

A Model Too Far

Boris Freidlin, Edward L. Korn

Correspondence to: Boris Freidlin, PhD, Biometric Research Branch, National Cancer Institute, Bethesda, MD 20892 (e-mail: freidlinb@ctep.nci.nih.gov).

Screening for prostate cancer is a “double-edged tool” (1); it is associated with potential benefits (reduction in mortality) and potential harms (unnecessary treatment and emotional distress due to the detection of cancers that would never cause symptoms or death). Availability of good estimates of these benefits and harms is required to properly guide screening practice, both from the public health and individual patient perspectives. The central role in this process is played by randomized screening trials (2). However, it is well recognized that the generalizability of the results from randomized trials may be limited: the results may not be applicable to a population different than that included in the trial, and the results strictly apply only to the specific intervention tested in the trial, not to possible modifications of it. In the screening context, statistical and biological modeling approaches have the potential to increase the generalizability of randomized trial results.

In this issue of the Journal, Gulati et al. (3) use a microsimulation model for prediction of an individual's risk that his cancer has been overdiagnosed, given that he has a biopsied prostate-specific antigen (PSA) screen-detected prostate cancer. An overdiagnosed cancer, as defined by Gulati et al., is one that would not otherwise become symptomatic or clinically apparent during the patient's lifetime in the absence of screening. [Others have defined it as cancers that would otherwise not go on to cause symptoms or death (4,5)]. The most reliable and transparent approach to estimating overdiagnosis of a screening strategy for individuals without apparent prostate cancer is with an appropriate analysis of data from a randomized screening trial. The Gulati et al. approach to estimating the overdiagnosis risk for an individual diagnosed with prostate cancer is based on a microsimulation model that uses the following modeling steps: First, a model for the natural history of prostate cancer progression and clinical detection is developed based on data from a prostate cancer prevention trial. Second, parameters of this model are calibrated to match the Surveillance, Epidemiology, and End Results (SEER) registry incidence data for the period from 1975 to 2005 (6). Third, the calibrated model is used to simulate 10 000 hypothetical life histories (including time of clinical diagnosis and death), with life histories where death precedes clinical diagnosis designated as overdiagnosis. Finally, this simulated data is fit to a logistic regression model to construct a prediction model of the risk of overdiagnosis given age, Gleason score, and PSA level at the time of the diagnosis. It must be recognized that each step of this modeling process makes multiple unverifiable assumptions that can produce bias. These potential biases may be magnified in subsequent modeling steps. Moreover, the model was developed and calibrated on past data (2005 and before). Considering 1) the recent changes in screening patterns [because of new recommendations, including

those of the US Preventative Services Task Force (7)], 2) evolving biopsy referral standards and detection rates, and 3) recent advances in treatment for early and metastatic cancers, it is possible that this prediction model may not apply to current patients.

Gaining reliable randomized evidence for guiding clinical practice and public health policy for screening for indolent diseases such as prostate cancer is a frustratingly drawn-out and expensive process. Modeling, on the other hand, gives answers with relatively modest time and resource requirements. However, the level of evidence that can be generated by modeling is more suited for augmentation of questions directly addressed in a randomized screening trial rather than as a primary source for guiding public health policy. For example, Heijnsdijk et al. (8) and Gulati et al. (6) used microsimulation models with randomized trial data to estimate the relative advantages and disadvantages of various screening strategies. These results could be used to help decide which screening strategies should be tested in future randomized screening trials. It can also be argued that these models can be used to explore and motivate minor adjustments to established screening strategies (eg, a small adjustment to the screening interval/ages or to the threshold for biopsy referral).

The debate about the role of modeling in shaping public health policy will undoubtedly press on (9,10). However, even if the limitations of the modeling could be addressed, by far the most critical question for assessing the current Gulati et al. proposal is whether a model estimating the risk of overdiagnosis is actually helpful for guiding treatment of patients with screen-detected prostate cancer. [Before contemplating screening for prostate cancer, individuals should be advised about the risk of overdiagnosis (11)]. Consider a man who is given a diagnosis of prostate cancer based on a biopsy suggested from a PSA screening test. The microsimulation model estimates the patient's chance of never receiving this diagnosis under a hypothetical assumption that he had not been screened. Gulati et al. hope that this contrafactual estimate may “provide useful information for patients and their physicians seeking to weigh the likely harms and benefits of the treatment options available” (3). The rationale for this expectation is unclear: the risk of overdiagnosis is an indirect reflection of the patient's prognosis. However, for the purpose of guiding patient treatment decisions, the most useful and directly relevant information is, for each possible treatment, its morbidity and the probability of having symptoms from, or dying from, prostate cancer at various times in the future given the patient's prognostic information (including information not used by Gulati et al. because of limitations of the SEER data.) These probabilities can be estimated from treatment arms of randomized trials using conventional statistical methods that do not require complex modeling.

In summary, microsimulation could be useful to expand the applicability of randomized trial results on the benefits and harms of a PSA screening regimen, including the probability of overdiagnosis. However, once an individual has been screened and found to have prostate cancer, the relevant question is the outcomes of various treatments (treatment morbidity, prostate cancer symptoms, and death), not the probability of an event that could have happened if the individual had not been screened.

References

1. Dunn BK, Srivastava S, Kramer BS. The word “cancer”: how language can corrupt thought: antiquated nomenclature is misleading and needs to be revised. *BMJ*. 2013;347(7925):f5328.
2. Kramer BS, Croswell JM. Cancer screening: the clash of science and intuition. *Ann Rev Med*. 2009;60:125–137.
3. Gulati R, Inoue LYT, Gore JL, Katcher J, Etzioni R. Individualized estimates of overdiagnosis in screen-detected prostate cancer. *J Natl Cancer Inst*. 2013;XX(XX):XXX–XXX.
4. Welch HG, Black WC. Overdiagnosis in cancer. *J Natl Cancer Inst*. 2010;102(9):605–613.
5. Woloshin S, Schwartz LM, Black WC, Kramer BS. Cancer screening campaigns—getting past uninformative persuasion. *New Engl J Med*. 2012;367(18):1677–1679.
6. Gulati R, Gore JL, Etzioni R. Comparative effectiveness of alternative prostate-specific antigen-based prostate cancer screening strategies: model estimates of potential benefits and harms. *Ann Intern Med*. 2013;158(3):145–153.
7. Moyer VA. Screening for prostate cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med*. 2012;157(2):120–134.
8. Heijnsdijk EAM, Wever EM, Auvinen A, et al. Quality-of-life effects of prostate-specific antigen screening. *New Engl J Med*. 2012;367(7):595–605.
9. Etzioni R, Gulati R, Cooperberg M, Penson DM, Weiss NS, Thompson IM. Limitations of basing screening policies on screening trials: the US Preventive Services Task Force and prostate cancer screening. *Med Care*. 2013;51(4):295–300.
10. Melnikow J, Lefevre M, Wilt T, Moyer V. Counterpoint: randomized trials provide the strongest evidence for clinical guidelines: the US Preventive Services Task Force and prostate cancer screening. *Med Care*. 2013;51(4):301–303.
11. Brawley OW. Prostate cancer screening: what we know, don’t know, and believe. *Ann Intern Med*. 2012;157(2):135–136.

Note

The authors have no conflicts of interest to declare.

Affiliation of authors: Biometric Research Branch, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD (BF, ELK).