

What is overdiagnosis and why should we take it seriously in cancer screening?

Stacy M Carter^{a,c} and Alexandra Barratt^b

^a Centre for Values, Ethics and the Law in Medicine, University of Sydney, NSW, Australia

^b Sydney School of Public Health, University of Sydney, NSW, Australia

^c Corresponding author: stacy.carter@sydney.edu.au

Article history

Publication date: July 2017

Carter SM, Barratt A. What is overdiagnosis and why should we take it seriously in cancer screening? *Public Health Res Pract.* 2017;27(3):e2731722.

<https://doi.org/10.17061/phrp2731722>

Key points

- An overdiagnosed cancer is correctly diagnosed but would not have produced symptoms or premature death if left undetected; treating these cancers results in harm instead of benefit
- Cancer screening is one driver of cancer overdiagnosis
- Overdiagnosis and false positives are not equivalent
- Much cancer overdiagnosis occurs because cancer is heterogeneous, and screening is more likely to detect cancers that are least likely to cause cancer death
- We should be concerned about overdiagnosis because it harms healthy people, creates opportunity costs by misdirecting healthcare resources, and potentially undermines the ultimate goal of screening, which is to reduce suffering and premature death

Abstract

Overdiagnosis occurs in a population when conditions are diagnosed correctly but the diagnosis produces an unfavourable balance between benefits and harms. In cancer screening, overdiagnosed cancers are those that did not need to be found because they would not have produced symptoms or led to premature death. These overdiagnosed cancers can be distinguished from false positives, which occur when an initial screening test suggests that a person is at high risk but follow-up testing shows them to be at normal risk. The cancers most likely to be overdiagnosed through screening are those of the prostate, thyroid, breast and lung. Overdiagnosis in cancer screening arises largely from the paradoxical problem that screening is most likely to find the slow-growing or dormant cancers that are least likely to harm us, and less likely to find the aggressive, fast-growing cancers that cause cancer mortality. This central paradox has become clearer over recent decades. The more overdiagnosis is produced by a screening program, the less likely the program is to serve its ultimate goal of reducing illness and premature death from cancer. Thus, it is vital that health professionals and researchers continue an open, scientific inquiry into the extent and consequences of overdiagnosis, and devise appropriate responses to it.

Introduction

A cancer screening program tests a large population of healthy people in a defined age group, ideally using a simple, affordable test. This test, much like a sieve, 'catches' those people who are apparently at higher cancer risk and directs them to further testing. If diagnostic testing confirms the presence of cancer precursors or cancer, preventive or curative treatment is provided. This should mean easier, more effective treatment and fewer people progressing to late-stage, advanced cancer. The ultimate goal is to reduce the suffering and death caused by that cancer in that population without introducing any significant additional harms. These are good goals, and some screening programs deliver on them (e.g. cervical screening programs¹).

This article explains the problem of overdiagnosis in cancer screening. Because overdiagnosis produces harm, a discussion of overdiagnosis may be perceived to be a general condemnation of cancer screening. We do not intend

this, and point to other papers in this issue that focus on some of the benefits screening may offer.

Overdiagnosis is an important concern because it potentially breaks the all-important link between screening programs and their ultimate goal. The more overdiagnosis a screening program produces, the more screening is likely to increase rather than decrease suffering and even death in populations. Most cancer authorities now acknowledge that overdiagnosis is a consideration in screening, including in their communication to consumers (e.g. Cancer Research UK², US Preventive Services Task Force³). However, significant expert disagreement exists over the extent of overdiagnosis¹, and screening programs with different characteristics are likely to produce overdiagnosis to different degrees.⁴ In this article, we consider what overdiagnosis is, why it occurs, what forms of screening it is most relevant to, and how we should respond to it.

Overdiagnosis is not a false positive

A common misunderstanding is to conflate overdiagnosis and false positives in screening. A false positive occurs when the 'sieve' of screening 'catches' a person who is at normal risk of cancer and incorrectly suggests that they may be at high risk. After a sometimes anxious wait for further testing⁵, the person is shown to be at normal risk. The extent of false positives (and false negatives) in a screening program is partly determined by test characteristics (a good test produces fewer false results) and partly by agreed standards within the program (e.g. the agreed cut-off value for a biomarker).

