

Better decision making in preventive health screening

Balancing benefits and harms

Neil R. Bell MD SM CCFP FCFP Roland Grad MD CM MSc CCFP FCFP James A. Dickinson MBBS CCFP PhD FRACGP
Harminder Singh MD MPH FRCPC Ainsley Elizabeth Moore MD MSc CCFP Danielle Kasperavicius MPH Kaylyn L. Kretschmer MSc

Decision making around preventive health screening by family physicians, other health care practitioners, and patients has become increasingly complex and controversial in recent years. Preventive health screening has long been advocated as one of the most important health care strategies to facilitate early diagnosis and treatment, improve quality of life, and prevent premature death. For many years, professional organizations, patient advocacy groups, and clinicians used a combination of public policy, persuasive advertising, and direct clinical messages to patients as methods to increase uptake of screening in specific patient populations and with individual patients.¹ Canadians have been consistently told that the more proactive you are about being screened and the more frequently that you

are screened, the healthier and more protected you will be. This belief has become deeply entrenched in society and has been reinforced by both empirical evidence and anecdotes. For example, the Papanicolaou test, introduced more than 50 years ago, has been highly effective in reducing the morbidity and mortality that was associated with cervical cancer.² There is widespread belief among men and women, supported by promotion from various organizations, in the benefits of screening for prostate cancer with the prostate-specific antigen (PSA) test and for breast cancer with mammography.

In spite of the perceived benefits of preventive screening, there has also been growing awareness and concern about the potential for preventive screening to cause harm to some patients.^{3,4} Examples of these conditions include screening for breast cancer with mammography,⁵⁻⁷ screening for prostate cancer with the PSA test,⁸ and screening for gestational diabetes.^{4,9}

Decisions around screening have been made more difficult owing to the proliferation of preventive screening guidelines by advocacy groups, professional organizations, and government agencies, often with conflicting recommendations on the same topic. Physicians and patients have their own clinical and personal experiences related to cancer and other health conditions that shape their beliefs about the value of preventive health screening for different conditions. Patients and physicians often overestimate the benefits and underestimate the harms associated with preventive health screening.^{10,11} Recommendations around screening strategies have often been accompanied by highly charged emotional debates related to beliefs about the potential benefits of a screening strategy.^{1,5,12}

Factors to consider

How can practitioners better inform preventive health screening decisions? The goal of screening is to improve health outcomes that matter to the patient, not simply to discover a disease state. Understanding the potential benefits and harms associated with a screening strategy is central to informed decision making by physicians and patients. This decision making requires understanding the key factors that influence the balance between benefits and harms, the quality of the evidence that supports the screening maneuver, and the common measures of

KEY POINTS

- Decision making about preventive health screening requires consideration of both the benefits and the harms.
- In screening for cancer, appropriate outcome measures for benefits include overall and disease-specific mortality; inappropriate measures include incidence (new cases) and 5- or 10-year survival.
- In order to adequately inform their patients, physicians need a detailed understanding of the magnitude of the benefits and harms of preventive screening maneuvers and testing.
- Overdiagnosis creates the greatest potential for harms owing to the adverse effects of further investigation and treatment of patients who will not benefit from treatment.
- In circumstances where there is a close trade-off between the benefits and harms of screening, shared decision making between physicians and patients should be undertaken. Shared decision making should include understanding patient values and preferences and the potential harms and benefits of different preventive health screening options.



This article is eligible for Mainpro+ certified Self-Learning credits. To earn credits, go to www.cfp.ca and click on the Mainpro+ link.

Cet article se trouve aussi en français à la page 525.

magnitude used to express the size of the benefits and harms. High-quality guidelines should provide information on both the benefits and harms associated with preventive health screening strategies and knowledge translation tools to support shared decision making with patients.

How do we determine the benefits of preventive screening?

The highest-quality evidence for screening interventions comes from randomized clinical trials that compare the screening intervention with usual care or no screening. In circumstances when evidence from randomized clinical trials is not available, lower-quality evidence from observational studies might also be used to inform the development of screening recommendations. Although many methods of evaluating the quality of evidence have been developed since the 1970s, the Canadian Task Force on Preventive Health Care, along with many other national and international guideline-development organizations, has adopted GRADE (Grading of Recommendations Assessment, Development and Evaluation) to evaluate the quality of evidence.^{13,14} The GRADE system ranks the continuum of the quality of evidence on a 4-point scale that ranges from very low quality to high quality.^{13,14}

The most important health outcome measures for cancer screening are disease-specific mortality and overall mortality. In screening for other health conditions, such as cognitive impairment or developmental delay, outcome measures such as quality of life, improved cognitive function, or academic performance are appropriate outcome measures.

