In practice

Assessing the efficacy of cancer screening

Gemma Jacklyn[a], Katy Bell[a,b] and Andrew Hayen[c,d]

[a] Wiser Healthcare, Sydney School of Public Health, University of Sydney, NSW, Australia
[b] Centre for Research in Evidence-Based Practice, Bond University, Gold Coast, Qld, Australia
[c] Australian Centre for Public and Population Health Research, University of Technology Sydney, NSW, Australia
[d] Corresponding author: andrew.hayen@uts.edu.au

Article history

Key points

- An evidence based approach to cancer screening is essential
- Both benefits and harms (including false positives, overdiagnosis and overtreatment) need to be estimated
- Evaluating cancer screening remains difficult because of important biases; randomised trials are the most effective way to control for lead-time, length-time, overdiagnosis and volunteer bias
- Communication about screening should always present benefits and harms using absolute risks and mortality

Abstract

Background: Population-based cancer screening has been established for several types of cancer in Australia and internationally. Screening may perform differently in practice from randomised controlled trials, which makes evaluating programs complex.

Materials and methods: We discuss how to assess the evidence of benefits and harms of cancer screening, including the main biases that can mislead clinicians and policy makers (such as volunteer, lead-time, length-time and overdiagnosis bias). We also discuss ways in which communication of risks can inform or mislead the community.

Results: The evaluation of cancer screening programs should involve balancing the benefits and harms. When considering the overall worth of an intervention and allocation of scarce health resources, decisions should focus on the net benefits and be informed by systematic reviews. Communication of screening outcomes can be misleading. Many messages highlight the benefits while downplaying the harms, and often use relative risks and 5-year survival to persuade people to screen rather than support informed choice.

Lessons learned: An evidence based approach is essential when evaluating and communicating the benefits and harms of cancer screening, to minimise misleading biases and the reliance on intuition.

Introduction

Cancer screening is the systematic search for cancer in people who have no signs or symptoms of the disease, to identify those who probably have cancer and those who do not. Individuals with a positive test result undergo further investigations, and those who are diagnosed with cancer receive treatment that aims to improve the length and quality of their life. Because screening involves testing people who are asymptomatic, there is the potential to turn otherwise healthy people into patients by sending them through a cascade of events that can be beneficial as well as harmful. Thus, there is a strong ethical imperative to ensure that a cancer screening program results in net benefit to the population and that we are using scarce resources efficiently and rationally.

The health benefits of cancer screening include minimising morbidity and mortality due to cancer, and reducing the risk of developing the disease.

However, there is also the potential for harm. The most frequent harm is a false-positive result – where an abnormality is detected but, with further investigation, no cancer is found. The cumulative risk of a false-positive result increases with each screening mammogram and can lead to unnecessary investigations that can cause physical pain and scarring, and negatively affect quality of life.[1]

The most serious harm is overdiagnosis – where an individual is diagnosed with cancer, and typically receives treatment, even though the cancer was not destined to cause symptoms or death. Overdiagnosis occurs because some screen-detected cancers grow very slowly, or are nonprogressive or even regressive, and competing mortality means that an individual may die from another cause before the cancer causes symptoms.

Potential for bias in cancer screening

Two critical forms of bias affect estimates of the effects of cancer screening: lead-time bias and length-time bias. Lead-time bias arises when we compare the survival – the time from diagnosis to death – of people screened for cancer to those not screened (Figure 1). Screening detects cancer earlier in the natural history of the disease and therefore moves the diagnosis to an earlier point in time. So even if cancer screening is not effective, and therefore does not reduce mortality, it may appear to increase survival because of lead-time bias.