Defining overdiagnosis

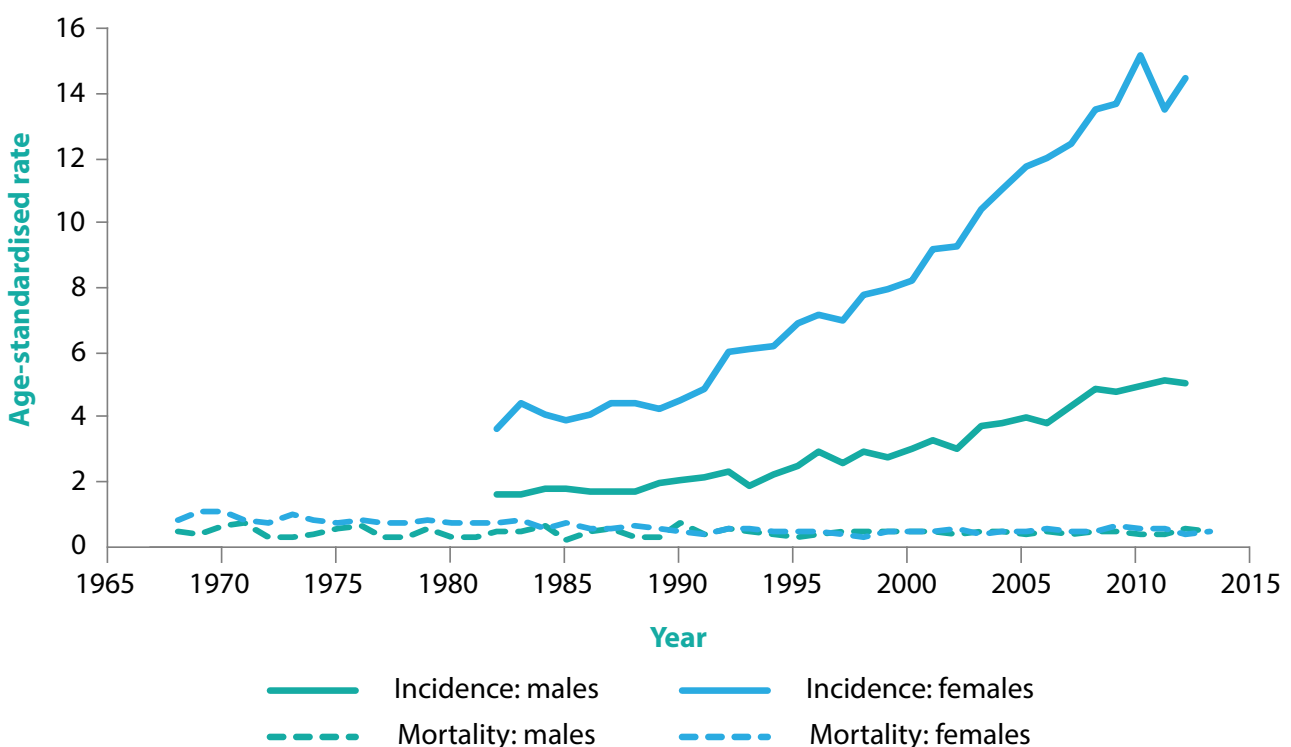
A false positive occurs when a person is incorrectly told that they may have cancer. Cancer overdiagnosis, in contrast, occurs when cancers are correctly diagnosed but those cancers would not have produced symptoms or been identified clinically.^{6,7} It is very difficult to determine whether a particular individual has been overdiagnosed⁸, particularly in cancer screening, because once a person has been diagnosed with cancer and treated, no-one can know what would have happened without that treatment.

Cancer overdiagnosis can therefore only be measured or statistically estimated in populations.⁸ Typically, a screening program that produces overdiagnosis in a population will greatly increase the incidence of early-stage cancer or pre-cancer, without reducing the incidence of late-stage cancer or mortality from that cancer. This is illustrated in Figure 1, using the example of thyroid cancer.^{7,9,10}

The paradox of screening for cancer(s)

An important dual paradox drives the overdiagnosis problem, one that has become more evident over time. First, cancer is complex and heterogeneous.^{12,13} Second, screening programs are more likely to detect slow-growing, less aggressive cancers and less likely to detect fast-growing, more aggressive cancers. This is simply a function of time. Screening occurs at regular intervals (e.g. 2, 3 or

Figure 1. Age-standardised thyroid cancer incidence and mortality rates per 100 000 males/females in Australia, 1968 to 2013



Source: Australian Institute of Health and Welfare¹¹

5 years) calculated to provide the best cost:benefit ratio and least harm. A slow-growing cancer is asymptotically present in the body for much longer, so is more likely to be present at a screening point. In contrast, a fast-growing, more aggressive cancer produces symptoms, so is likely to prompt the person to see a doctor and be clinically diagnosed in the gap between scheduled screenings.¹⁴ These ‘interval cancers’ are not a sign that the screening program has failed – they simply demonstrate that a small proportion of cancer is extremely aggressive (more so than any reasonable screening schedule could catch). This relationship is shown in Figure 2.