Preferred measures of effect size or magnitude include both natural frequencies, which present the outcomes in a population of individuals undergoing screening compared with an equivalent population that is not screened, and absolute risk reduction (ARR) or risk reduction that are calculated as straight linear equations by subtracting the rate (benefit or harm) in the screened group from that in the unscreened group. Other commonly used measures include the number needed to treat (NNT) and relative risk reduction. The NNT is the number of people who need to be treated before 1 person experiences the benefit or harm; it is calculated as the inverse of the ARR (1/ARR). Relative risk reduction is calculated by dividing the rate of interest in the screened group by the rate in the control group. Research has found that clinician and patient understanding is improved by expressing measures of magnitude as natural frequencies or ARR; they might have difficulty understanding other commonly used measures of effect size such as NNT or relative risk.¹⁵⁻¹⁹ Further, the manner in which the effect size or magnitude is presented can be misleading and overestimate the benefit of the screening interventions.¹⁵⁻¹⁹

How can preventive screening cause harms? False-positive (FP) results: False-positive results are positive

test results in patients in whom the disease or condition is absent. Harms related to FP test results include anxiety and stress related to the diagnosis of the disease or condition, as well as direct harms associated with the extra testing required to determine the presence or absence of disease. One of the most striking examples of the potential for FP results in preventive health screening comes from the National Lung Screening Trial.^{20,21} This randomized clinical trial on screening for lung cancer compared low-dose computed tomography (LDCT) with chest x-ray scans in patients aged 55 to 74 years of age with at least a 30-pack-year history of smoking. It showed that for every 1000 men screened 3 times with LDCT, 391 would have at least 1 positive result, of which 351 (90%) of these would subsequently be found to be FP results.²¹ In this circumstance, it would be important for practitioners to have discussions with their patients about the benefits and harms of screening for lung cancer with LDCT that include information about FP results and the adverse effects of invasive follow-up. Because of the potential for screening-related harms, subsequent management should be done in health care settings where there is expertise in early diagnosis and treatment of lung cancer.²¹

Overdiagnosis: For many primary care practitioners and their patients, the concept of overdiagnosis continues to be both counterintuitive and conceptually challenging. This issue has now been identified as the most important potential cause of harm to healthy people in preventive health screening.^{3,4,22-24} *Overdiagnosis* is defined as the detection of an asymptomatic “abnormality” or “condition” that would ultimately not go on to cause symptoms or death.^{3,4,24} Overdiagnosis occurs in a variety of conditions spanning the spectrum of patients presenting to primary care physicians. This can include breast, prostate, or thyroid cancer, asthma, ischemic heart disease, chronic kidney disease, gestational diabetes, pulmonary embolism, and attention deficit hyperactivity disorder.^{3,4,6,7,8,22-30}

In cancer, overdiagnosis is linked to the concept of heterogeneous disease progression. Some cancers can progress at rates that will cause symptoms or death. Other cancers never progress or progress so slowly that the patient dies of other causes before symptoms appear. Recent evidence also suggests that some “cancers” regress, either because they were not truly cancer or because immune or other mechanisms reverse the pathological process.²⁴ Identification of patients with these nonprogressive or very slowly progressing cancers during preventive health screening would result in overdiagnosis.^{3,24} Definitions of *overdiagnosis* have also broadened to consider the social and ethical issues related to this concept.²⁵

Factors driving overdiagnosis include advances in technology with increasingly sensitive imaging and tests, widening of disease definitions or treatment thresholds that yield much larger populations with a disease, and

“incidentalomas” or incidental findings in individuals being investigated for other reasons.^{3,4,24}

The harms of overdiagnosis occur among patients given a diagnosis of a condition or disease from which they would have never suffered any harm. These patients therefore suffer the harms related to additional diagnostic testing and treatments but do not receive any benefits.^{3,4,23-25} The unfortunate conundrum is that at the time of diagnosis we cannot determine which individual patients are overdiagnosed.^{3,24}

The potential effects of overdiagnosis for a number of cancers are shown in **Table 1**.^{6,8,21,26-29} These results suggest that screening for these conditions has the potential to cause harm to large populations of patients.

