

# **The Bernard G. Greenberg Lectures**

**University of North Carolina**

**May 12-13, 2010**

**Marvin Zelen**

**Harvard University**

**Part I . Models in Action : How simple models may lead  
to a better understanding of observed phenomenon**

**Part II. The Early Detection of Disease**

# My Collaborators



**Sandra J. Lee, Dana-Farber Cancer Institute  
and Harvard University**



**Ori Davidov, University of Haifa  
(now visiting NIEHS)**



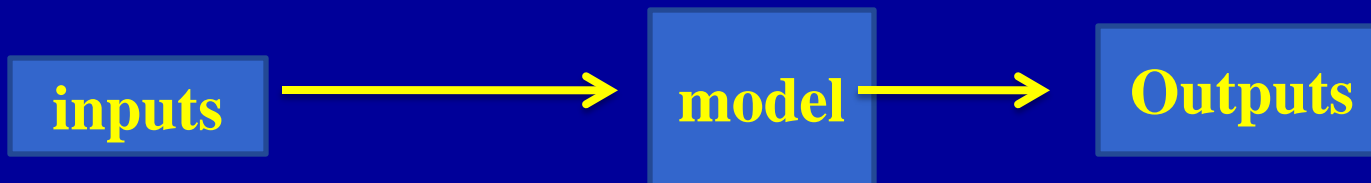
**Hui Huang , Dana-Farber Cancer Institute**

# **Part I : Models in Action. How simple models lead to a better understanding of observed phenomenon.**

- **Introduction and general remarks**
- **Illustrations**
  - 1. Waiting for a bus. Why does it take so long?**
  - 2. Family history and disease. Are family members at risk if there is at least one member of family with disease.**
  - 3. Cell kinetics. Why do people become refractory to therapy?**
  - 4. Prevalence and incidence of disease. What can we learn knowing prevalence and incidence.**
  - 5. Data base studies**
  - 6. Recurrence time distributions**
  - 7. Length biased sampling**

# Models

- **Many areas in science and technology cannot seek reliable answers to relevant questions by experiments or empirical studies --- not feasible --- too costly --- takes too long.**
- **Models may be developed to describe the phenomenon. The goals of such models are to use known inputs to predict outcomes and to determine how changes in the inputs affect outcomes.**



- **Models may serve as a guide for data collection and suggest experiments.**

# Models in Science

Science has many different kinds of models

Animal Models : To study human disease

Physical models : To demonstrate the apparent truth of a principle.

Mathematical Models : To gain insight and understanding of a phenomenon which is being observed.

# Models and Their Uses

- **Nearly all models are simplifications of the phenomenon being observed . However in order for a model to be useful, it must capture the important features of the phenomenon being observed.**
- **Models can be used to guide the design of experiments**
- **Models may suggest the kinds of data to be collected**

# Limitations of Models

- **Models may not entirely reflect features of the observed phenomenon; e.g.:**
  - animal disease may not be same as human disease
  - prototype model may not scale up to actual model
  - equations may not properly describe phenomenon
- **Parameters of model may not exist or estimates may be unreliable; e.g. growth of tumor**

# Probability Models

**A probability model is a mathematical model which incorporates the probability aspects of the observed phenomenon when the observed phenomenon has outputs which are subject to probability laws.**

**If the observed process has a time parameter ( or is indexed by something equivalent) we refer to the probability model as a stochastic process.**



# Data Models

- **Underlying all statistical analyses is a model of how the data was generated.**
- **The most common models relate to linear models where there are unknown parameters and a data structure characterizing the expected values and second moments ( variances and covariances) of the data.**
- **Example :**

**Measurement = True value + random error**

# Mechanistic models

- **Another kind of model is “mechanistic models”. Such models incorporate the main features of an observed phenomenon and provide a way for predicting outcomes when the input variables are changed. These models attempt to impart an understanding of how outcomes are changed with changing inputs.**

**“They drive our thinking”**

- **In this series of lectures we will attempt to illustrate some applications of mechanistic models.**

# Why Use Mechanistic Models?

Many areas in science and technology cannot seek reliable answers to relevant questions by experiments or empirical studies--not feasible--too costly--takes too long -- may be unethical.

Examples : How often should a person visit their physician?

Should a therapy be withheld from a man diagnosed with prostate cancer ( no signs or symptoms ) because there is a high probability that it will never be clinical in his lifetime?

Model may contain as inputs : Individual health history, risk factors for potential diseases and possible outcomes if disease is discovered early .

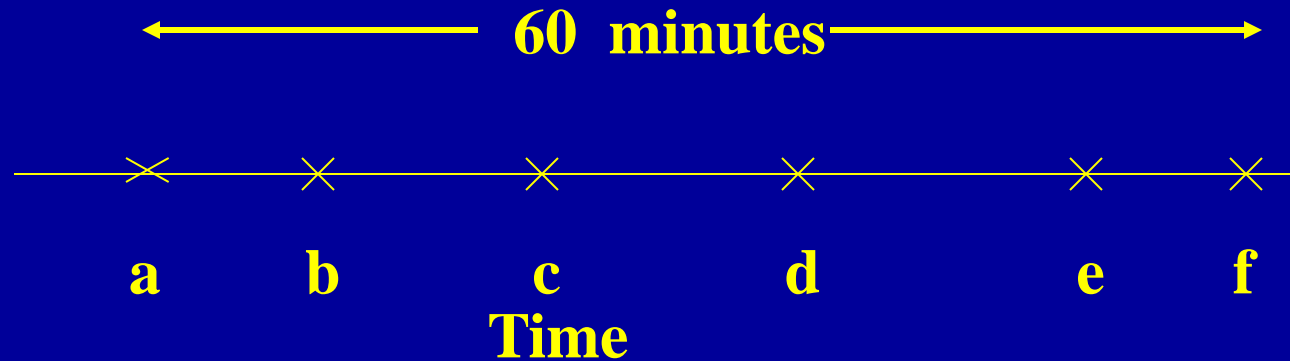
# Waiting for a Bus

**Problem: Why is it when I wait for a bus it always seems to take an unusually long time? It seems that I just missed the bus.**

**Was I just born unlucky? – Or do I have a psychological hang-up? – Or is it really true.**

# Hypothetical Situation

Observed time between bus arrivals

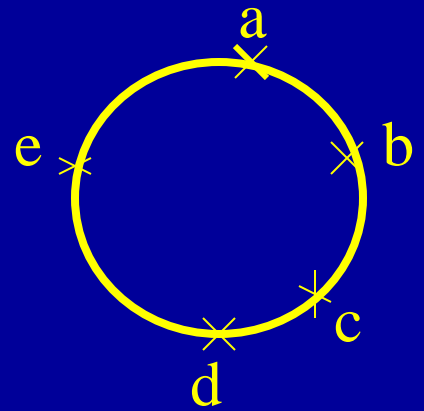


**5 Intervals in 60 minutes**

**MBTA = Mean time between arrivals**

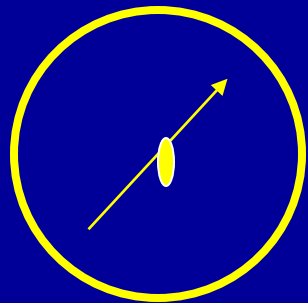
**= Total observed time/ No. of intervals =  $60/5 = 12$  minutes**

