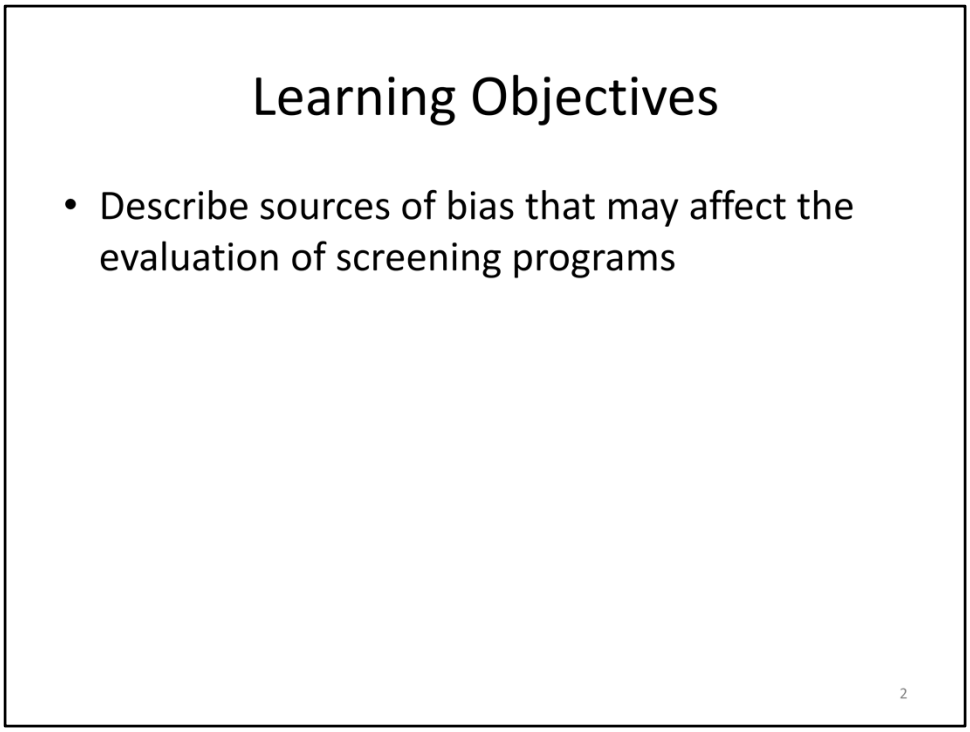
# Screening for Disease
# Part IV

**Julie Stoner, PhD**
Professor
Department of Biostatistics and Epidemiology
College of Public Health

(most slides courtesy of Laura Beebe, PhD)

1

Welcome to this four-part series focused on epidemiologic and biostatistical methods related to disease screening. In this fourth and final segment, we will discuss sources of bias that may impact our evaluation of screening and diagnostic tests.




After viewing this segment, you will be able to describe sources of bias that may affect the evaluation of screening programs.

# Evaluation of a Screening Program

- Does early detection of disease result in benefits to the individuals being screened?
- Is the screening program effective in reducing morbidity and mortality from disease?

3

## Outcomes Measures:

1. Reduction of mortality in the population being screened.
2. Reduction of case-fatality in screened persons.
3. Increase in percent of cases detected at earlier stages.

4

This slide includes a listing of commonly used outcome measures that we can use in our evaluation of screening programs.

To evaluate the impact of a screening program, we may want to determine the program results in a reduction of mortality in the population being screened.

Furthermore, it may be that the screening program results in a reduction of the case-fatality ratio in screened persons because the disease process can be identified earlier and treatment started earlier thereby reducing the risk of mortality.

Finally, the program may result in an increase in the percentage of cases who are detected at earlier stages of disease and can be more effectively treated and managed to minimize the risk of morbidity and mortality.