Length-time bias can also arise in analyses of cancer screening (Figure 2). Cancer screening tests are typically able to detect slow-growing cancers with a longer preclinical phase (that is, a longer time before signs and symptoms develop), which tend to have a better prognosis. In contrast, more rapidly growing cancers with a shorter preclinical phase and poorer prognosis are more likely to present clinically than be detected by screening. An extreme form of length-time bias is called overdiagnosis bias, where survival rates improve because, as a result of attending screening, people are diagnosed with a cancer that would never have caused symptoms during their lifetime. Randomised controlled trials in which the endpoint is mortality rates (not survival) can largely overcome lead- and length-time biases.
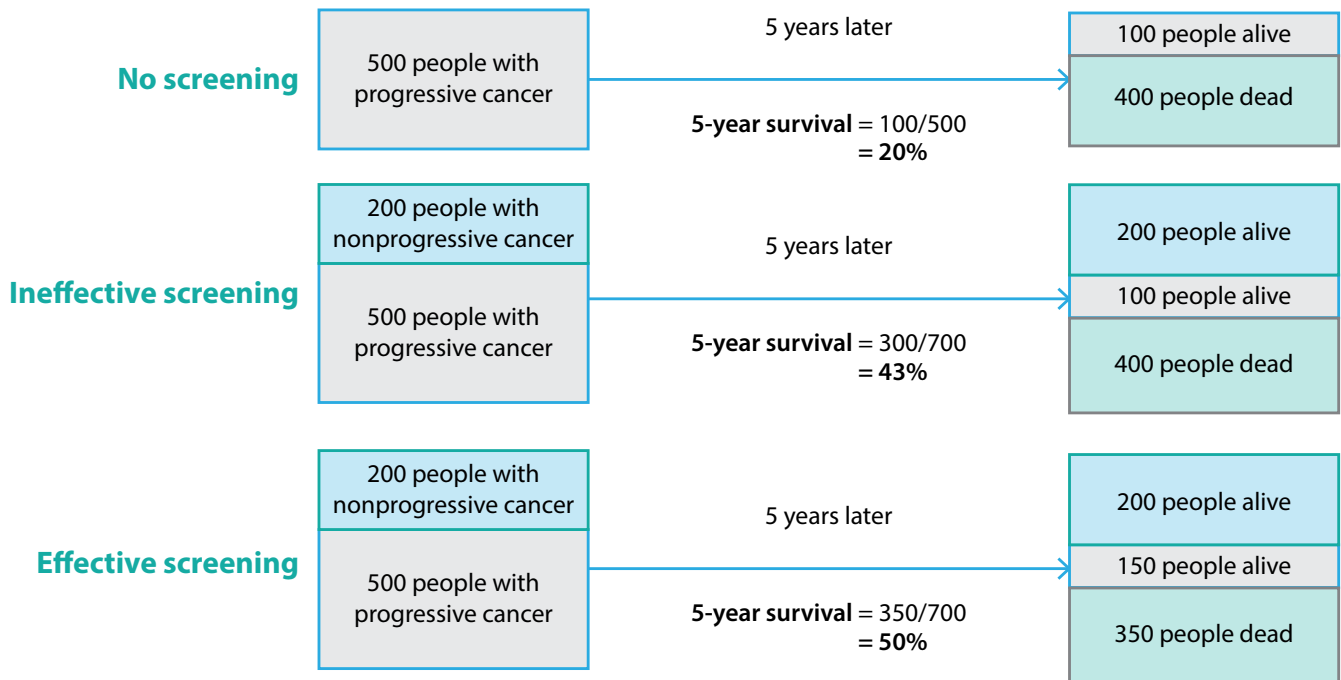
Communication of cancer screening

A systematic review has shown that the public overestimates the benefits and underestimates the harms of screening for breast, cervical and prostate cancer.[2] Advocates of screening may emphasise its advantages while minimising (or even ignoring) the downsides.[3] Efforts may go beyond persuasion and make people feel irresponsible and guilty to convince them to have cancer screening. A more balanced communication of potential benefits and harms facilitates informed choice, where individuals make their own decision about participating in screening.[4] There is arguably an ethical imperative to provide evidence about benefits and harms of screening to potential participants, and to policy makers who must choose between different population health strategies within budgetary constraints.[5,6]