**MBTA = 12 minutes**



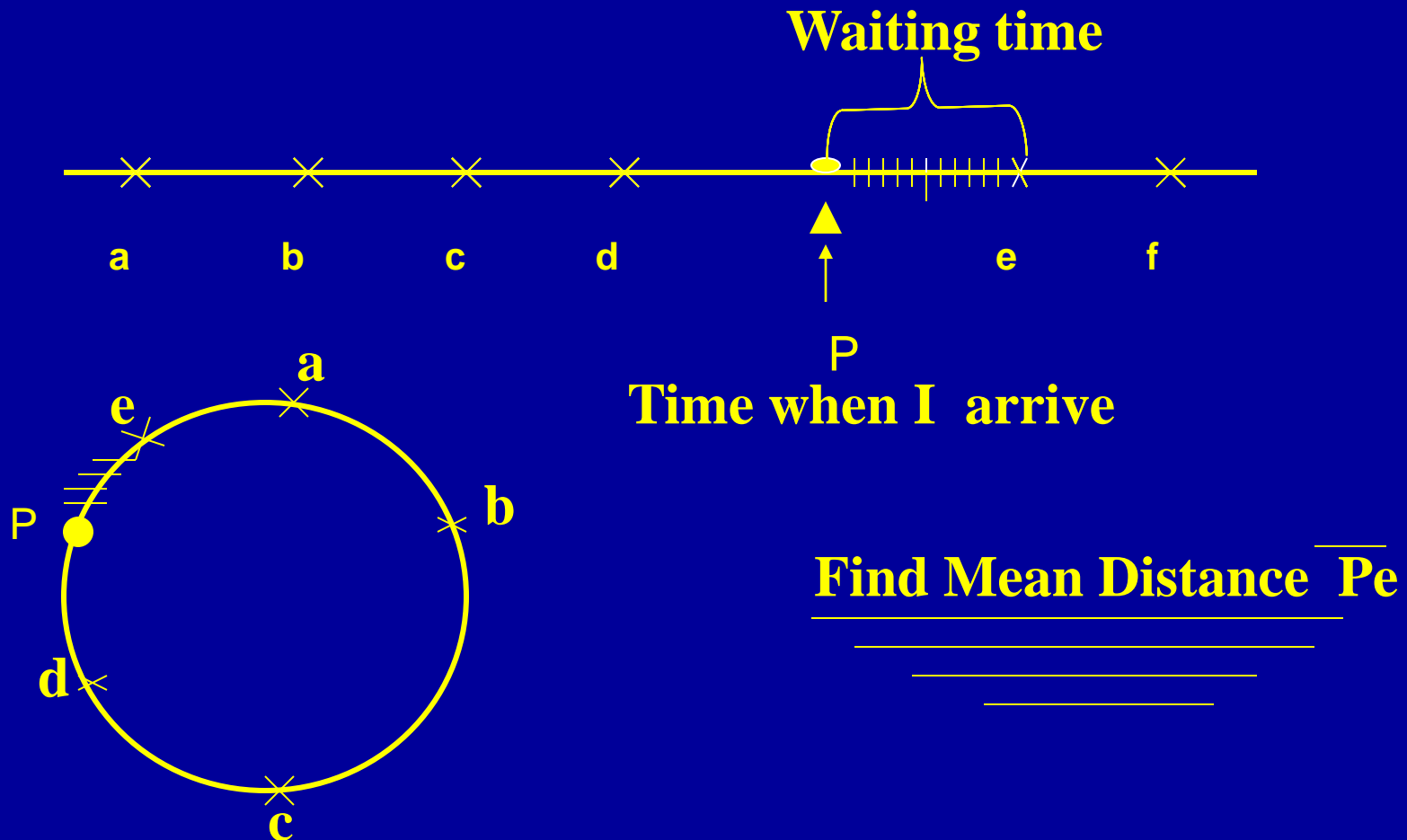
**5 intervals**  
**60 minute clock**

**Equivalent to placing 5 points at random on clock; i.e.**



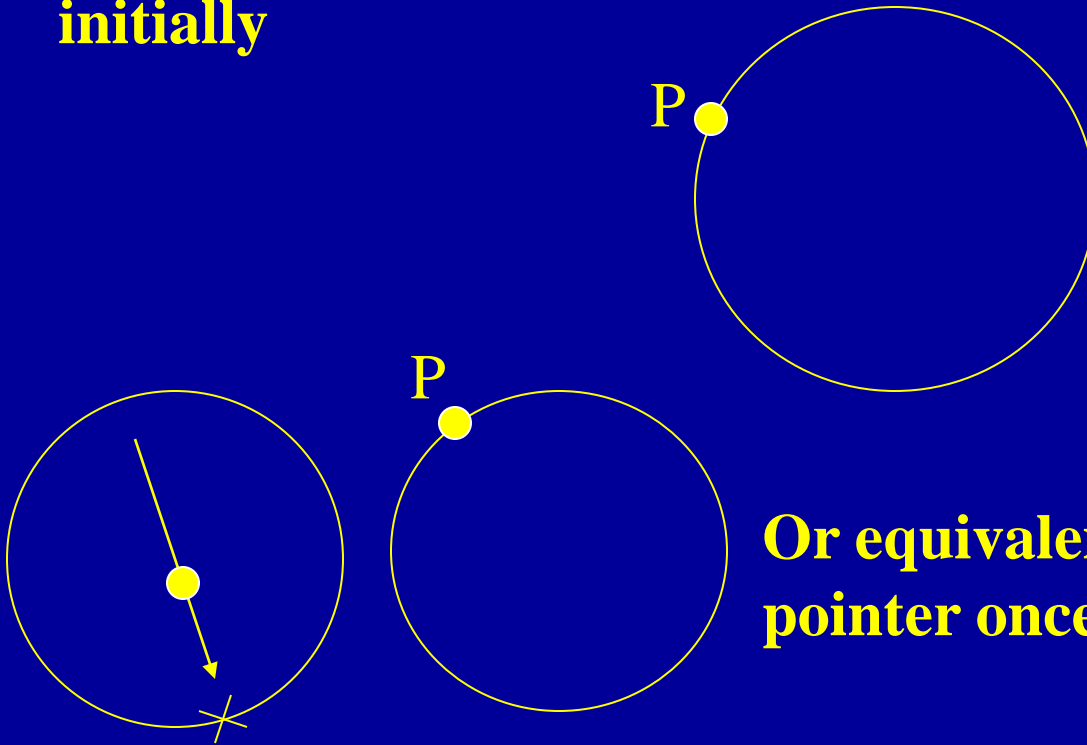
**Flip pointer 5 times**

Suppose I decide to catch a bus sometime during the observational period. How long do I have to wait for the bus to arrive (on the average)? What is the expected mean waiting time if MBTA = 12 minutes.



## Informal Analysis

Suppose the arrival point is placed on the clock initially



Or equivalently spin the pointer once and move  $x$  to  $P$ .



## Conclusion:

Placing a point on a clock is the same as spinning a pointer once. But we have 5 spins and one placed point. This is equivalent to 6 spins and results in 6 intervals. Hence average time between point is

$$\frac{60 \text{ minutes}}{6} = 10$$

Therefore average waiting time = 10 minutes.

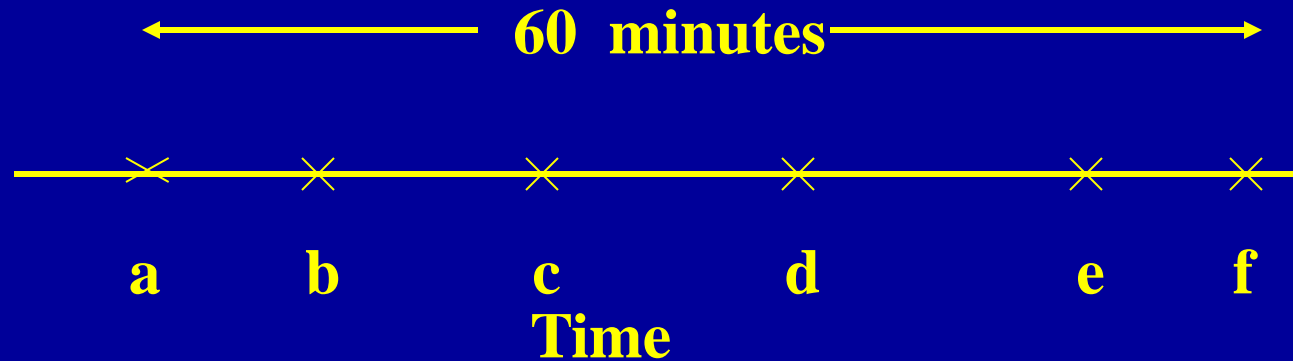
**Reason for unexpected result is that my arrival at the bus stop is equivalent to choosing an inter-arrival interval. The probability of choosing an interval is proportional to the length of the interval.**

**In reality I am sampling possible intervals. The longer the interval the greater the chance of choosing it. This kind of sampling is called Length Biased Sampling.**

**Length Biased Sampling occurs often in modeling biomedical phenomenon.**

# Choosing an interval in proportion to its length

Observed time between bus arrivals



Choosing a point at random on the line is equivalent to intersecting an interval proportion to its length

# What happens to my waiting time if the time between arrivals is smaller?

- Suppose the time between bus arrivals is 6 minutes (double the number of buses ).
- This is equivalent to observing 10 intervals (11 bus arrivals within the hour). Hence by our model, we will have 10 points (buses) and one placed point on the circle.
- The mean waiting time is  $60/11 = 5.5$  minutes.  
“Equivalent” to missing the bus by half a minute.

## 2. Drug Therapy and Cell Kinetics

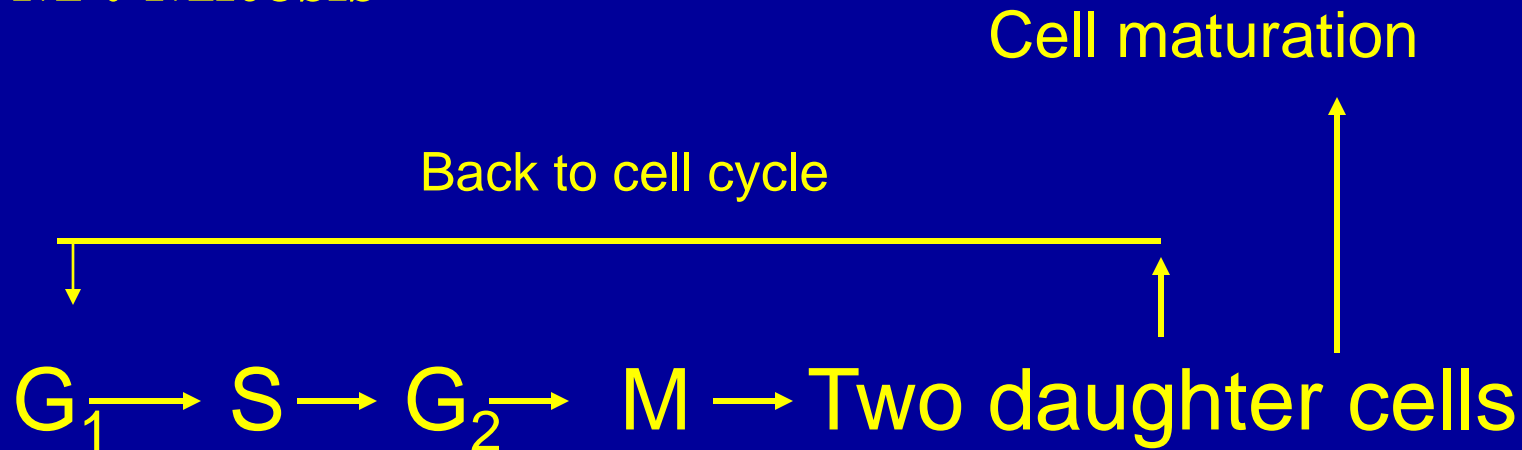
- Life cycle of cell consists of four states

**G<sub>1</sub>** : pre-DNA synthesis

**S** : DNA synthesis

**G<sub>2</sub>**: post-DNA synthesis

**M** : Mitosis



# Drug Therapy

- **Some drugs (chemotherapy against cancer) may only destroy tumor cells when they are in the S phase.**
- **As a result the longer a tumor cell is in the S phase the more likely it will be destroyed.**
- **Consequently the probability of a tumor cell being destroyed may change with additional courses of therapy as the cells with longer S phases tend to be removed from the population. Eventually the tumor cells remaining will have very short S phases and it will be difficult to destroy them .**
- **With repeated courses of therapy , the drug will not seem to have any therapeutic effect and the patient will be described as being refractory to treatment with this drug.**
- **If therapy can be continually infused, this may increase the probability of destroying tumor cells from a changing population.**