## Outcomes Measures:

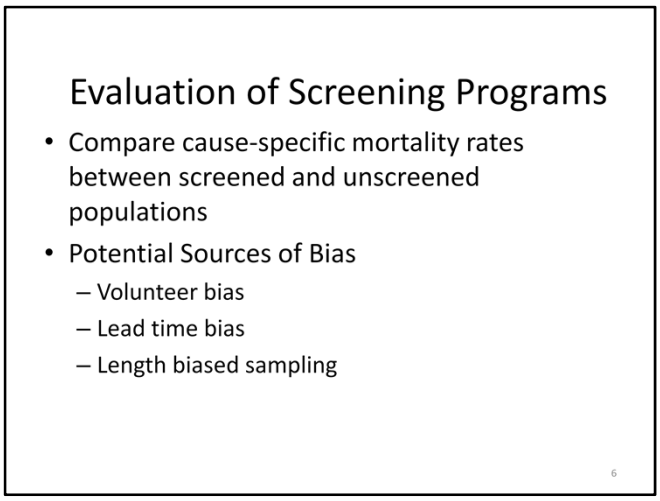
4. Reduction in complications.
5. Prevention of/reduction in recurrences or metastases.
6. Improvement in quality of life in screened individuals.

5

An effective screening program may result in a lower burden of complications from the disease by identifying the disease earlier in the natural history of the disease or through earlier detection, complications due to more invasive treatment approaches, reserved for later stages of disease, can be avoided.

An effective screening program may result in prevention or reduction of disease recurrence or metastasis, again by identifying cases earlier in the course of disease.

Finally, the program may be effective in improving the quality of life of screened individuals through earlier diagnosis of disease.




One approach to evaluating screening programs is to compare cause-specific mortality rates between screened and unscreened populations.

When using this approach, there are several sources of bias that should be kept in mind, including volunteer bias, lead time bias, and length biased sampling. We will now discuss each of these possible sources of bias.

## Volunteer Bias

- Screening population made up of:
  - Self-selected volunteers
  - "Worried well"
  - May be healthier or at higher risk of developing the disease than those that don't participate

7

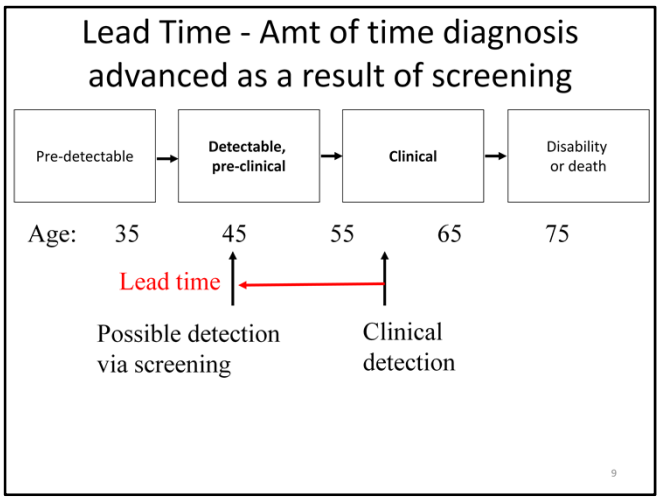
Volunteer bias occurs when the screening population is made of self-selected volunteers, the "worried well", or populations that are different in their health status, either healthier or who are at higher risk of developing the disease, than those who do not participate. Comparting outcomes between the screened and non-screened populations may result in a difference in cause-specific mortality; however, this difference may be due to differences in the initial health status of the populations and not entirely due to the screening program itself.

## Lead Time Bias

- Lead Time - Amount of time diagnosis advanced as a result of screening
- Lead Time Bias - Survival may falsely appear to be increased among screened group
  - simply because the diagnosis was made earlier in the course of the disease

8

Lead time bias occurs when the time from diagnosis to death, for example, is prolonged because the patient is diagnosed at an earlier time point in their disease process. It appears that the screened patients survived for a longer period of time, but the survival duration is prolonged because of the early identification of the diseased case.

