

Over Diagnosis in Screening: Does It Make Sense?

Dongfeng Wu*

Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, KY 40202, USA

Cancer screenings have been carried out in developed countries over the past four decades. While it may seem to be effective in detecting tumors early, there are many arguments or concerns in the issue of over diagnosis; namely, the diagnosis of cancer that will not cause symptoms, nor death in a patient's lifetime. Will regular screening contribute to a greater chance of over diagnosis? What is the possibility of over-diagnosis for those individuals diagnosed with cancer early by screening? Or what is the percentage/proportion of over-diagnosis among the screen-detected cancer patients? How should we estimate this percentage?

There are mainly two approaches to estimate the percentage of over diagnosis: one is to use the incidence rate in cohort studies or cancer registry, the other is to use probability modeling. There are some methods that may try to combine these two, but in fact they usually fall into the first approach, with some slight usage of probability concept in the cohort studies. Let's use breast cancer screening as an example to describe these two approaches.

The first approach usually compares the differences in the cumulative incidence rates between the study (screened) group and the control group over a long follow-up period in a cohort studies, and based on the differences in the incidence rates to infer the percentage of over diagnosis. The idea is based on a simple assumption: screening will remove most cancerous cases (including over diagnosed cases), and when screening is stopped for the study group, the incidence of breast cancer should decrease over time, while the control group has a relatively stable incidence rate; by studying the difference, the rate of over diagnosis can be obtained. There are some variations in the application of this method. For example, some researchers use the cumulative incidence rates over the same time period for the two groups [1], while others use different time period, and make some adjustments for the lead time to compensate [2]. Hence the inference results varies dramatically, estimates of over diagnosis vary from 7% to 52%. While long term cohort studies are valuable and provide much important information in many areas, there are a few flaws in this approach: a) Results based on one cohort study cannot be extended to other scenarios. The reason is that for this one particular cohort study, with one specific screening interval, the result may be correct; however, one cannot use this result to make inference for studies with different screening intervals. Meanwhile, it is of great value for policy makers or general public to know how the proportion of over diagnosis will change with different screening frequencies. b) This approach usually needs a long follow-up time period to collect the cumulative incidence data from both the study and the control groups, often up to 10 - 15 years, in order to compare the differences in the cumulative incidence rates. The answer may be correct and can be used as a general criterion for that specific cancer with the specific screening interval; however, it is not cost effective. c) This approach may greatly overestimate the rate of over diagnosis if it is used inappropriately. For example, some research compared the cumulative incidence rate over different time period, using existing data from cancer registry. The problem is that in the 1990s breast cancer has a higher incidence in many developed countries due to higher usage of hormone replacement therapy (HRT), and the incidence rate dramatically goes down after 2002 when the

HRT was reduced or stopped [3]. The result is that the over diagnosis is severely over estimated, to be above 50% [2,3].

The second approach is to use probability modeling to overcome the limitations mentioned above in the incidence rate studies. This approach assumes that all cancer patients will develop through a three states model: disease-free, preclinical, and clinical, where the disease-free state means that there is no cancer or the state in which the cancer cannot be detected; the preclinical state means that an asymptomatic individual unknowingly has cancer that a screening exam can detect; the clinical state refers to the state that the cancer manifests itself in clinical symptoms. This three-state model is commonly used and accepted by researchers since the end of 1960s. There are three key parameters in this model: the screening sensitivity, the sojourn time (or time spent) in the preclinical state, and the transition probability from the disease-free state to the preclinical state. These three key parameters can be estimated from likelihood functions based on cohort cancer screening data, such as number of people screened, number of confirmed cancer cases at each screening and number of incidence between two screenings [4]. Instead of using incidence rates in cohort studies, this approach studies the development of cancer with screening frequencies and expected human lifetime. This is because over diagnosis means those patients whose clinical symptoms would not have appeared before death. The probability of over diagnosis can be expressed mathematically as a function of the three key parameters, the (future) screening schedule or frequency, and the expected lifetime of patients at their current age [4-7]. This approach builds up a strict probability formula for the estimation or prediction of over diagnosis. For example, policy makers can use this method to study how the percentage/probability of over diagnosis changes with future screening frequency, and how it changes for different cohorts of people and for different kinds of cancer screening, how it changes when the screening modality (hence the sensitivity) changes, or the sojourn time changes. The advantages of this approach are that long term follow up data is unnecessary, because it uses the existing screening data to obtain information on the three key parameters; and predictive estimation for different screening frequencies in the future is possible. There are some limitations of this model: a) the accuracy of the estimation of over diagnosis depends on the accuracy of the three key parameters directly. b) the expected lifetime needs to be estimated accurately [6]. Much research has been done to estimate the three key parameters, and new approaches still appear constantly these days [5]. The expected lifetime distribution for males and females can be derived from the actuarial life table in the US Government Social Security Administration's website

*Corresponding author: DongfengWu, Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, KY40202, USA, E-mail: dongfeng.wu@louisville.edu

Received June 04, 2012; Accepted June 04, 2012; Published June 05, 2012

Citation: Wu D (2012) Over Diagnosis in Screening: Does It Make Sense? J Biomet Biostat 3:e110. doi:[10.4172/2155-6180.1000e110](https://doi.org/10.4172/2155-6180.1000e110)

Copyright: © 2012 Wu D. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

[6]. The estimated probability of over diagnosis for breast cancer is about 6- 9%, very close to the results in the cumulative incidence study by Zackrisson, et al [1], showing that the probability modeling is correct besides theoretical proof.

In summary, over diagnosis are a big challenge in early detection of cancer screening and other disease testing. We endorse the probability modeling because of its accuracy and effectiveness in predictive estimation, although the incidence rates in cohort studies provide an intuitive answer and rough estimates for this issue. One important thing we need to keep in mind is to put research in the right perspective. Just as Dr. Len's Cancer Blog [8] pointed out: when one is diagnosed with cancer, who would be so confident to say that his/her case is an over diagnosed one, and refuse to receive any treatment?

References

1. Zackrisson S, Andersson I, Janzon L, Manjer J, Garne JP (2006) Rate of over-diagnosis of breast cancer 15 years after end of Malmö mammographic screening trial: follow-up study. *BMJ* 332: 689-692.
2. Jørgensen KJ, Gøtzsche PC (2009) Overdiagnosis in publicly organised mammography screening programmes: systematic review of incidence trends. *BMJ* 339: b2587.
3. Duffy SW, Lynge E, Jonsson H, Ayyaz S, Olsen AH (2008) Complexities in the estimation of overdiagnosis in breast cancer screening. *Br J Cancer* 99: 1176-1178.
4. Wu D, Rosner GL (2010) A projection of true-early-detection, no-early-detection, overdiagnosis and not-so-necessary probabilities in tumor screening. *JSM Proceedings, Biopharmaceutical Section, VA*.
5. Wu D, Rosner GL, Broemeling L (2005) MLE and Bayesian inference of age-dependent sensitivity and transition probability in periodic screening. *Biometrics* 61: 1056-1063.
6. Wu D, Kafadar K, Rosner GL, Broemeling LD (2012) The lead time distribution when lifetime is subject to competing risks in cancer screening. *Int J Biostat* 8.
7. Wu D, Pérez A (2011) A limited review of over diagnosis methods and long term effects in breast cancer screening. *Oncol Rev* 5: 143-147.
8. Lichtenfeld JL (2010) Crux of the dilemma: whose cancer is over diagnosed? Dr. Len's Cancer Blog, American Cancer Society.