Lead Time and Overdiagnosis

Stuart G. Baker, Philip C. Prorok, Barnett S. Kramer

Correspondence to: Stuart G. Baker, National Cancer Institute, 9609 Medical Center Dr, 5E638 MSC 9789, Bethesda, MD 20892–9789 (e-mail: sb16@nih.gov).

In this issue of the Journal, Etzioni et al. (1) suggest a novel approach to estimating overdiagnosis in breast cancer. However, there is a key assumption in their method regarding the nature of overdiagnosis that is unlikely to hold and leads to an underestimate of the overdiagnosis rate. To understand this assumption, we need to go back to basic definitions. Lead time is the time from detection of preclinical cancer by screening to detection of clinical (symptomatic) cancer in the absence of screening. An overdiagnosed cancer is a cancer detected by screening that would not have presented clinically during that person’s lifetime in the absence of screening.

We define two types of preclinical cancer detected on screening, progressive and nonprogressive, which are linked to two types of overdiagnosis described by Welch and Black (2). A progressive preclinical cancer would present as clinical cancer in the person’s lifetime in the absence of both screening and death from a competing risk. A person with progressive preclinical cancer is overdiagnosed if the person dies from a competing risk before the cancer presents clinically. In contrast, a nonprogressive preclinical cancer would not present as clinical cancer in a person’s lifetime in the absence of both screening and death from a competing risk. A person with nonprogressive preclinical cancer is always overdiagnosed because the cancer never progresses to a clinical state, or even regresses. (Spontaneous breast cancer regression has been observed [3] and could result from a variety of mechanisms, including normalization of disrupted signals from adjacent tissues [4].)

Importantly, only progressive preclinical cancers are related to lead time, because the definition of lead time requires clinical presentation of cancer in the absence of screening. In computing the overdiagnosis rate, Etzioni et al. (1) use estimates of mean lead time based on models that assume that all preclinical cancers are progressive. (Technically these models allow for an infinite lead time, which could, in theory, capture nonprogressive preclinical cancer; however, in practice, the usual lead time distributions, such as exponential, have little probability mass in the upper tails, so practically the preclinical cancers under these models are progressive.)

Although Etzioni et al. recognize this limitation, they do not acknowledge that it could potentially yield a large bias.

To quantitatively investigate this bias, we extended the classic mathematical model for cancer development in three screening states (healthy, progressive preclinical, and clinical) (5,6) to include progressive and nonprogressive preclinical states. We generated hypothetical data under this model, and then we fit the model under the assumption of only progressive preclinical cancer. Let $w(x) = \kappa x^4$ denote the probability of entering the preclinical state at age $x$. Let $\lambda \exp(-\lambda \ u)$ denote the distribution of time $u$ in the progressive preclinical state; this is an exponential distribution, so the mean lead time in the progressive preclinical state is $1/\lambda$. Let $\pi$ denote the probability a preclinical cancer is nonprogressive. Let $\beta$ denote the probability of detecting preclinical cancer on the screening. For simplicity we consider a single screening of a cross-section of persons ages $a = 50, 55, 60, 65$ with follow-up for five years after screening.

Under our model, the probability that screening at age $a$ detects preclinical cancer is:

$$P_a = (1-\pi)\beta \int_0^a w(x) \exp\{-\lambda (a-x)\} \, dx + \pi \beta \int_a^\infty w(x) \, dx. \tag{1}$$

The first term in Equation 1 is the probability of entering the progressive preclinical state before age $a$, no incidence of clinical cancer before age $a$, and detection on screening. The second term in Equation 1 is the probability of entering the nonprogressive preclinical state before age $a$ and detection on screening.

Under our model, the probability of interval cancer between ages $a$ and $a+5$ after no detection on screening at age $a$ is:

$$I_a = (1-\beta) (1-\pi) \int_0^a w(x) \left[ \exp\{-\lambda (a-x)\} - \exp\{-\lambda (a+5-x)\} \right] \, dx + (1-\pi) \int_a^\infty w(x) \left[ 1 - \exp\{-\lambda (a+5-x)\} \right] \, dx. \tag{2}$$

The first term in Equation 2 is the probability of entering the progressive preclinical state before age $a$, not detecting preclinical cancer on screening, and clinical cancer incidence in the next five years. The second term in Equation 2 is the probability of entering the preclinical preclinical state after screening at age $a$, and clinical cancer incidence up to five years after screening.

Let $S_a$ denote the probability of surviving from age 50 to age $a$ for the US population (7), which we use as an approximation for the survival rate for competing risks. For simplicity, we computed these survival probabilities at five year intervals from 50 to 100, a set of ages we denote by $A$: the values are 1, 0.97, 0.94, 0.89, 0.82, 0.73, 0.6, 0.44, 0.25, 0.1, 0.02. Under our model, the probability of overdiagnosis given screen detection at age $a$ is:

$$\theta = \sum_{a \in A} (S_{a+5} - S_a) \exp\{-\lambda (a+5-50)\}(1-\pi) + \pi. \tag{3}$$
The first term in Equation 3 is the probability of death from competing risk before cancer incidence among those with progressive preclinical cancer. The second term in Equation 3 is the probability of nonprogressive preclinical cancer.