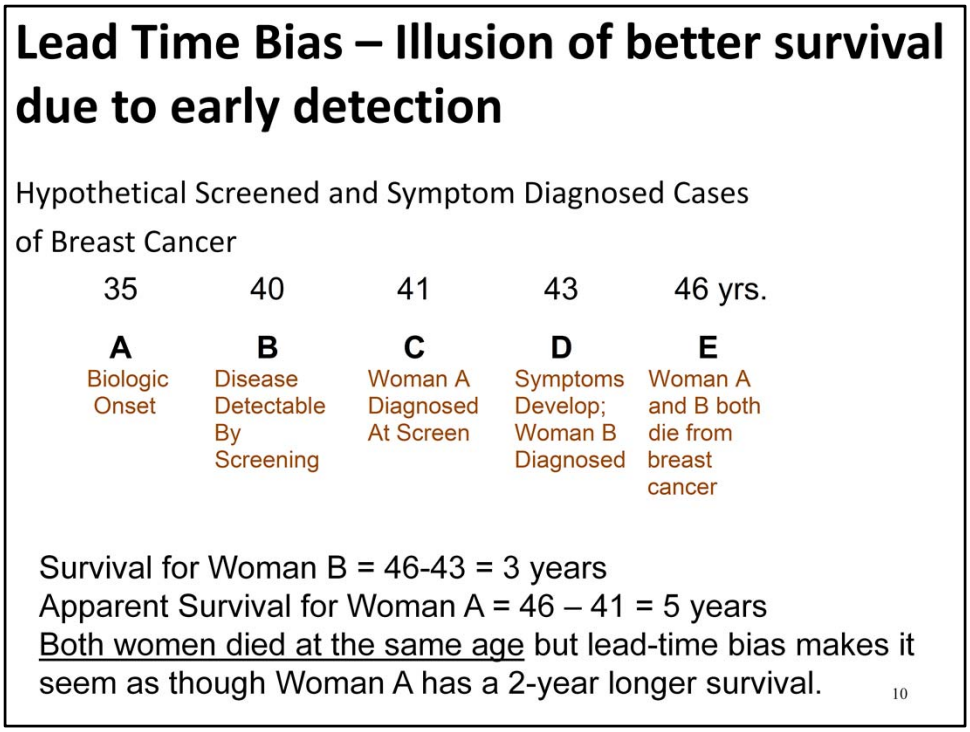



This schematic diagram presents an example of lead time bias.

Assume that clinical detection, after the onset of symptoms, occurs at the age of 59.

With an accurate screening test, we may be able to identify a disease case earlier, before the onset of symptoms, at the age of 45.

The time gap between the age of 45 and 59 is the potential lead time.




Consider another example where two women die from breast cancer at the age of 46 years.

Woman A undergoes early screening and her disease is diagnosed at the age of 41, resulting in an apparent post-diagnosis survival time of 5 years.

Woman B does not undergo early screening and her disease is diagnosed at the age of 43, once symptoms began to develop. Her post-diagnosis survival time is only 3 years.

It appears that Woman A survived 2 years longer after her diagnosis compared to woman B; however, both women survival to the age of 46 years.

This is an example of lead time bias. The post-diagnosis survival is longer due to the early timing of diagnosis and is not longer because of an actual survival benefit.