Slow-growing, less aggressive cancers are an important source of overdiagnosis in screening programs, as they may not progress, or may even regress¹⁵, and so do not need to be found. To quote Otis Brawley, the Chief Medical Officer of the American Cancer Society, they are “tumours that appear cancerous under the microscope but are behaviourally of no clinical threat”.¹³ These cancers bias the perceived outcome of screening in a potentially confusing way.¹⁶ It is common to see claims such as the following: “95% of women whose breast cancer is detected by screening are still alive after 5 years, versus only 25% of women whose cancer is diagnosed clinically”. This may be so, but it is at least partly a function of cancer heterogeneity, rather than a benefit of screening. Women with slow-growing (screen-detected) cancers are likely to survive for 5 years with or without screening. In contrast, women with aggressive (clinical, symptomatic) cancers are, sadly, less likely to survive 5 years because the

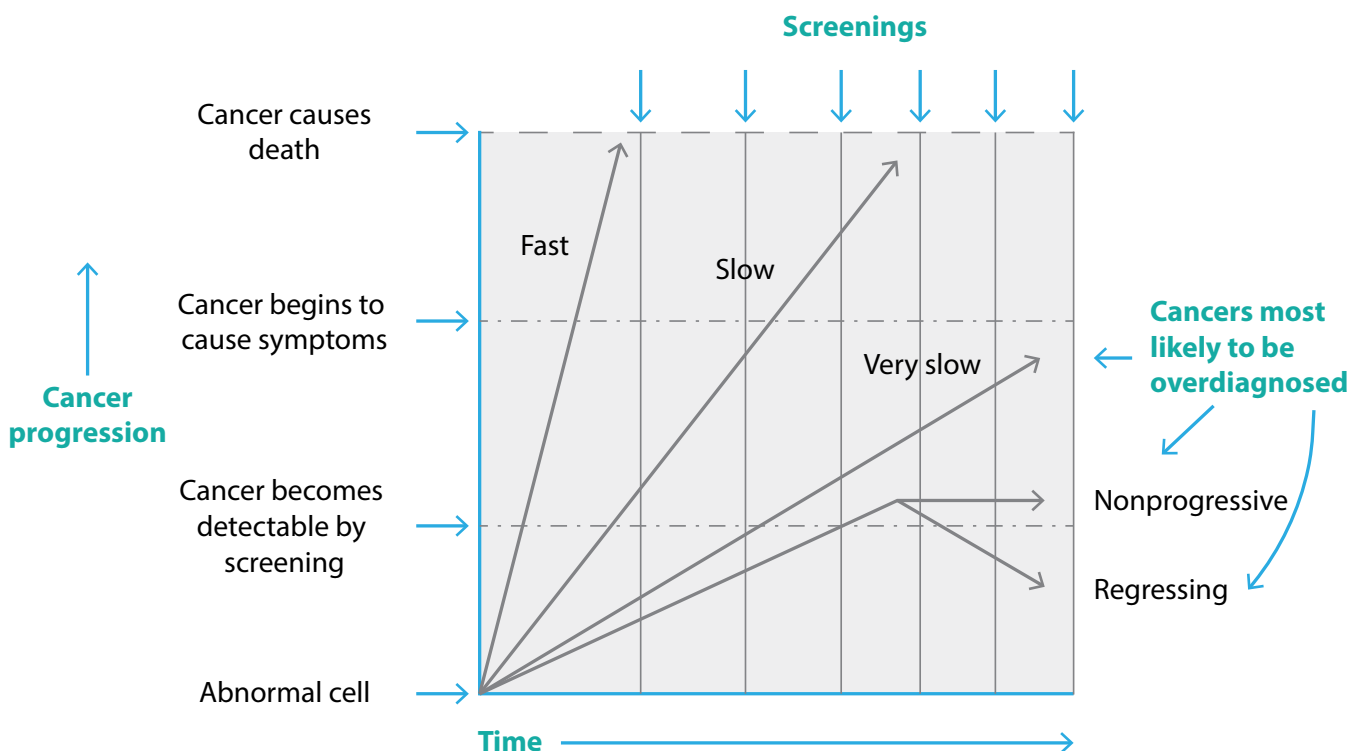
cancer is aggressive (and thus less likely to be detected by screening). Five-year survival will therefore always be higher in screen-detected cohorts than in clinically detected cohorts. This does not, however, prove that screening reduces suffering or premature cancer death.