# Process of Cell Destruction

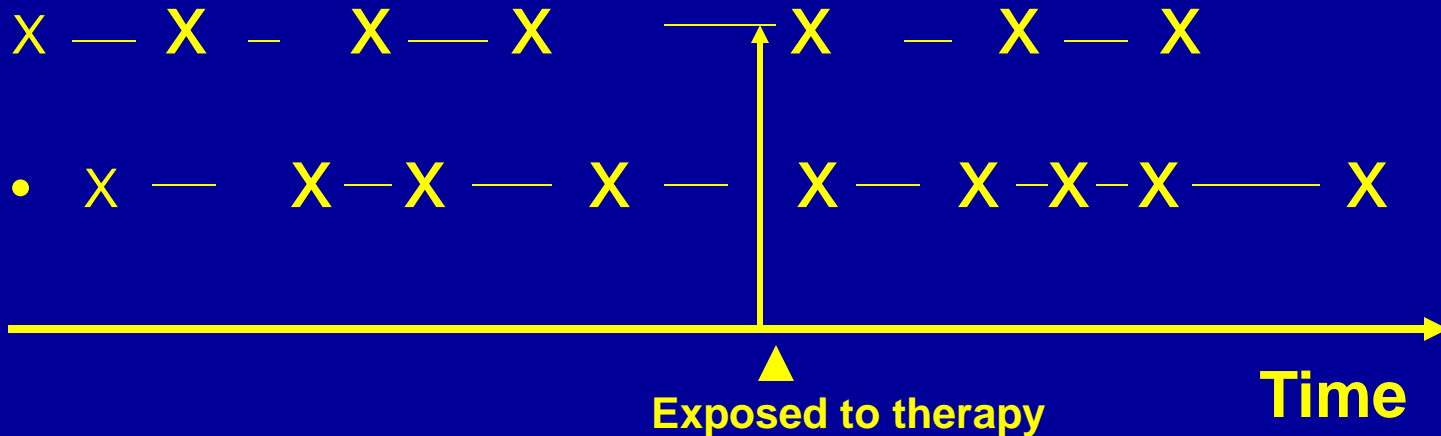
## Idealized process



Duration of S Phase



Duration of intermitotic times (Cell cycle time)



# **3. FAMILY HISTORY AND DISEASE**

**Some chronic diseases have been found to be associated with family history of disease.**

**Individuals having family members with disease may be at higher risk for disease.**

**Examples: Prostate Cancer, Breast Cancer, BRCA-1/BRCA-2 gene mutations.**



# Family History and Breast Cancer

- **Breast cancer risk is higher among women whose close blood relatives have this disease.**
- **“Having one first-degree relative (mother, sister, or daughter) with breast cancer approximately doubles a woman's risk. Having 2 first-degree relatives increases her risk about 5-fold”.**

# STUDY DESIGNS for EVALUATING FAMILY HISTORY

- Random Sample of Families from Population
- Case Control Study With Two Probands

# Study Design :Random Sample of Families from Population

- Carry out random telephone calls.

Determine if probability of having more than one disease case in family is larger than expected. If disease incidence is low, this method may not be feasible.

(Breast cancer: disease incidence is 80-100/100,000 per year depending on age). Most of phone calls will not have any disease in family.

# Study Design : Case control study

- **Collect pairs of observations in which one member of pair has disease and the other does not. These are called probands.**
- **Then determine if remaining family members have disease for each proband. Calculate observed proportion of having at least one of the remaining family members with disease.**

# Study Design: Case control study

The proband member of pair with disease can be recruited from hospital . (Referent registry)--- ( R=1)

The proband member without disease can be recruited by a random sample. --- (R=0)

# Some Calculations

Assume disease incidence is independent of family

$\theta = P\{\text{individual is incident with disease}\}$  (Assume  $\theta$  is very small)

$s = \text{No. of people in family with disease}$

$N = \text{Family Size}$ ,  $P\{N = n \mid R=0\} = p(n)$  for population (proband without disease)

$P\{S \geq 1 \mid N = n\} = 1 - (1 - \theta)^n \sim n\theta$  ignoring terms of order  $\theta^2$ .

$P\{S \geq 1, N = n\} \sim n\theta p(n)$ , ignoring terms of order  $\theta^2$

$P\{N = n \mid S \geq 1\} = n\theta p(n) / \theta \sum n p(n) = n p(n) / m$  ( $m = \text{mean family size}$ )

$P\{N = n \mid R=1\} = n p(n) / m$  where  $R=1$  indicates case control proband

Note :  $E[N \mid R=1] = E[N^2] / m = (\sigma^2 + m^2) / m = m [1 + C^2]$ ,  $C^2 = \sigma^2 / m^2$

$$P\{N=n \mid R=1\} = n p(n) / m$$



The distribution of family size for a referent registry is different than the distribution in the population

$$P\{N=n \mid R=0\} = p(n)$$

Assume that individuals in family are independent with respect to disease incidence. Then in the family sampled from referent registry having (n-1) remaining members has

- $P\{S \geq 2 \mid R=1, N=n\} = p\{S \geq 1 \mid R=1, N=n-1\} = 1 - (1-\theta)^{n-1} \sim (n-1)\theta \sim n\theta$
- $P\{S \geq 2, N=n \mid R=1\} \sim \theta n^2 p(n)/m$  and
- $P\{S \geq 2 \mid R=1\} \sim \theta \sum n^2 p(n)/m = \theta [E(N^2)]/m = \theta [m(1+C^2)]$
- Whereas if the population was sampled, taking into account one family member without disease,
- $P\{S \geq 1 \mid N=n, R=0\} \sim (n-1)\theta \sim \theta n$  and  $P\{S \geq 1 \mid R=0\} \sim \theta \sum n p(n) = \theta m$  Hence the ratio is
- $P\{S \geq 2 \mid R=1\} / P\{S \geq 1 \mid R=0\} \sim 1 + C^2$
- Thus assuming independence of family members having disease, the calculations show a bias in which, conditional on having one family member with disease, the probability of having additional family members with disease is greater than a randomly chosen family.

# Conclusions

**Note:** Larger families have a greater probability of having at least one member with disease than smaller families.

**Conclusion:** Referent registries tend to have larger size families than the general population.

- The probability of having more than one family member with disease is higher for families in referent registry compared to families in general population. This will be true even if family history is not a risk factor for disease.
- Referent registries yield biased results



# EXTENT OF BIAS IN FAMILIES FOR RELATIVE RISK

<u>Family</u>	<u>Bias</u> $(1 + C^2)$
Single gender families	1.75
Maternal line	1.5

## 4. Prevalence and Incidence

- Prevalence refers to how many people have disease at some point in time .
- Incidence refers to how many new cases of disease occurred in a defined time period.

Example : In the U.S. in 2007, 2.6 million women were alive with a diagnosis of breast cancer ; 192,000 new breast cancer cases

Since there are approximately 66,000,000 women over the age of 40, we can standardize these numbers by stating in this age group the prevalence is 39 per thousand ; incidence is 2.9 per thousand per year

# Prevalence and Incidence

- Prevalence may be indicated by  $P(t)$
- Incidence may be indicated by  $I(t)$
- $t$  may refer to chronological time or age.
- In probability terms,

Prevalence can refer to the probability of being in a disease state (and alive).

Incidence refers to the probability of a transition from no disease to disease over a given period of time.

# Models: Stable and Non-stable Disease

Stable disease model :  $P'(t) = 0$

Implies  $P(t) = P$  ,  $I(t) = I$

$P = I m$  where  $m$  is mean time with disease.

Breast cancer :  $m = P/I = 39 / 2.9 = 13.5$  years

Non-stable disease model :  $P'(t) \neq 0$

Prevalence and incidence are functions of time.

Note : If  $t$  refers to chronological time, the stable disease model may be true for some diseases. Origin can be chosen such that  $P(0) = P$

# Prevalence , Incidence and Mortality Models

**$I(t)$  : Incidence of disease at time  $t$ . ( $t$  may refer to age or chronological time).**

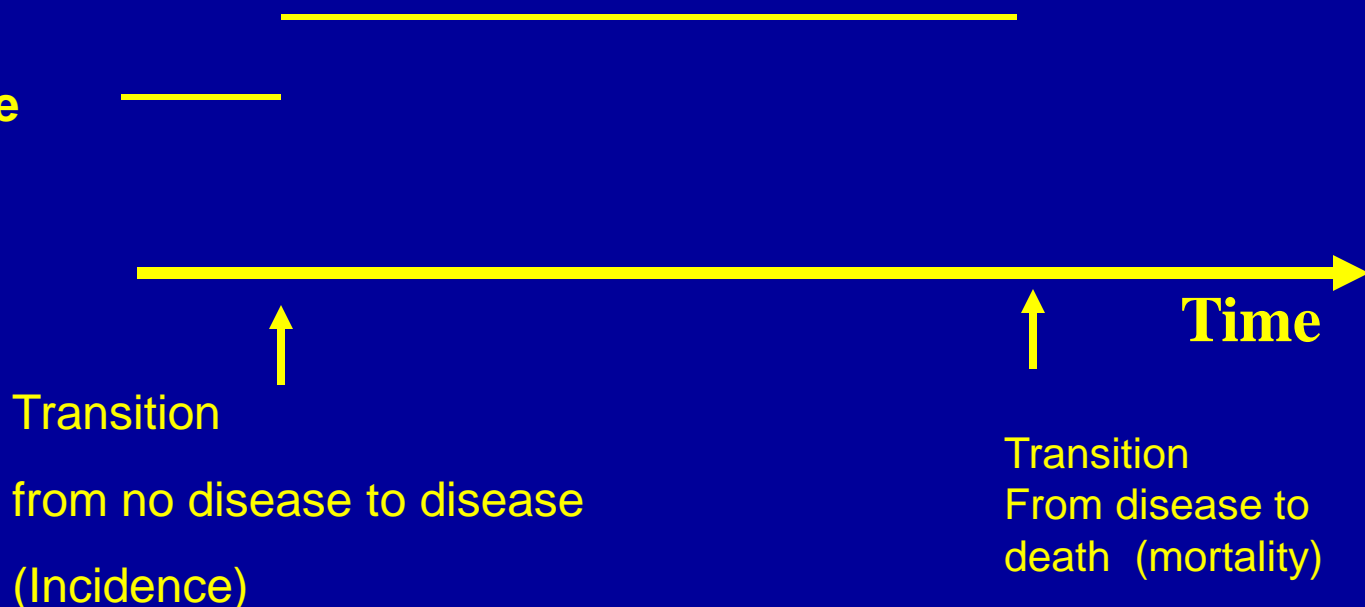
**$M(t)$  : Mortality of disease at time  $t$**