We created perfect fit data under a true model with \( \kappa = 10^{-11}, \beta = 0.9, \pi = 0.3, \) and \( \lambda = 1 \) (implying a mean lead time of one year). We specified \( N = 10000 \) persons screened at each age of \( a = 50, 55, 60, \) and 65. Let \( n_a = N P_a \) denote the number of persons detected with preclinical cancer on screening at age \( a \). Let \( n_a = N I_a \) denote the number of incident cancers within five years of the screening at age \( a \). Making the incorrect assumption of only a progressive preclinical state (so \( \pi = 0 \)), for \( \beta = 1, 0.9, 0.8, 0.7, \) we computed maximum likelihood estimates of \( \kappa \) and \( \lambda \) using the likelihood kernel:

\[
L(\kappa, \lambda | \beta) = \sum \kappa_m \log(P_m) + \sum \kappa_n \log(I_n) \\
+ \sum \kappa_s (N - n_s) \log(1 - P_s - I_s).
\] (4)

Substituting the estimate of \( \lambda \) into Equation 3 with \( \pi = 0 \) gives an estimate of the overdiagnosis rate. The incorrectly specified model assuming only progressive preclinical cancers fits well, as evidenced by the small deviances, but the estimated overdiagnosis rates are much smaller (1% to 5%) than for the true model (30%) that included the additional nonprogressive preclinical cancer (Table 1). Hence, the assumption that preclinical cancer is only progressive, which is implicit in the models used by Etzioni et al. to estimate mean lead time, can yield substantially lower estimates of the overdiagnosis rate than reality.

What about other methods of estimating the overdiagnosis rate of breast cancer? Etzioni et al. focused on the estimate of 31% by Bleyer and Welch (8). Because the study of Bleyer and Welch is observational, there is a possibility of bias from incorrectly specifying the baseline incidence of breast cancer in the absence of screening. To mitigate this potential bias, Bleyer and Welch performed a sensitivity analysis and estimated an overdiagnosis rate as low as 22% under their most extreme case.

The most convincing studies of overdiagnosis come from randomized trials with long-term follow-up after screening stops. Etzioni et al. voiced concern about biased results from randomized trials but this may be related only to those trials with only short-term follow-up. Etzioni et al. may also be concerned that the participants in randomized trials may differ from the general population in ways related to the overdiagnosis, but this would also be an issue with using mean lead time estimates from randomized trials in Etzioni et al. In a randomized screening trial with sufficient follow-up after the last screening, a good estimate of the overdiagnosis rate is a persistent excess incidence in the control arm vs the screened arm divided by number screen-detected (2). Using this method, Welch and Black (2) estimated an overdiagnosis rate of 24% in the Malmo mammography trial involving 15 years of follow-up after the 10-year trial. Based on a binomial distribution, the 95% confidence interval for this estimate of 24% is 20% to 28%. In contrast, in the “Health Insurance Plan (HIP) of Greater New York” mammography trial there was no excess incidence after seven years (9), suggesting no overdiagnosis. It may be that mammography in the Malmo trial (with screening from 1976 to 1986 [10]) was detecting lesions that would have been missed by the mammography extant at the time of HIP trial (with screening from 1963 to 1969 [9]).

---

Table 1. Estimates based on fitting an incorrect model for only progressive preclinical cancer to perfect fit data generated from a true model involving both progressive and nonprogressive preclinical cancer.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>True model</th>
<th>Estimates under incorrect models, given various values of ( \beta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviation</td>
<td>0.13</td>
<td>0.12</td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.11</td>
<td>0.09</td>
</tr>
<tr>
<td>( \pi )</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>( \kappa )</td>
<td>0.9</td>
<td>0.9</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>1.00 \times 10^{-11}</td>
<td>1.32 \times 10^{-11}</td>
</tr>
<tr>
<td>( \lambda_{\text{mean lead time}} )</td>
<td>1.32 \times 10^{-11}</td>
<td>1.32 \times 10^{-11}</td>
</tr>
<tr>
<td>( \theta = \text{overdiagnosis rate} )</td>
<td>0.300</td>
<td>0.012</td>
</tr>
</tbody>
</table>
In summary, based on our concerns about key assumptions of Etzioni et al. about progressive preclinical cancers and our critical analysis of this field, we would not rule out an overdiagnosis rate as high as 28% for mammography based on data from the era of mammography randomized trials. The higher sensitivity of modern screening modalities may yield an even higher rate of overdiagnosis. Accurate estimation of the overdiagnosis rate with today’s screening modalities is an urgent research challenge.

References


Funding

This work was supported by the Division of Cancer Prevention of the National Cancer Institute.

Notes

The funder did not have a role in the writing of the article or the decision to submit the article for publication. The opinions presented should not be viewed as official opinions or positions of the National Cancer Institute.

Affiliation of authors: Division of Cancer Prevention, National Cancer Institute, Bethesda, MD.