A valid measure of the benefit of screening is the difference in absolute mortality rates between screened and unscreened groups over a particular time. Differences in absolute mortality tend to suggest a more delicate balance of benefits and harms than many might expect.¹⁷ Nonetheless, the balance of benefits and harms in any type of screening is often contested¹, and will vary depending on the disease, program design, protocol design, technology and quality assurance.⁴

Some cancers are more likely to be overdiagnosed than others.^{7,13} In particular, screening is more likely to overdiagnose cancer if it detects early-stage cancer and triggers cancer treatments (as opposed to screening that detects – and prompts removal of – changes that can later become cancer, e.g. cervical and colorectal screening). To quote Esserman et al.:¹⁸

Physicians, patients, and the general public must recognize that overdiagnosis is common and occurs more frequently with cancer screening. Overdiagnosis, or identification of indolent cancer, is common in breast, lung, prostate, and thyroid cancer. Whenever screening is used, the fraction of tumours in this category increases. By acknowledging this consequence of screening, approaches that mitigate the problem can be tested.

Figure 2. Cancer heterogeneity and multiple screenings drive overdiagnosis



Source: Adapted from Welch and Black.⁷

Consistent with this, those cancers most likely to be overdiagnosed are prostate cancer, from prostate-specific antigen testing of asymptomatic men in primary care; breast cancer, from organised mammography screening; thyroid cancer, from ultrasound of the thyroid and/or adjoining structures; and lung cancer, from screening smokers using computed tomography.^{7,10,13,19-21}

Discussing overdiagnosis is good science, not an accusation of malpractice

It is the nature of science to change: a scientific attitude entails scepticism and openness to questioning. Overdiagnosis in cancer screening has become topical because, within the science of cancer, it has become increasingly clear that some cancers do not need to be found, and that the benefit/harm balance of screening is less favourable than originally hoped. Unfortunately, at present, overdiagnosis debates sometimes fail to take a scientific approach, instead becoming polarised into perceived 'pro-screening' or 'anti-screening' camps, where defensiveness collides with zealotry.¹³

In future, we believe it is critical for healthcare practitioners and researchers to hold open, rigorous scientific conversations about overdiagnosis, staying focused on the central fact that overdiagnosis is correct diagnosis. The people who may cause overdiagnosis are rarely maleficent or incompetent: they are, in large part, professionals doing their job to a high standard and in the way their profession has agreed it should be done, with a desire to prevent suffering and premature death from cancer. The problem is that, within the healthcare system, standards for practice have been set at a point that may produce more harm than good. This harms individuals while appearing to help them.¹³ In addition, any amount of overdiagnosis will increase the costs and decrease the cost-effectiveness of screening programs, thus leading to opportunity costs.²²⁻²⁴ System standards need to be constantly and collaboratively monitored to ensure that screening produces more good than harm.^{24,25} More research is needed to test whether treating the lowest risk screen-detected cancers differently will reduce harm (e.g. Francis et al.²⁵, Lane et al.²⁶).

Citizens expect their healthcare systems to help, not harm. Health professionals make their career choices because they are motivated to relieve suffering, not cause it. If we keep hold of these values, it should be possible for us to continue to have a constructive and productive conversation about overdiagnosis and how to minimise it in cancer screening programs.

Acknowledgements

SC and AB received funding to support this work from the National Health and Medical Research Council under grant number 1104136.

Competing interests

SC received travel funding related to this work from the Preventing Overdiagnosis Conference 2017, hosted by the Quebec Medical Association. AB received travel funding related to this work from the Australian Government and the Cochrane Collaboration.

Author contributions

SC led the writing. SC and AB contributed equally to the conceptualisation and analysis.