Figure 1. Lead-time bias

2
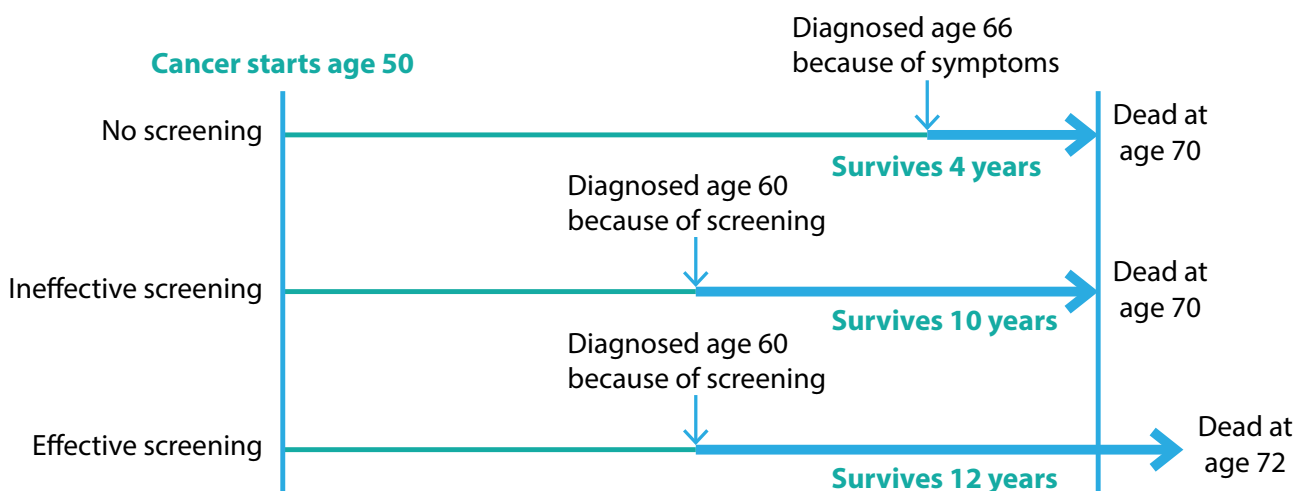
Figure 2. Length-time bias

When communicating the potential mortality benefit of cancer screening, the first step is to ensure that an appropriate outcome is used. Survival rates (typically reported as 5-year survival) should be avoided because they can be misleading. Five-year survival is defined as the proportion of people who are still alive 5 years after diagnosis of cancer. Although survival has an intuitive appeal, it cannot be used to establish the efficacy of screening because of lead- and length-time bias (see Figures 1 and 2). The mortality rate, which is the number of people who die from the cancer divided by the size of the population, is a better outcome to evaluate the effectiveness of cancer screening. Mortality does not consider the time from diagnosis to death and is unaffected by length-time bias.

Even when the mortality rate is used to communicate benefit, the way this is reported can influence health behaviours.[7] Health statistics are often reported as relative risks rather than absolute risks, which affect perception of the benefits and harms of screening.[8] For example, screening for prostate cancer using the prostate-specific antigen (PSA) test gives a 20% relative reduction in prostate cancer mortality among men aged 55 to 69 years according to one estimate.[9] If we consider the absolute risk reduction, however, the benefit is less impressive. For every 1000 men aged 55 to 69 years who are screened every 2 to 7 years across a 10-year period, 0 to 1 men avoid dying from prostate cancer compared with unscreened men. This represents a reduction from 5 deaths among every 1000 unscreened men to 4 to 5 deaths among 1000 screened men.[9]

Evaluating whether to implement a cancer screening program

Evaluating a cancer screening program involves estimating benefits and harms to determine net benefits, and then considering whether the magnitude of net benefits justifies the resources required to run a program.[10] An evaluation should be conducted by an independent organisation that considers both high-quality scientific evidence and other factors such as available resources and population preferences. This requires data on patient-important health outcomes, cost-effectiveness and public engagement to elicit values and preferences, using deliberative methods.[11] The best method for estimating the benefit and harms of cancer screening is a randomised controlled trial, but this may lack applicability to clinical practice.

Cancer screening trials require a large sample size – often hundreds of thousands of participants – to ensure control of potential confounding factors and detection of a difference between groups. This is resource intensive and makes it difficult to achieve adequate recruitment. Participants may prefer a particular group than the one to which they are randomised, which can influence screening attendance.[12] Resulting loss of fidelity causes bias towards the null, reducing statistical power.[13] Long-term follow-up over many years is necessary to evaluate both the benefit of mortality reduction and harms due to overdiagnosis. These prerequisites of cancer screening trials – long duration, large sample size and adequate adherence to the trial protocol – can make them impractical, and they may be particularly challenging to undertake once a program is established.

Any decision to implement a cancer screening program should be based on evidence of net benefit to the population, demonstrated by a reduction in all-cause mortality – that is, evidence that people are, on balance, likely to live longer if they participate in screening. A reduction in overall mortality may not be found because any decrease in cancer-specific mortality is offset or even overtaken by deaths due to downstream effects of the screening test, diagnosis or treatment.[14] Increased mortality from other causes might occur as a result of:

- The invasive test needed to confirm diagnosis after a positive screening test (e.g. prostate biopsies)[15,16]
- Psychological consequences of the disease label (e.g. increased rates of myocardial infarction and of suicide after prostate cancer diagnosis)[17]
- Consequences of treatment of overdiagnosed cancers (e.g. deaths due to surgical complications and radiation effects after treatment for breast cancer).[18]