How does the practitioner balance benefits and harms to inform decision making? In circumstances in which preventive health guideline recommendations provide strong evidence that the desirable effects of an intervention outweigh the undesirable effects or the undesirable effects outweigh the benefits, clinicians can be confident that most patients would be best served by following the guideline recommendation. In the GRADE system, these types of recommendations would be given a strong recommendation. Examples include screening for cervical cancer,² screening for hypertension,³¹ and not screening for cognitive impairment in older patients³² or developmental delay in children.³³

For most conditions, the trade-off between benefits and harms can be much less clear. In these circumstances, practitioners and patients need to consider both the balance between the benefits and harms of the intervention and the preferences and values of the individual patient. In the GRADE system, these types of recommendations are given weak recommendations. It is important to note that weak recommendations do not indicate that physicians should or should not perform the action or intervention; rather, it means that physicians and patients should engage in shared decision making to determine the most appropriate screening

decision for each patient. In this situation, patients with similar health states for which the recommendation is intended might make different choices on whether to undertake or decline preventive screening maneuvers based on their values and preferences. Examples of such weak recommendations include screening for breast, prostate, or lung cancer.

Conclusion

To improve decision making in preventive health screening, physicians must be familiar with and fluent in many key concepts that inform evidence-based decision making. Understanding these concepts provides physicians with the essential skills to understand and address the complexity and controversy surrounding preventive health screening with their patients. This article, the first in a series on preventive screening, outlines and addresses the concepts related to potential benefits and harms of screening decisions. Future articles will expand on these concepts to provide a foundation of skills that can be used by physicians in shared decision making with their patients. Some of the topics to be addressed in future articles include shared decision making, patient values and preferences, knowledge translation tools, and outcome measures and magnitude of effect.

Dr Bell is Professor in the Department of Family Medicine at the University of Alberta in Edmonton. **Dr Grad** is Associate Professor in the Department of Family Medicine at McGill University in Montreal, Que. **Dr Dickinson** is Professor in the Department of Family Medicine and the Department of Community Health Sciences at the University of Calgary in Alberta. **Dr Singh** is Associate Professor in the Department of Internal Medicine and in the Department of Community Health Sciences at the University of Manitoba in Winnipeg and in the Department of Hematology and Oncology for CancerCare Manitoba. **Dr Moore** is Associate Professor in the Department of Family Medicine at McMaster University in Hamilton, Ont. **Ms Kasperavicius** is a research coordinator for the Knowledge Translation program at St Michael's Hospital in Toronto, Ont. **Ms Kretschmer** is Business Partner, Stakeholder Engagement at the BC Safety Authority in New Westminster, BC.

Competing interests

All authors have completed the International Committee of Medical Journal Editors' Unified Competing Interest form (available on request from the corresponding author) and declare that they have no competing interests.

Correspondence

Danielle Kasperavicius; e-mail kasperavici@smh.ca

Table 1. Potential effects of overdiagnosis of cancer

CANCER	SCREENING TEST	RANGE OF ESTIMATES OF OVERDIAGNOSIS*	COMMENTS
Prostate cancer ⁸	PSA test	40% to 56%	Estimates from the ERSPC trial
Breast cancer ^{6,26}	Mammography	0% to > 50%	Lack of consensus on rates and most appropriate methods to obtain estimates. Issues include study design, length of follow-up, and tumour size
Thyroid cancer ^{27,28}	Ultrasound, CT, MRI	50% to 90%	Overdiagnosis linked to increased incidence of papillary thyroid cancer
Lung cancer ^{21,29}	Low-dose CT	18.5% (95% CI 5.4% to 30.6%)	Estimates from NLST

CT—computed tomography, ERSPC—European Randomized Study on Screening for Prostate Cancer, MRI—magnetic resonance imaging, NLST—National Lung Screening Trial, PSA—prostate-specific antigen.

*Rates of overdiagnosis can vary because of different methodologic approaches used to calculate estimates.²⁹

