore, do not result in disease. We have only just begun to unravel the subtleties of the human pathological process and it is easy to fall into the trap of treating every possible 'abnormality' that can be found. The problem is compounded by newer imaging techniques and molecular markers that can find seeds of disease that may never bloom in the person's lifetime. The cost of screening is not just financial, but sociological. Once the diagnosis of cancer loses all its associated clinical and social taboos, it would be easier to accept overdiagnosis. Before any test, that would 'logically' save lives, is accepted for general use it needs to go through the rigors of the randomised trials process that asks the right questions — does it really work? And at what cost? Clinicians and health policy makers should remember the differential effectiveness of an apparently universal intervention such as screening and the high non-apparent costs associated with it.

References

1. Wilson JM, Jungner YG. Principles and practice of mass screening for disease. *Bol Oficina Sanit Panam* October 1968;**65**(4):281—393.
2. Fleming C, Whitlock EP, Beil TL, Lederle FA. Screening for abdominal aortic aneurysm: a best-evidence systematic review for the U.S. preventive services task force. *Ann Intern Med* February 1, 2005;**142**(3):203—11.
3. Multicentre Aneurysm Screening Study Group. Multicentre aneurysm screening study (MASS): cost effectiveness analysis of screening for abdominal aortic aneurysms based on four year results from randomised controlled trial. *BMJ* November 16, 2002;**325**(7373):1135.
4. Earnshaw JJ, Shaw E, Whyman MR, Poskitt KR, Heather BP. Screening for abdominal aortic aneurysms in men. *BMJ* May 8, 2004;**328**(7448):1122—4.
5. Vaidya JS, Baum M. Screening and mortality from cervical cancer. Does screening really reduce mortality? *BMJ* September 4, 1999;**319**(7210):642.
6. Quinn M, Babb P, Jones J, Allen E. Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. *BMJ* April 3, 1999;**318**(7188):904.
7. Kronborg O, Jorgensen OD, Fenger C, Rasmussen M. Randomized study of biennial screening with a faecal occult blood test: results after nine screening rounds. *Scand J Gastroenterol* September 2004;**39**(9):846—51.
8. Sankaranarayanan R, Ramadas K, Thomas G, Muwonge R, Thara S, Mathew B, et al. Effect of screening on oral cancer mortality in Kerala, India: a cluster-randomised controlled trial. *Lancet* June 21, 2005;**365**(9475):1927—33.
9. Tsubono Y, Hisamichi S. A halt to neuroblastoma screening in Japan. *N Engl J Med* May 6, 2004;**350**(19):2010—1.
10. Schilling FH, Spix C, Berthold F, Erttmann R, Fehse N, Hero B, et al. Neuroblastoma screening at one year of age. *N Engl J Med* April 4, 2002;**346**(14):1047—53.
11. Sone S, Takashima S, Li F, Yang Z, Honda T, Maruyama Y, et al. Mass screening for lung cancer with mobile spiral computed tomography scanner. *Lancet* April 25, 1998; **351**(9111):1242—5.
12. Vaidya JS, Baum M. Low-dose spiral computed tomography for lung-cancer screening [letter]. *Lancet* 1998;**352**(9123): 236 [discussion 236—7].
13. Manser R, Dalton A, Carter R, Byrnes G, Elwood M, Campbell DA. Cost-effectiveness analysis of screening for lung cancer with low dose spiral CT (computed tomography) in the Australian setting. *Lung Cancer* May 2005;**48**(2): 171—85.
14. Bill-Axelson A, Holmberg L, Ruutu M, Haggman M, Andersson SO, Bratell S, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. *N Engl J Med* May 12, 2005;**352**(19):1977—84.
15. Holmberg L, Bill-Axelson A, Helgesen F, Salo JO, Folmerz P, Haggman M, et al. A randomized trial comparing radical prostatectomy with watchful waiting in early prostate cancer. *N Engl J Med* September 12, 2002;**347**(11):781—9.