**$P(t)$  : Prevalence of disease at time  $t$  ( Probability of having disease at time  $t$ )**

**What are relations between prevalence , mortality and incidence? Can the numbers estimate how long people live with disease?**

**Disease state**

**Disease free  
state**

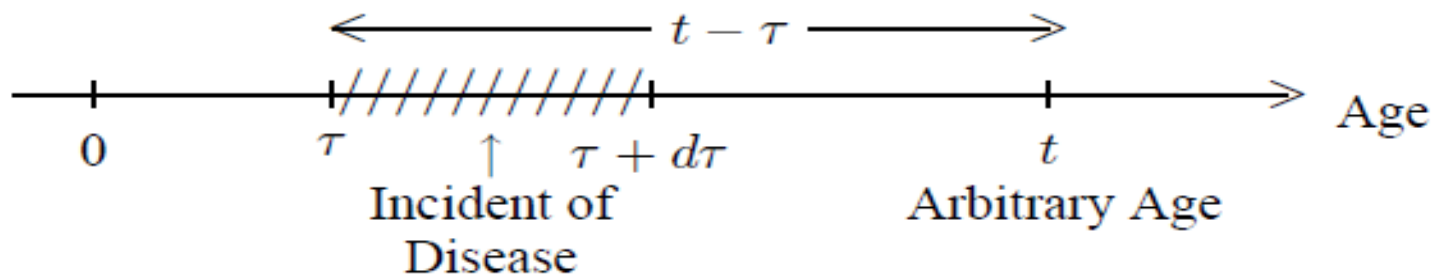


# Modeling Prevalence and Incidence: Stable Disease Model

Define

$$q(t) : \text{ pdf of disease survival}$$

$$Q(t) = \int_t^{\infty} q(x) dx = P\{T > t\}$$



$$P(t) = \int_0^t I(\tau) Q(t - \tau) d\tau$$

**$I(t) = I$  and  $t \rightarrow \infty$ , then  $P = mI$**

where  $m$  is the expected value of  $T$   
(survival)

# Non –Stable Disease Model

Suppose  $P(t)$  depends on  $t$ . Then it can be shown that approximately

$$P(t) \sim m I(t) \left\{ 1 - \frac{1}{2} m (C^2 - 1) I'(t)/I(t) \right\}$$

where  $C = \sigma/m$ .

If the time with disease follows the exponential distribution ( $C = 1$ ) and  $P(t) = m I(t)$  exactly.

Actually : The N & S condition for  $P(t) = m I(t)$  is that  $T$  follows an exponential distribution

# 5. Backwards and Forward Recurrence Times

**Backward and Forward Recurrence Times arise in many applications.**

**Examples :**

**Studying the natural history of a disease**

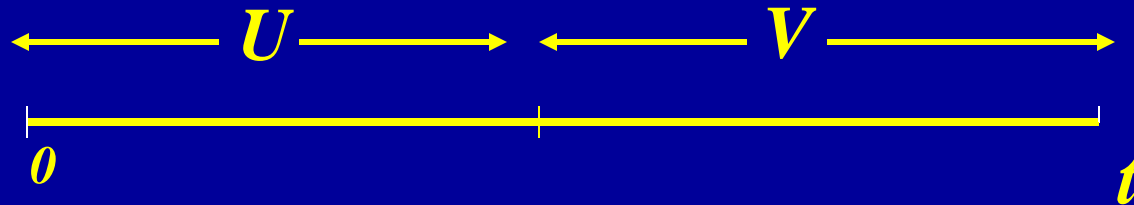
**Early detection of disease**

**Over diagnosis**



# Definition: Backward and Forward Recurrence Times

Let  $T$  be a non-negative random variable and  $T = t$  be a possible outcome. Suppose  $t$  is randomly divided into two parts:  $t = U + V$

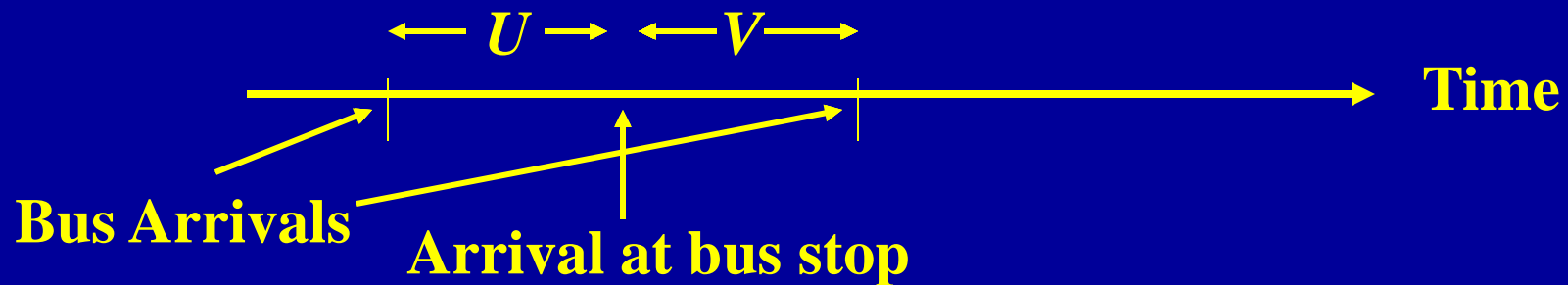


$U =$  Backward Recurrence Time

$V =$  Forward Recurrence Time

# Bus Waiting

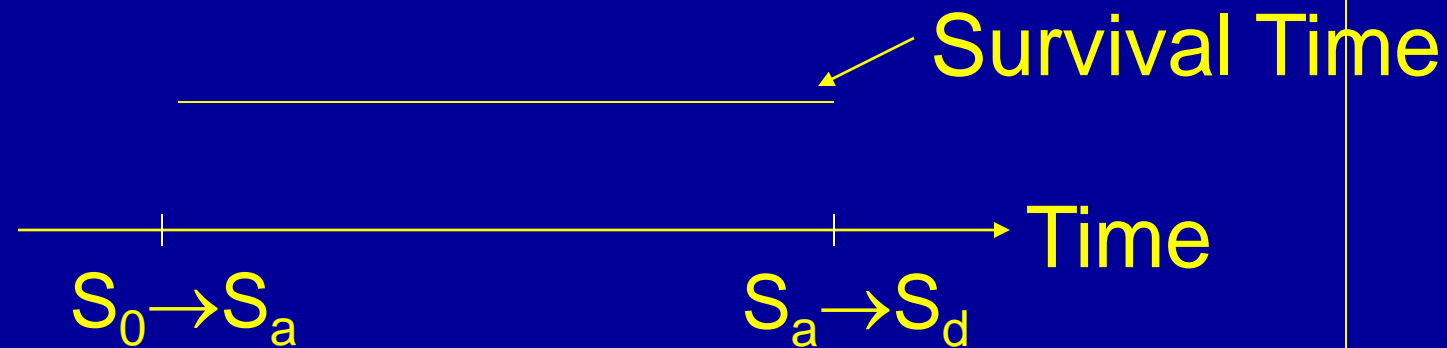
A person goes to a bus stop to take a bus. The backward recurrence time is how long the person missed the bus. The forward recurrence time is how long the person has to wait for the next bus to arrive



- If bus schedule does not change with time, distribution of  $U$  and  $V$  does not change with time.
- If bus schedule changes with time, distribution of  $U$  and  $V$  will be time dependent.

# Application to Chronic Disease Modeling

No disease  $(S_0)$   $\longrightarrow$  Alive with Disease  $(S_a)$   $\longrightarrow$  Death  $(S_d)$

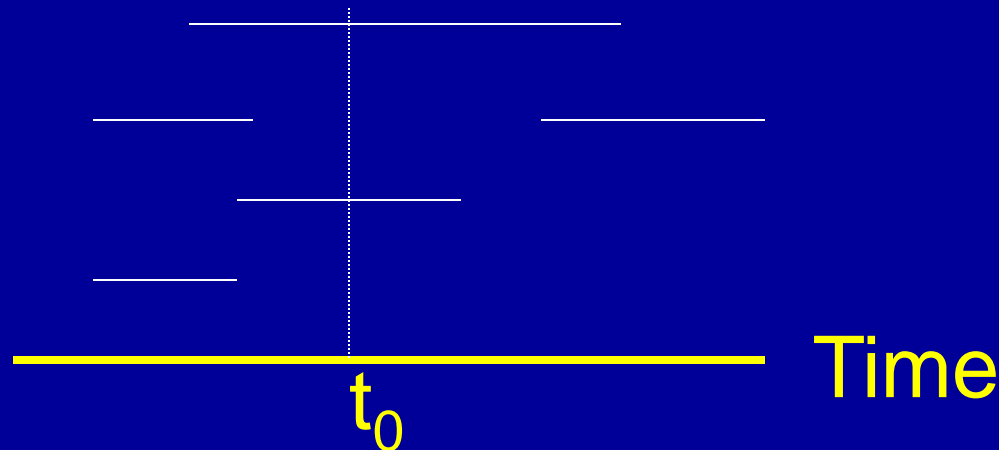


Time: Chronological Time or Age

Point Incidence:  $S_0 \rightarrow S_a$

Point Mortality:  $S_a \rightarrow S_d$

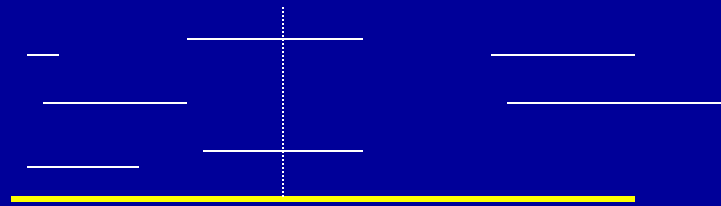
Suppose at chronologic time  $t_0$  the disease process is beginning to be observed. The process has been going on a long time.