References

1. Carter SM, Williams J, Parker L, Pickles K, Jacklyn G, Rychetnik L, et al. Screening for cervical, prostate, and breast cancer. *Am J Prev Med.* 2015;49(2):274–85.
2. Cancer Research UK. Let's beat cancer sooner. London: Cancer Research UK; 2013. Understanding cancer screening; 2015 Apr 2 [cited 2017 Jun 29]; [about 5 screens]. Available from: www.cancerresearchuk.org/about-cancer/screening/understanding-cancer-screening
3. U.S. Preventive Services Task Force. Rockville, MD: USPST; 2017. Final recommendation statement breast cancer: screening; 2016 Jan [cited 2017 Jun 29]; [about 25 screens]. Available from: www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/breast-cancer-screening1
4. Etzioni R, Gulati R, Mallinger L, Mandelblatt J. Influence of study features and methods on overdiagnosis estimates in breast and prostate cancer screening. *Ann Intern Med.* 2013;158(11):831–8.
5. Brodersen J, Siersma VD. Long-term psychosocial consequences of false-positive screening mammography. *Ann Fam Med.* 2013;11(2):106–15.
6. Marcus PM, Prorok PC, Miller AB, DeVoto EJ, Kramer BS. Conceptualizing overdiagnosis in cancer screening. *J Natl Cancer Inst.* 2015;107(4):djv014.
7. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst.* 2010;102(9):605–13.
8. Carter SM, Degeling C, Doust J, Barratt A. A definition and ethical evaluation of overdiagnosis. *J Med Ethics.* 2016;42(11):705–14.
9. Jørgensen K, Gøtzsche PC, Kalager M, Zahl P. Breast cancer screening in Denmark: a cohort study of tumor size and overdiagnosis. *Ann Intern Med.* 2017;166(5):313–23.

10. Ahn HS, Kim HJ, Welch HG. Korea's thyroid-cancer "epidemic" – screening and overdiagnosis. *N Engl J Med*. 2014;371(19):1765–7.
11. Australian cancer incidence and mortality (ACIM) books: thyroid cancer. Canberra: AIHW; 2016 [cited 2017 Jun 29]. Available from: www.aihw.gov.au/acim-books
12. Meacham CE, Morrison SJ. Tumour heterogeneity and cancer cell plasticity. *Nature*. 2013;501(7467):328–37.
13. Brawley OW. Accepting the existence of breast cancer overdiagnosis. *Ann Intern Med*. 2017;166(5):364–5.
14. Barratt A, Irwig L, Glasziou P, Cumming RG, Raffle A, Hicks N, et al. Users' guides to the medical literature - XVII. How to use guidelines and recommendations about screening. *JAMA*. 1999;281(21):2029–34.
15. Zahl P, Mæhlen J, Welch H. The natural history of invasive breast cancers detected by screening mammography. *Arch Intern Med*. 2008;168(21):2311–6.
16. Bell K, Barratt A, Hayen A. When talking about cancer screening, survival rates mislead. *The Conversation*. 2015 Jan 28. Available from: theconversation.com/when-talking-about-cancer-screening-survival-rates-mislead-30395
17. Biller-Andorno N, Jüni P. Abolishing mammography screening programs? A view from the Swiss Medical Board. *N Engl J Med*. 2014;370(21):1965–7.
18. Esserman LJ, Thompson IM Jr, Reid B. Overdiagnosis and overtreatment in cancer: an opportunity for improvement. *JAMA*. 2013;310(8):797–8.
19. Davies L, Welch H. Current thyroid cancer trends in the United States. *JAMA Otolaryngol Head Neck Surg*. 2014;140(4):317–22.
20. Patz EF, Pinsky P, Gatsonis C, Sicks JD, Kramer BS, Tammemägi MC, et al. Overdiagnosis in low-dose computed tomography screening for lung cancer. *JAMA Intern Med*. 2014;174(2):269–74.
21. Veronesi G, Maisonneuve P, Bellomi M, Rampinelli C, Durlì I, Bertolotti R, et al. Estimating overdiagnosis in low-dose computed tomography screening for lung cancer: a cohort study. *Ann Intern Med*. 2012;157(11):776–84.
22. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ*. 2012;344:e3502.
23. Ong MS, Mandl KD. National expenditure for false-positive mammograms and breast cancer overdiagnoses estimated at \$4 billion a year. *Health Aff (Millwood)*. 2015;34(4):576–83.
24. Pharoah PDP, Sewell B, Fitzsimmons D, Bennett HS, Pashayan N. Cost effectiveness of the NHS breast screening programme: life table model. *BMJ*. 2013;346:f2618.
25. Francis A, Thomas J, Fallowfield L, Wallis M, Bartlett JMS, Brookes C, et al. Addressing overtreatment of screen detected DCIS; the LORIS trial. *Eur J Cancer*. 2015;51(16):2296–303.
26. Lane JA, Donovan JL, Davis M, Walsh E, Dedman D, Down L, et al. Active monitoring, radical prostatectomy, or radiotherapy for localised prostate cancer: study design and diagnostic and baseline results of the ProtecT randomised phase 3 trial. *Lancet Oncol*. 2014;15(10):1109–18.