Few randomised controlled trials conducted to date have had sufficient power to show an effect on all-cause mortality. Solutions to this include pooling trial data by meta-analysis or analysing epidemiological trends. Where there is uncertainty, new randomised controlled trials that are powered to detect a difference in all-cause mortality may be needed. Although such trials need to be very large, they are justified given the expense of population screening programs when we are uncertain if there is true benefit to society. Trial costs may be substantially reduced if they are conducted within large, national observational registries, which may even make them comparable to the cost of current screening trials.[19]

Evaluating a cancer screening program after implementation

Although randomised trial data are often used to inform decisions about whether a screening program should be implemented, they are rarely used for evaluation of benefits and harms of the program once it is established. Uncertainty may remain about specific components of the strategy, and the screening technology itself often changes over time – for example, 3D mammography for breast cancer screening[20] or liquid cytology for cervical cancer screening.[21] Advances in technology are often incorporated into clinical practice without adequate evaluation of their effects on important health outcomes such as cancer-specific mortality or overdiagnosis.[20,22]

There may also be changes in cancer prevention and treatment after the implementation of a screening program. For example, improvements in breast cancer treatment[23] have meant that the mortality benefit of current mammography screening is likely to be smaller than that first observed in the original screening trials. Improvements in prevention can also decrease the incidence of cancer. For example, implementation of human papillomavirus vaccination programs has brought a decline in high-grade cervical abnormalities and reduced estimated mortality benefits from cervical cancer screening.[24] Given the improved outlook, these cancer screening programs may not be as necessary as they once were.

Ongoing evaluation of screening programs is possible using pragmatic randomised trials within the program, with the eligible population randomised to different screening strategies.[25] The comparison may be between two screening interventions[26,27] or screening to no screening where there is proposed expansion to the program.[28] Evaluating the benefits and harms of screening through randomised comparisons could theoretically be widely adopted within new and existing screening programs, if there is the political will – the previous examples show that communities readily accept the randomisation.[26-28] Observational studies are more commonly used to evaluate existing cancer screening programs. Although they may be more applicable than randomised controlled trials, they are more prone to biases (which usually favour screening). Nevertheless, observational studies may be useful for monitoring benefits and harms over time.[29] Regardless of the study design, the critical requirement for screening evaluation is that the evidence enables us, with minimal bias and reasonable certainty, to quantify the benefits and harms.

Conclusions

An evidence based approach to cancer screening is essential to maximise benefits (improved length and quality of life) while minimising the harms to individuals (false positives, overdiagnosis and overtreatment) and opportunity costs to society. Evaluating cancer screening remains difficult due to important biases, so we must continue to implement randomised controlled trials to generate the best evidence on the magnitude of benefits and harms with reasonable certainty. Those responsible for communicating cancer screening can do better by always providing absolute risks and ensuring transparent reporting of benefits and harms without using 5-year survival rates or relative risks.

Acknowledgements

Competing interests

None declared

Author contributions

All authors contributed to drafting the manuscript.

References

1. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. Ann Fam Med. 2013;11(2):106–15.

2. Hoffmann TC, Del Mar C. Patients' expectations of the benefits and harms of treatments, screening, and tests: a systematic review. JAMA Intern Med. 2015;175(2):274–86.

3. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaigns – getting past uninformative persuasion. N Engl J Med. 2012;367(18):1677–9.

4. Hersch J, Barratt A, Jansen J, Irwig L, McGeechan K, Jacklyn G, et al. Use of a decision aid including information on overdetection to support informed choice about breast cancer screening: a randomised controlled trial. Lancet. 2015;385(9978):1642–52.

5. Carter SM, Degeling C, Doust J, Barratt A. A definition and ethical evaluation of overdiagnosis. J Med Ethics. 2016;42(11):705–14.

6. Salmi LR, Coureau G, Bailhache M, Mathoulin-Pélissier S. To screen or not to screen: reconciling individual and population perspectives on screening. Mayo Clin Proc. 2016;91(11):1594–605.

7. Gigerenzer G. Making sense of health statistics. Bull World Health Organ. 2009;87(8):567.

8. Trevena LJ, Zikmund-Fisher BJ, Edwards A, Gaissmaier W, Galesic M, Han PK, et al. Presenting quantitative information about decision outcomes: a risk communication primer for patient decision aid developers. BMC Med Inform Decis Mak. 2013;13(Suppl 2):S7.