**Some individuals will be in disease free state ( $S_0$ ), others will be alive and in disease state ( $S_a$ ) at time  $t_0$ .**

**Backwards Recurrence Time: How long a person has disease up to time  $t_0$ .**

**Forward Recurrence Time: Time to death with origin at time  $t_0$ .  
Only will be observed with additional follow-up time.**



$t_0$

### Length Biased Sampling

Individuals in  $S_a$  at time  $t_0$  are not a random sample of the population with disease. The longer they are alive, the more likely they will be in  $S_a$  at time  $t_0$ .

### Issues

1. Probability of being in  $S_a$  at time  $t_0$ ? Does proportion of people in  $S_a$  contain information on survival?
2. How long do people in  $S_a$  at time  $t_0$  live relative to people who become incident with disease after  $t_0$ ? (Forward Recurrence Time)
3. Total time in  $S_a$ ? (Survival)
4. Role of disease incidence?

# Theory: Length Biased Sampling

$a(t) = 1$  if time with disease is  $t$  in sample

$P\{ a(t) = 1 | T = t \} \sim t$  (Length biased sampling)

$P\{ a(t) = 1 | T = t \} = \text{independent of } t$  (random sampling)

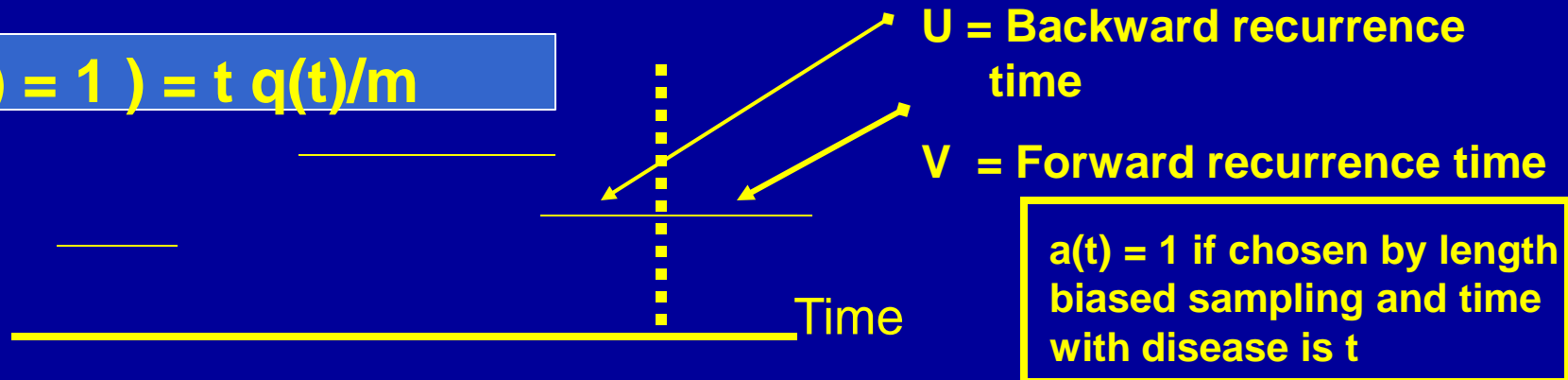
$P\{ a(t) = 1, t < T \leq t + dt \} \sim t q(t) dt$ ,  $T$  has pdf  $q(t)$

$P\{ t < T \leq t + dt | a(t) = 1 \} = t q(t) dt / m$ ,  $m = E(T)$

**Note:**  $E(T | a(t) = 1) = E(T^2)/m = m(1 + \sigma^2/m^2) > m$

# Theory: length biased sampling and recurrence times

$$f(t | a(t) = 1) = t q(t)/m$$



$T = U + V$  is time with disease having pdf  $q(t)$

## Forward recurrence time distribution:

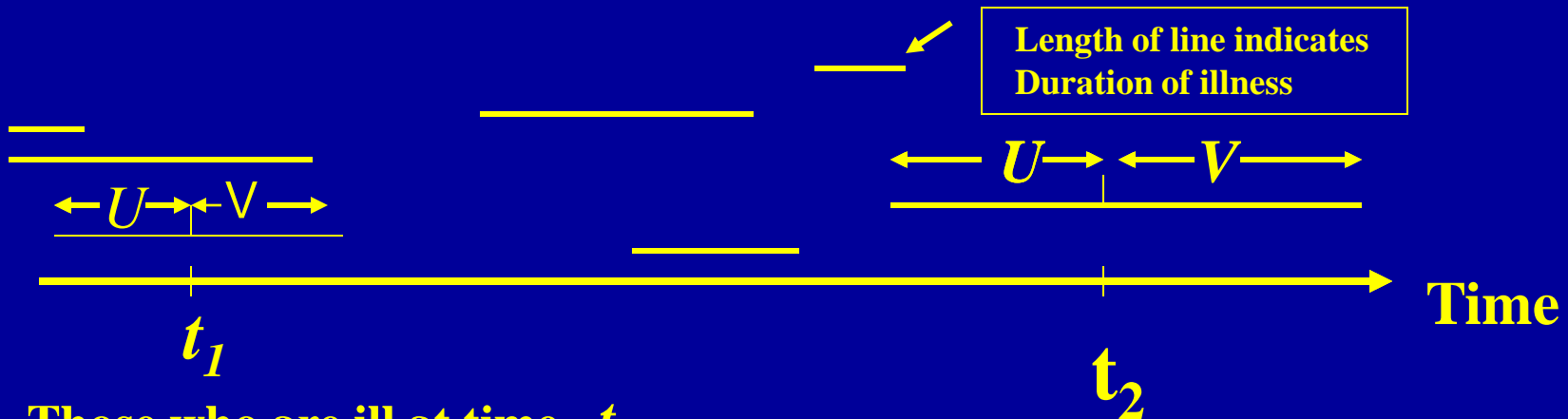
- $f(v | a(t)=1, T=t) = 1/t$  ,  $0 \leq v < t$  ,

$$f(v, T=t | a(t)=1) = (1/t) t q(t)/m = q(t)/m , 0 \leq v < t$$

$$f(v) = \int_v^{\infty} (q(t)/m) dt = Q(v)/m, \text{ where } Q(v) = \int_v^{\infty} q(t) dt$$

# Observational Study on a Population

Consider observing a population of individuals. Suppose the observational period begins at time  $t_1$  and ends at time  $t_2$ . Suppose the individuals in the population are sick or well. Those who are sick have a probability distribution for length of sickness period.



Those who are ill at time  $t_1$

$U$  = how long in sick state before being observed.

$V$  = length of time being observed beginning at  $t_1$ .

Those who are ill at time  $t_2$

$U$  = observed period of sick time. (Sick time is right censored)

$V$  = residual sick time after observations have ceased.

Distribution of  $U$  and  $V$  depend on whether incidence of sickness is independent or dependent on time



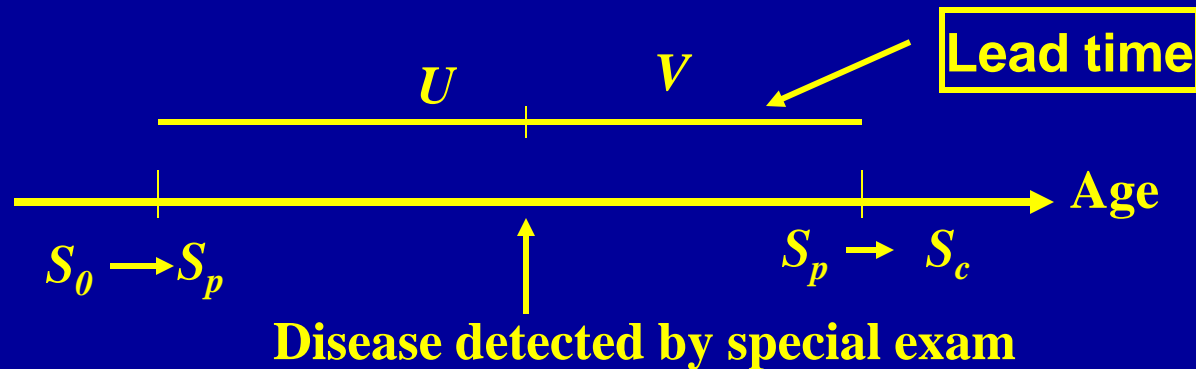
## Example: Early Detection of Disease

Suppose a person is in three states with respect to a particular disease ; i.e.

$S_0$ : Disease free state or if with disease cannot be detected.

$S_p$ : Pre-clinical state. Disease may be diagnosed with special examination only. No symptoms.

$S_c$ : Clinical state. Disease shows clinical symptoms and is diagnosed.