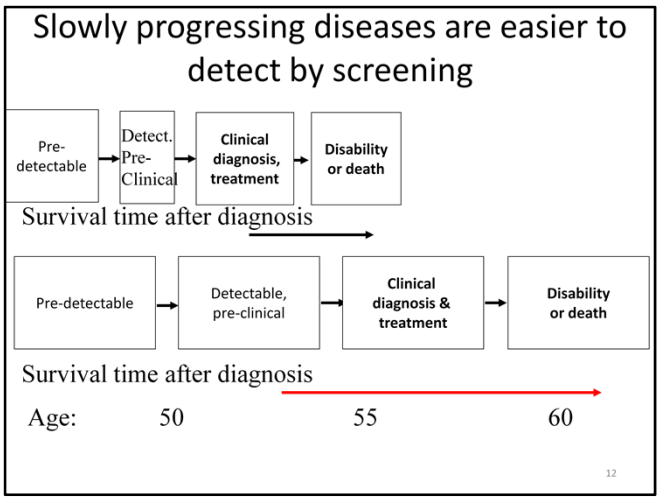
## Length Biased Sampling

- Less aggressive forms of a disease are more likely to be picked up by screening because have a longer preclinical phase
- Less aggressive forms of disease usually have better survival
- Thus, screen detected cases appear to have better survival

11

Length biased sampling may also impact our evaluation of a screening program.

For example, if less aggressive forms of a disease are more likely to be picked up by screening because they have a longer preclinical phase, and less aggressive forms have a longer survival, it may appear that the screened cases have better survival. However, this result is driven by the length biased sampling approach.




This series of diagrams illustrates the potential impact of length biased sampling.

The first patient has a short pre-clinical detection period and a short survival time after diagnosis.

In contrast, the second patient has a longer pre-clinical detection period and has a longer survival time.

The screening program is more likely to enroll a participant with a longer pre-clinical detection period than a patient who progresses from the pre-clinical period to the clinical, symptomatic, period more quickly.

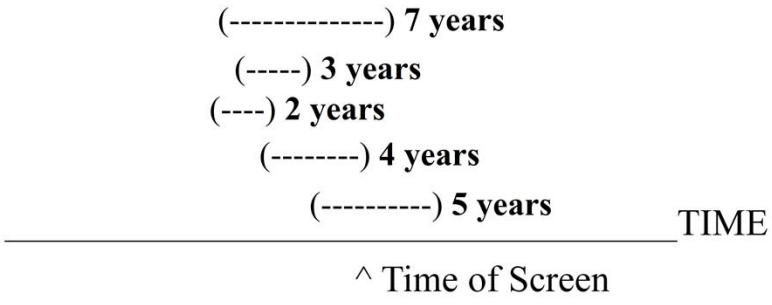
As a result, the comparison between the screened and unscreened populations may result in longer survival for the screened population because of the length biased sampling and not the impact of the screening program itself.

## Length- Biased Sampling

Consider the following screening program that was administered to five individuals.

(-------) is the preclinical phase for a particular person.

Which individuals are picked up at the screening?




13

As another example, consider this diagram showing the preclinical period for five patients.

At the time of screening, we are more likely to screen patients with the longer preclinical phase compared to patients with short preclinical phases.

Therefore, it may appear that the screening program results in prolonged survival, but this may be driven by the fact that those with longer preclinical phases, who are more likely to be enrolled in the screening program, also have a longer survival due to the slower course of their disease and not due to the impact of the screening and subsequent treatment programs.

## Characteristics of Screening test

- Reliable – get same result each time (precise)
- Validity – get the correct result (accurate)
  - Sensitive – correctly classify cases
  - Specific – correctly classify noncases

## Interpretation of Screening Test

- PPV – proportion of positive test results identifying cases
- NPV – proportion of negative test results identifying non-cases

14

In summary, when developing a screening program, we want to utilize a screening test that is both reliable/precise and valid/accurate. We have learned about measures to quantify both reliability and validity of a screening test and have focused on sensitivity and specificity as measures of accuracy.

We have also learned how to calculate positive and negative predictive values, which are often of greatest interest to patients in that they convey the probability of disease among those who test positive and the probability of no disease among those with a negative test result, respectively.

## Summary

- Identified measures for evaluating screening program
- Describe sources of bias that may affect the evaluation of screening programs

15

In summary, we have discussed outcome measures that are useful in evaluating the impact of screening programs. We have also discussed common sources of bias that may impact our evaluation of a screening program.

This includes the module focused on screening test evaluation methods.