9. Moyer VA; U.S. Preventive Services Task Force. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;157(2):120–34.

10. Harris R, Sawaya GF, Moyer VA, Calonge N. Reconsidering the criteria for evaluating proposed screening programs: reflections from 4 current and former members of the U.S. Preventive Services Task Force. Epidemiol Rev. 2011;33:20–35.

11. Degeling C, Carter SM, Rychetnik L. Which public and why deliberate?–A scoping review of public deliberation in public health and health policy research. Soc Sci Med. 2015;131:114–21.

12. Ilic D, Neuberger MM, Djulbegovic M, Dahm P. Screening for prostate cancer. Cochrane Database Syst Rev. 2013(1):CD004720.

13. Sommer A, Zeger SL. On estimating efficacy from clinical trials. Stat Med. 1991;10(1):45–52.

14. Prasad V, Lenzer J, Newman DH. Why cancer screening has never been shown to "save lives"—and what we can do about it. BMJ. 2016;352:h6080.

15. Loeb S, Carter HB, Berndt SI, Ricker W, Schaeffer EM. Complications after prostate biopsy: data from SEER-Medicare. J Urol. 2011;186(5):1830–4.

16. Gallina A, Suardi N, Montorsi F, Capitanio U, Jeldres C, Saad F, et al. Mortality at 120 days after prostatic biopsy: a population-based study of 22,175 men. Int J Cancer. 2008;123(3):647–52.

17. Fang F, Keating NL, Mucci LA, Adami HO, Stampfer MJ, Valdimarsdóttir U, Fall K. Immediate risk of suicide and cardiovascular death after a prostate cancer diagnosis: cohort study in the United States. J Natl Cancer Inst. 2010;102(5):307–14.

18. Marmot MG, Altman DG, Cameron DA, Dewar JA, Thompson SG, Wilcox M. The benefits and harms of breast cancer screening: an independent review. Br J Cancer. 2013;108(11):2205–40.

19. Lauer MS, D'Agostino Sr RB. The randomized registry trial–the next disruptive technology in clinical research? N Engl J Med. 2013;369(17):1579–81.

20. Bernardi D, Macaskill P, Pellegrini M, Valentini M, Fantò C, Ostillio L, et al. Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study. Lancet Oncol. 2016;17(8):1105–13.

21. Davey E, Barratt A, Irwig L, Chan SF, Macaskill P, Mannes P, Saville AM. Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional cervical cytology: a systematic review. Lancet. 2006;367(9505):122–32.

22. Pisano ED, Gatsonis C, Hendrick E, Yaffe M, Baum JK, Acharyya S, et al. Diagnostic performance of digital versus film mammography for breast-cancer screening. N Engl J Med. 2005;353(17):1773–83.

23. Burton RC, Bell RJ, Thiagarajah G, Stevenson C. Adjuvant therapy, not mammographic screening, accounts for most of the observed breast cancer specific mortality reductions in Australian women since the national screening program began in 1991. Breast Cancer Res Treat. 2012;131(3):949–55.

24. Brotherton JML, Fridman M, May CL, Chappell G, Saville AM, Gertig DM. Early effect of the HPV vaccination programme on cervical abnormalities in Victoria, Australia: an ecological study. Lancet. 2011;377(9783):2085–92.

25. Bell KJ, Bossuyt P, Glasziou P, Irwig L. Assessment of changes to screening programmes: why randomisation is important. BMJ. 2015;350:h1566.

26. Bowel Cancer Screening Pilot Monitoring and Evaluation Steering Committee. The Australian bowel cancer screening pilot program and beyond: final evaluation report. Screening Monograph No.6/2005. Canberra: Commonwealth of Australia; 2005 [cited 2017 May 19]. Available from: www.cancerscreening.gov.au/internet/screening/publishing.nsf/Content/9C0493AFEB3FD33CCA257D720005C9F2/$File/final-eval-1.pdf

27. van Roon AH, Goede SL, van Ballegooijen M, van Vuuren AJ, Looman CW, Biermann K, et al. Random comparison of repeated faecal immunochemical testing at different intervals for population-based colorectal cancer screening. Gut. 2013;62(3):409–15.

28. Moser K, Sellars S, Wheaton M, Cooke J, Duncan A, Maxwell A, et al. Extending the age range for breast screening in England: pilot study to assess the feasibility and acceptability of randomization. J Med Screen. 2011;18(2):96–102.

29. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. BMJ. 2015;350:g7773.