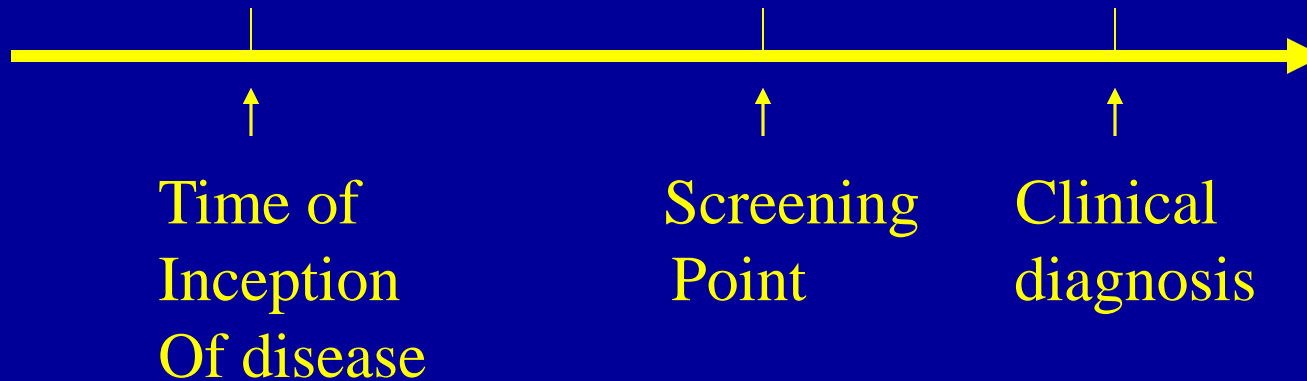
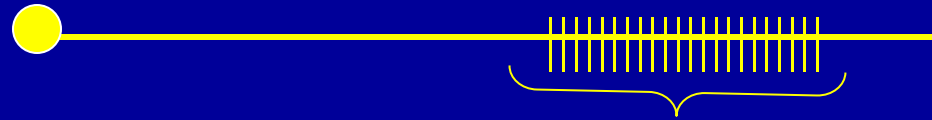


$U$ : How long person has had disease before diagnosis

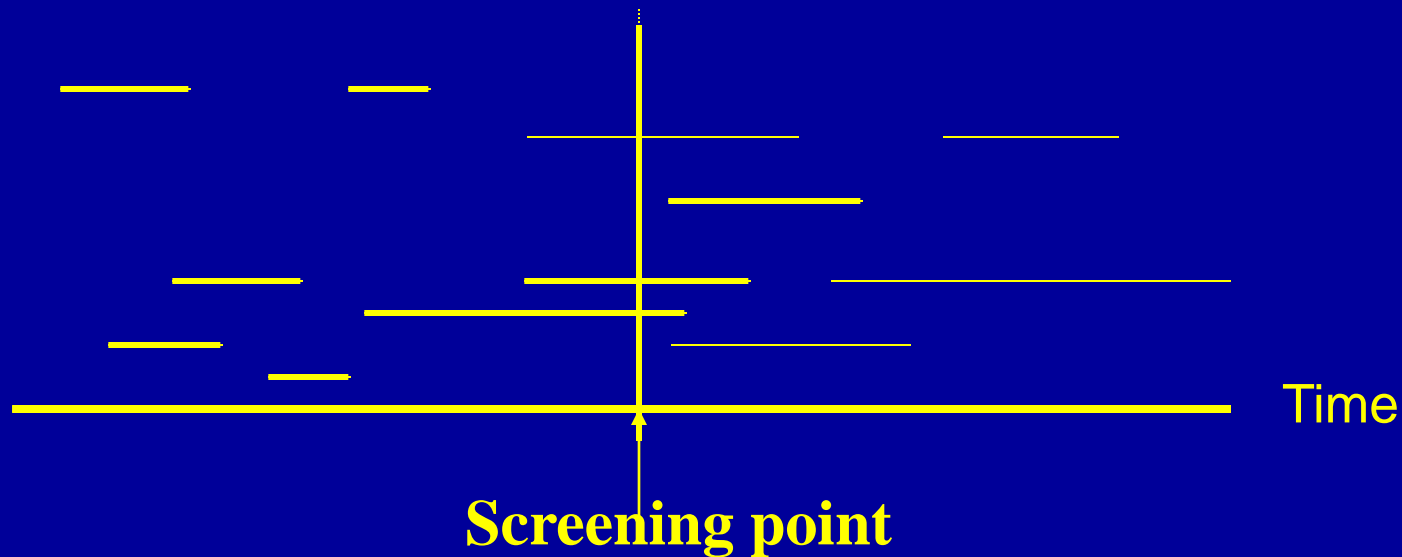
$V$ : Time gained by earlier diagnosis

If incidence of disease ( $S_0 \rightarrow S_p$ ) changes with age, then distribution of  $U$  and  $V$  will depend on age.

# Early detection of disease



**Consider a population of individuals who are screened**



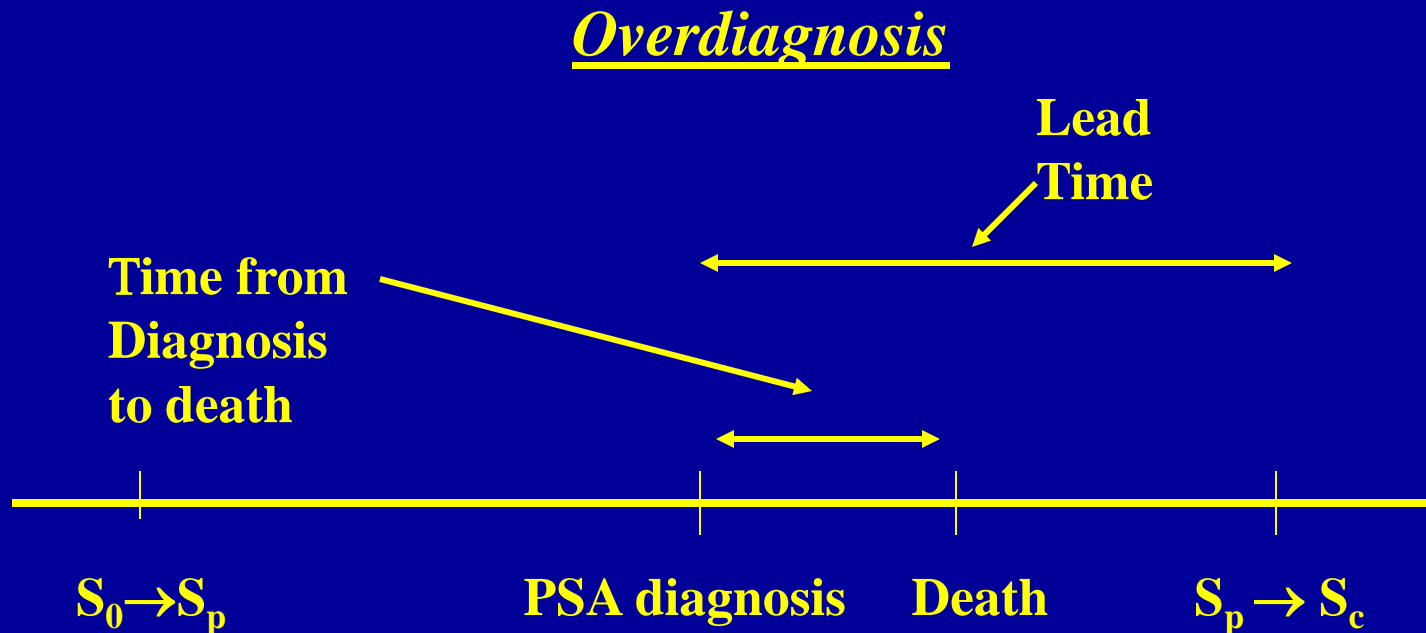
**Each horizontal line represent duration of pre-clinical disease for an individual**

**OBSERVE: Vertical dotted line has a higher probability of intersecting horizontal line; i.e.: the screening procedure finds those individuals having longer pre-clinical durations. This kind of sampling is called length biased sampling.**

## 2.9 Overdiagnosis: Prostate Cancer

Background: Prostate Specific Antigen (PSA) test is widely used to diagnose prostate cancer. Nearly all diagnosed cases by PSA are asymptomatic.

Question: Would the prostate cancer exhibit clinical symptoms during a man's lifetime? If no --- PSA diagnosis is an overdiagnosis

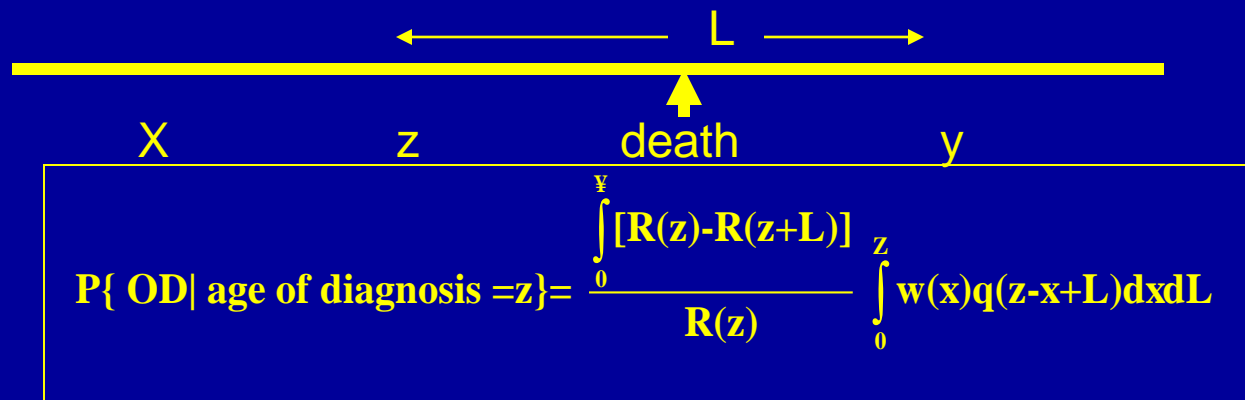


# Calculations for Over Diagnosis

- $S_0$  : disease free state,  $S_p$  : pre-clinical state,  $S_c$ : clinical state
- $w(x) = P\{ \text{transition } S_0 \rightarrow S_p \text{ at age } x\}$  ,  $y = \text{age entering clinical state}$
- $q(y) = \text{pdf in pre-clinical state}$
- $z = \text{age of diagnosis of disease by early detection program}$
- $P\{ T > t \} = R(t)$  where  $T = \text{survival without disease}$
- $L = y - z$  (Lead time)

## Sequence of events for over diagnosis conditional on being diagnosed at age $z$

- Enter  $S_p$  at age  $x$ , if alive enter  $S_c$  at age  $y$ ,
- Early diagnosis at age  $z$ , Die between  $(z, z+L)$  conditional on surviving at least to  $z$  where  $L = y - z$ .