References

1. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaign—getting past uninformative persuasion. *N Engl J Med* 2012;367(18):1677-9.
2. Canadian Task Force on Preventive Health Care. Recommendations on screening for cervical cancer. *CMAJ* 2013;185(1):35-45.
3. Welch HG, Schwartz LM, Woloshin S. *Over-diagnosed. Making people sick in the pursuit of health*. Boston, MA: Beacon Press Books; 2011.
4. Moynihan R, Doust J, Henry D. Preventing overdiagnosis: how to stop harming the healthy. *BMJ* 2012;344:e3502.
5. Rosenbaum L. Invisible risks, emotional choices—mammography and medical decision making. *N Engl J Med* 2014;371(16):1549-52.
6. Barratt A. Overdiagnosis in mammography screening: a 45 year journey from shadowy idea to acknowledged reality. *BMJ* 2015;350:h867.
7. 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. Epub 2014 Apr 16.
8. Canadian Task Force on Preventive Health Care. Recommendations on screening for prostate cancer with the prostate-specific antigen test. *CMAJ* 2014;186(16):1225-34.
9. International Association of Diabetes and Pregnancy Study Groups. Recommendations on the diagnosis and classification of hyperglycemia in pregnancy. *Diabetes Care* 2010;33(3):676-82.
10. 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.
11. Hoffmann TC, Del Mar C. Clinicians' expectations of the benefits and harms of treatments, screening, and tests. *JAMA Intern Med* 2017;177(3):407-19.
12. Klass W. *Our position on PSA screening*. Toronto, ON: Prostate Cancer Canada Network; 2014. Available from: <https://pcnntoronto.ca/2014/11/14/our-position-on-psa-screening/>. Accessed 2017 May 9.
13. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating of quality of evidence and strength of recommendations. *BMJ* 2008;336(7650):924-6.
14. Bell N, Connor Gorber S, Tonelli M, Pottie K, Singh H, Joffres M, et al. From ABCs to GRADE. Canadian Task Force on Preventive Health Care's new rating system for clinical practice guidelines. *Can Fam Physician* 2013;59:1282-9.
15. Zipkin DA, Umscheid CA, Keating NL, Allen E, Aung K, Beyth R, et al. Evidence-based risk communication: a systematic review. *Ann Intern Med* 2014;161:270-80.
16. Johnston BC, Alonso-Coello P, Friedrich JO, Mustafa RA, Tikkinen KA, Neumann I, et al. Do clinicians understand the size of treatment effects? A randomized survey across 8 countries. *CMAJ* 2016;188(1):25-32.
17. Bartlett G, Gagnon J. Physicians and knowledge translation of statistics: mind the gap. *CMAJ* 2016;188(1):11-2.
18. Schwartz LM, Woloshin S, Black W, Welch GL. The role of numeracy in understanding the benefit of screening mammography. *Ann Intern Med* 1997;127(11):966-72.
19. Schwartz PH, Meslin EM. The ethics of information: absolute risk reduction and patient understanding of screening. *J Gen Intern Med* 2008;23(6):867-70. Epub 2008 Apr 18.
20. The National Lung Screening Trial Research Team. Reduced lung-cancer mortality with low-dose computed tomographic screening. *N Engl J Med* 2011;365(5):395-409. Epub 2011 Jun 29.
21. Canadian Task Force on Preventive Health Care. Recommendations on screening for lung cancer. *CMAJ* 2016;188(6):425-32.
22. McCaffery KJ, Jansen J, Scherer LD, Thornton H, Hersch J, Carter SM, et al. Walking the tightrope: communication overdiagnosis in modern health care. *BMJ* 2016;352:i348.
23. Heath I. Overdiagnosis: when good intentions meet vested interests—an essay by Iona Heath. *BMJ* 2013;347:f6361.
24. Welch GH, Black W. Overdiagnosis in cancer. *J Natl Cancer Inst* 2010;102:605-13.
25. Carter SM, Rogers W, Heath I, Degeling C, Doust J, Barratt A. The challenge of overdiagnosis begins with its definition. *BMJ* 2015;350:h869.
26. The Canadian Task Force on Preventive Health Care. Recommendations on screening for breast cancer in average-risk women aged 40-74 years. *CMAJ* 2011;183(17):1991-2001.
27. Vaccarella S, Franceschi S, Bray F, Wild C, Plummer M, Dal Maso L. Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. *N Engl J Med* 2016;357(7):614-7.
28. Brito JP, Morris JC, Montori VM. Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours. *BMJ* 2013;347:f4706.
29. Patz EF Jr, 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.
30. Carter JL, Coletti RJ, Harris RP. Quantifying and monitoring overdiagnosis in cancer screening: a systematic review of methods. *BMJ* 2015;350:g7773.
31. Lindsay P, Connor Gorber S, Joffres M, Birtwhistle R, McKay D, Cloutier L. Recommendations on screening for high blood pressure in Canadian adults. *Can Fam Physician* 2013;59:927-33 (Eng), e393-400 (Fr).
32. Canadian Task Force on Preventive Health Care. Recommendations on screening for cognitive impairment in older adults. *CMAJ* 2016;188(1):37-46.
33. Canadian Task Force on Preventive Health Care. Recommendations on screening for developmental delay. *CMAJ* 2016 Mar 29. Epub ahead of print.

Suggested additional reading

1. Welch HG, Schwartz LM, Woloshin S. *Over-diagnosed. Making people sick in the pursuit of health*. Boston, MA: Beacon Press Books; 2011.
2. The cancer tests you need and those you don't. *Consum Rep* 2013;78(3):28-33.