16. Steineck G, Helgesen F, Adolfsson J, Dickman PW, Johansson JE, Norlen BJ, et al. Quality of life after radical prostatectomy or watchful waiting. *N Engl J Med* September 12, 2002;**347**(11):790—6.
17. Labrie F, Candas B, Cusan L, Gomez JL, Belanger A, Brousseau G, et al. Screening decreases prostate cancer mortality: 11-year follow-up of the 1988 Quebec prospective randomized controlled trial. *Prostate* May 15, 2004;**59**(3):311—8.
18. Baum M, Chaplain MA, Anderson AR, Douek M, Vaidya JS. Does breast cancer exist in a state of chaos? *Eur J Cancer* June 1999;**35**(6):886—91.
19. Nielsen M, Thomsen JL, Primdahl S, Dyreborg U, Andersen JA. Breast cancer and atypia among young and middle-aged women: a study of 110 medicolegal autopsies. *Br J Cancer* December 1987;**56**(6):814—9.
20. 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* September 20, 2000;**92**(18):1490—9.
21. 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. A randomized screening trial of mammography in women age 40 to 49 years. *Ann Intern Med* September 3, 2002;**137**(5 Part 1):305–12.
22. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet* October 20, 2001;**358**(9290):1340–2.
23. Gotzsche PC, Olsen O. Is screening for breast cancer with mammography justifiable? *Lancet* January 8, 2000; **355**(9198):129–34.
24. Barratt A, Howard K, Irwig L, Salkeld G, Houssami N. Model of outcomes of screening mammography: information to support informed choices. *BMJ* April 23, 2005;**330**(7497):936.
25. Vaidya JS. Does screening mammography cause cancers? *BMJ* May 12, 2005 <http://bmj.bmjjournals.com/cgi/eletters/330/7497/936#106627>.
26. Sarkeala T, Hakama M, Saarenmaa I, Hakulinen T, Forsman H, Anttila A. Episode sensitivity in association with process indicators in the Finnish breast cancer screening program. *Int J Cancer* Jul 7, 2005 [Epub ahead of print].
27. Barton MB, Morley DS, Moore S, Allen JD, Kleinman KP, Emmons KM, et al. Decreasing women's anxieties after abnormal mammograms: a controlled trial. *J Natl Cancer Inst* April 7, 2004;**96**(7):529–38.
28. Sanders ME, Schuyler PA, Dupont WD, Page DL. The natural history of low-grade ductal carcinoma in situ of the breast in women treated by biopsy only revealed over 30 years of long-term follow-up. *Cancer* May 9, 2005.
29. Olsen O, Gotzsche PC. Cochrane review on screening for breast cancer with mammography. *Lancet* October 20, 2001;**358**(9290):1340–2.
30. Baum M. Rethink on screening for breast cancer. *Lancet* September 13, 1997;**350**(9080):810–1.
31. Baum M. Breast cancer screening for younger women is not an efficient use of resources. *BMJ* June 21, 1997;**314**(7097): 1834.
32. Mittra I. Breast cancer screening by physical examination. In: Jatoi I, editor. *Breast cancer screening*. Austin, TX: Landes Biosciences; 1997. p. 97–100.
33. Leech MOaMc. Screening in women at genetic risk of breast cancer: results of the UK multicentre study of magnetic resonance imaging (MRI) and X-ray mammography (XRM). ASCO annual meeting 2005 abstracts, 520; 2005.
34. Rommond EH, Perez EA, Bryant J, et al. Advances in monocional antibody therapy for breast cancer. Combined analysis of NSABP-B31/NCCTG-N9831. Proceedings of the 41st Annual Meeting of the American Society of Clinical Oncology. Orlando, Fla. May 2005; oral presentation.
35. Piccart-Gebhart MJ. Advances in monoclonal antibody therapy for breast cancer: first results of the HERA trial. Proceedings of the 41st Annual Meeting of the American Society of Clinical Oncology. Orlando, Fla. May 2005; oral presentation.

Jayant S.Vaidya
University of Dundee,
Department of Surgery and Molecular Oncology,
Level 6, Ninewells Hospital and Medical School,
Dundee DD1 9SY, UK
Tel.: +44 (0) 1382 633 936.
E-mail address: j.s.vaidya @dundee.ac.uk