# Comparisons of Probability of Over Diagnosis

<u>Age</u>	<u>Model</u>	<u>Prostate cancer Draisma</u>
<b>55</b>	<b>0.26</b>	<b>0.27</b>
<b>60</b>	<b>.33</b>	<b>.38</b>
<b>65</b>	<b>.40</b>	<b>.47</b>
<b>70</b>	<b>.48</b>	<b>.53</b>
<b>75</b>	<b>.57</b>	<b>.56</b>
<b>80</b>	<b>.65</b>	

Draisma et al. jnci (2003)—Based on European Trial; (sojourn time 15.4 yr, relevant cases),

Populations : model (U.S. ,all races), Others are white.



# **The Bernard G. Greenberg Lectures.**

## **Part II.**

### **The Early Detection of Disease**



# **Outline :Early detection of disease**

**1. Background, Motivation and Rationale**

**2. Early Detection Clinical Trials**

**3. The Early Detection Process**

**Models**

**Length and Lead time biases**

**Dynamics**

# Outline (continued)

## 4. Applications (breast cancer)

Mean lead time vs. disease stage

Comparisons of U.S. , U.K. and Nordic public health programs

U.S. Preventive Services Task Force recommendations

Risk based schedules vs. time based schedules

Maximum benefit from screening (?)

Is screening women under age 50 and over age 65 beneficial?

Screening high risk individuals

Over diagnosis (breast cancer)

## 5. Over diagnosis (prostate cancer)

# Motivation

- **This class of problems has been motivated by the belief that the early detection of disease may lead to higher cure rates and reduced mortality.**
- **Breast cancer has served as the disease to illustrate the ideas and apply the results.**
- **However the ideas are general and apply to many other cancers as well as possibly to other diseases.**

# Background and Rationale

- **Screening Programs** : Special exams to diagnose disease when it is asymptomatic.
- **Rationale** : Diagnosing and treating the disease early, before signs/symptoms appear, may result in more cures and lower mortality.
- **Caution** : If treatment is of little benefit there may be no reason to diagnose disease early.

# Examples of Screening Programs

- **Tuberculosis**
- **Diabetes**
- **Cancer**
  - **Breast Cancer**
  - **Cervical Cancer**
  - **Colorectal Cancer**
  - **Lung Cancer**
  - **Prostate Cancer**
- **Hypertension**
- **Coronary Artery Disease**
- **Thyroid Disease**
- **Osteoporosis**
- **HIV**

# Some Issues

- **Public Health Programs : Recommendations**

Initial age to begin screening, intervals between exams, schedules for high risk individuals. Recommendations should be made by risk status. --- How ?? Costs may be an important consideration.

- **Over diagnosis**

Disease is diagnosed early but would not be diagnosed by usual care in a person's lifetime --(Breast and prostate cancer)

- **Maximum benefit**

What is the maximum benefit that can be achieved from screening by searching for optimal exam schedules ????

# Example : Breast Cancer Screening Using Mammography

- **Breast Cancer-Clinical trials** The 8 randomized trials have mainly included women age 50-65. Do younger women ( under age 50 ) and older women (older than age 65) benefit from screening?
- The U.S. Preventive Service Task Force recommends that beginning at age 50 women have biennial exams. (2009)
- Nevertheless the **American Cancer Society** recommends that annual screening begin at age 40 for women at average risk.
- **United Kingdom** : The National Health Service offers screening beginning at age 50 with three year intervals for subsequent exams.
- **Nordic countries** : The recommendation is that screening begin at age 50 with two year intervals for subsequent exams.
- **What are the trade-offs for the different recommendations?**

# Scientific Evidence of Screening Benefit

- Diagnosing disease early does not necessarily result in benefit; e.g. diagnosing a primary cancer earlier may not be of benefit if the disease has already metastasized.
- If an effective treatment does not exist, there is no benefit in diagnosing disease early.
- The general consensus is that randomized clinical trials are the only way to evaluate screening programs for potential benefit.



# Early Detection Randomized Clinical Trials

- Typical trial consists of two groups . One group (control) receives usual care; the other group (study group) receives invitation to have a finite number of special examinations.
- Follow up for disease occurrence and death continues after the last exam.
- Endpoint is death from disease.
- Randomization may be carried out on an individual basis or by cluster randomization ;  
e. g. geographical region, physician practice.

# Early Detection vs. Therapeutic Trials

	Early Detection Trials	Therapeutic Trials
Eligibility	Subjects do not have disease	Subjects have disease
Follow-up	Long term follow-up may reduce power (optimal analysis time)	Long term follow-up always increases power
Endpoint	Mortality	Survival or related end point

**Statistical Problem : Design of Early Detection Clinical Trials. -- How many subjects, exams, exam spacing, follow up and optimal analysis time, etc.**

# Early Detection Clinical Trials

- Only subjects who are diagnosed with disease carry information about benefit.
- Trials need very large number of subjects
  - Relatively low incidence is characteristic of many chronic diseases; e.g. female breast cancer incidence is about 80-100 per 100,000 women per year depending on age.
  - Typical trial will require 10-20 years. During this time the technology for diagnosing disease may have changed.
    - Conclusions may be of limited interest.
- **Is it possible to carry out an early analysis, with limited follow up time?**

# Need for Models

- **Issues in the previous slide (optimal schedules, over diagnosis, benefit for women over age 65 and under age 50), high risk individuals, cannot be readily addressed by Randomized Clinical Trials.**
- **May not be feasible.**
- **Too many variables, takes too long, too costly , ethical concerns.**
- **Issues may be addressed by modeling the early detection process**

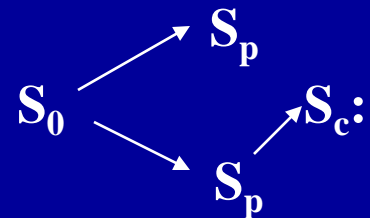
# Models

$S_0$ : **Disease free state** : Does not have disease or has disease which cannot be detected by exam.

$S_p$ : **Pre-clinical state**: Has disease but no signs or symptoms; capable of being detected by exam. Individual is asymptomatic.

$S_c$ : **Clinical state** : diagnosis by usual care.

$S_0 \rightarrow S_p \rightarrow S_c$  : **Progressive disease model (Breast cancer)**



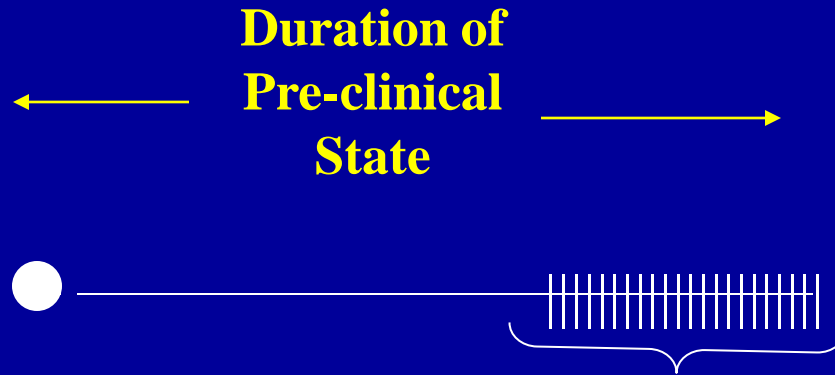
**Progressive disease model : subgroup never goes on to clinical disease -Prostate cancer**

$S_0 \leftrightarrow S_p \rightarrow S_c$  : **Non-progressive disease model (HPV ,Cervical cancer)**

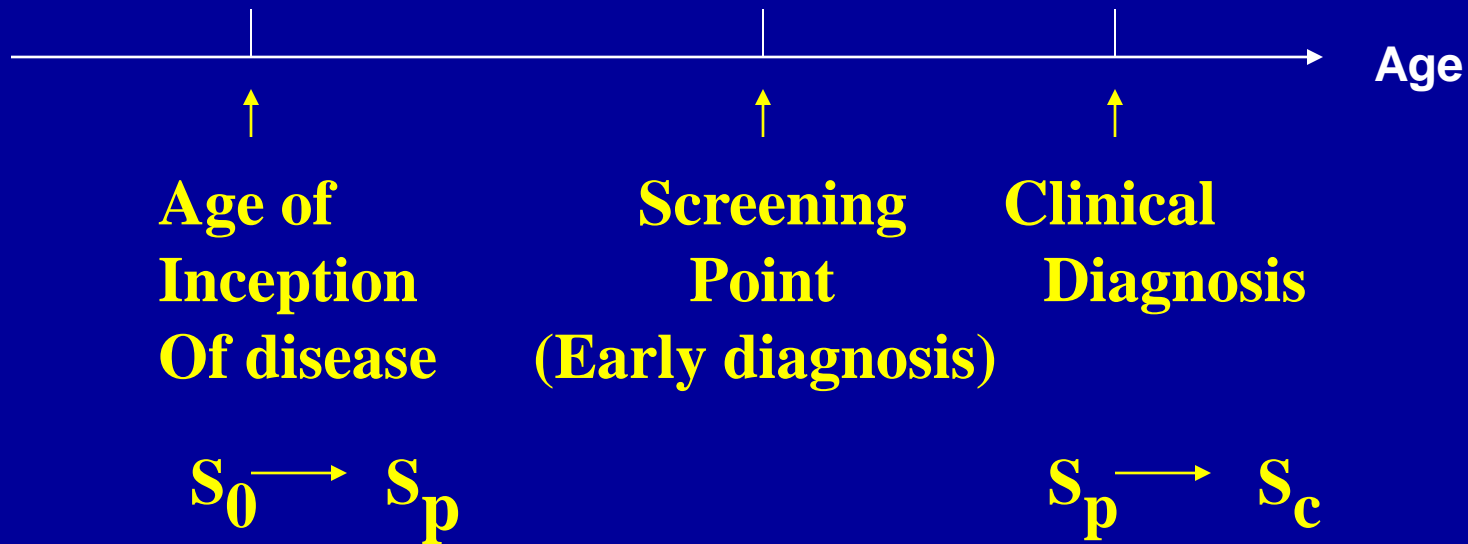
# Issues in the interpretation of data

- Suppose a group of subjects undergo screening for a particular disease and a number of subjects are diagnosed and treated.
- The subjects in this screened group have longer survival than a control group (no screening). Is this scientific evidence of the benefit of screening?
- No ---- Length biased sampling and lead time bias may introduce significant biases

# Natural History of Progressive Disease

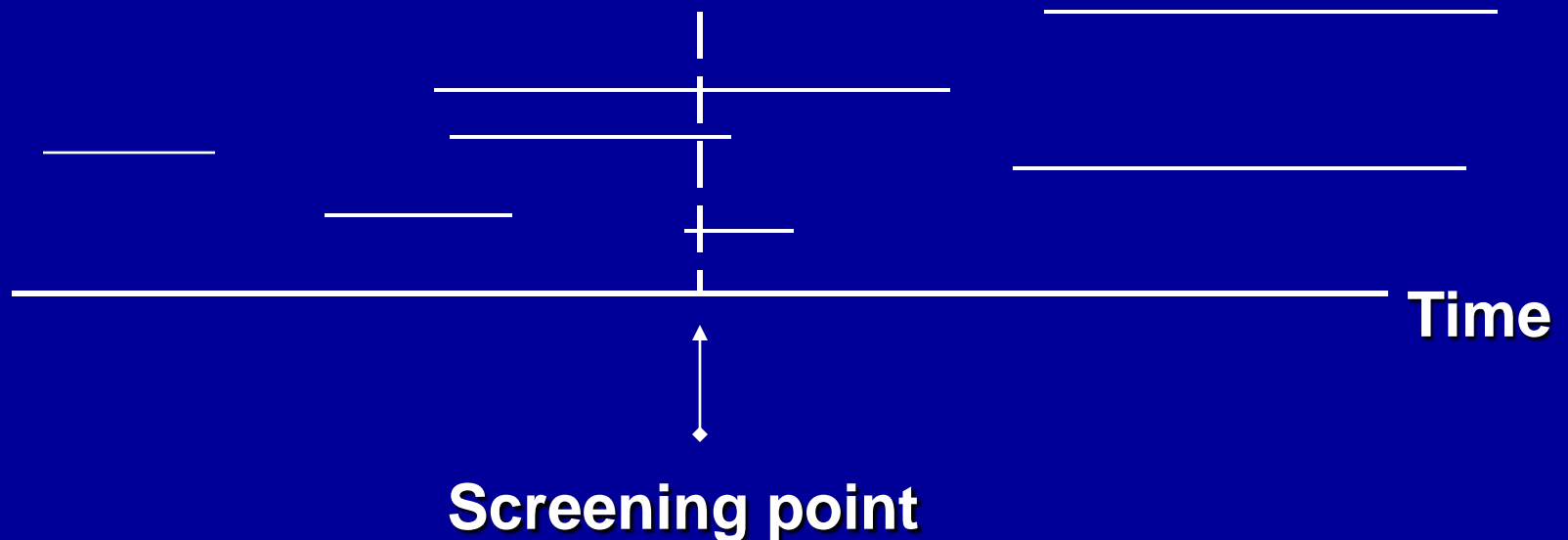


**S<sub>0</sub> = disease free**  
**S<sub>p</sub> = pre-clinical**  
**S<sub>c</sub> = clinical**



# Length biased sampling

Consider a population of cases



- Horizontal line : duration of time in pre-clinical state
- Diagnosis : equivalent to placing a random vertical line. Intersection represents case diagnosed.
- Vertical line is more likely to intersect longer horizontal lines ---people will tend to live longer.



# Why would screening result in benefit ?

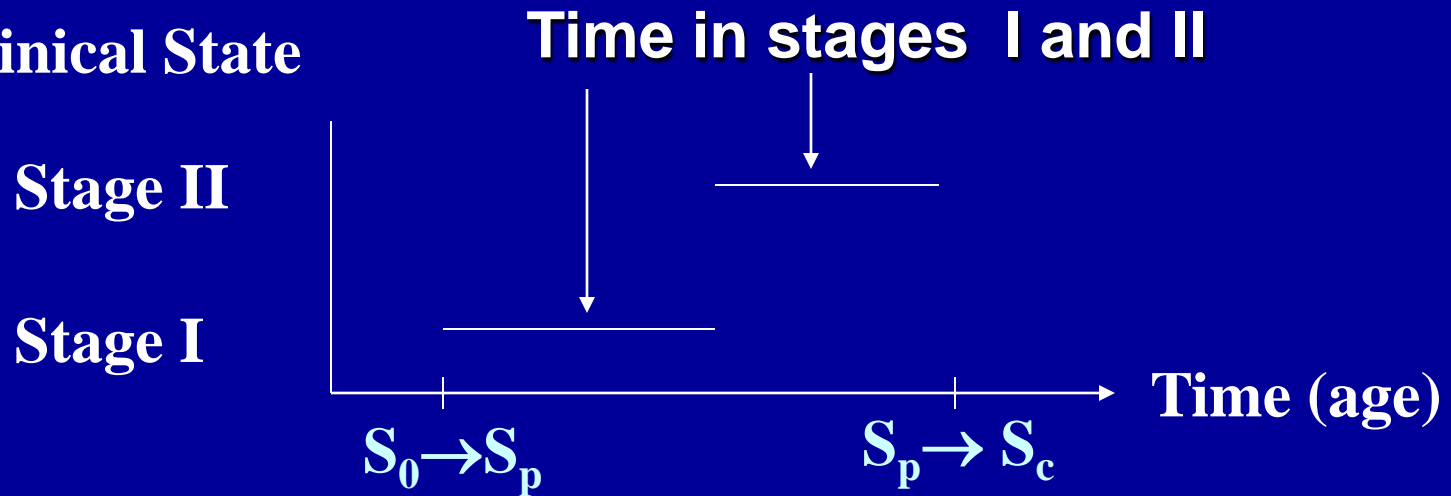
- If screen diagnosed cases are found in an earlier disease stage compared to usual care then there is likely to be benefit. This is referred to as a **stage shift**.
- Stage shift can be due to a **long lead time** ; i.e.cases are diagnosed before they transit to a more advanced prognostic stage.
- Stage shift may also arise from the **length biased sampling**. The selection of cases by screening may also be associated with earlier prognostic stages.

# Natural History of Disease

$S_0$ : Disease Free State or Cannot Be Detected

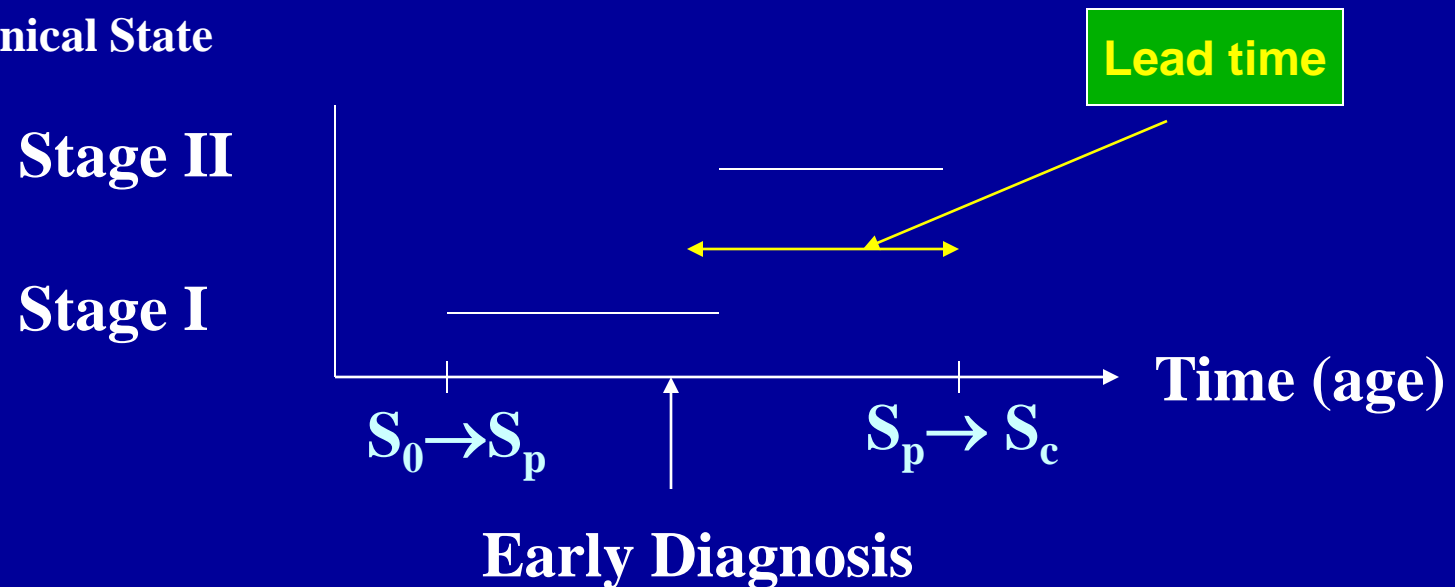
$S_p$ : Pre-clinical State

$S_c$ : Clinical State



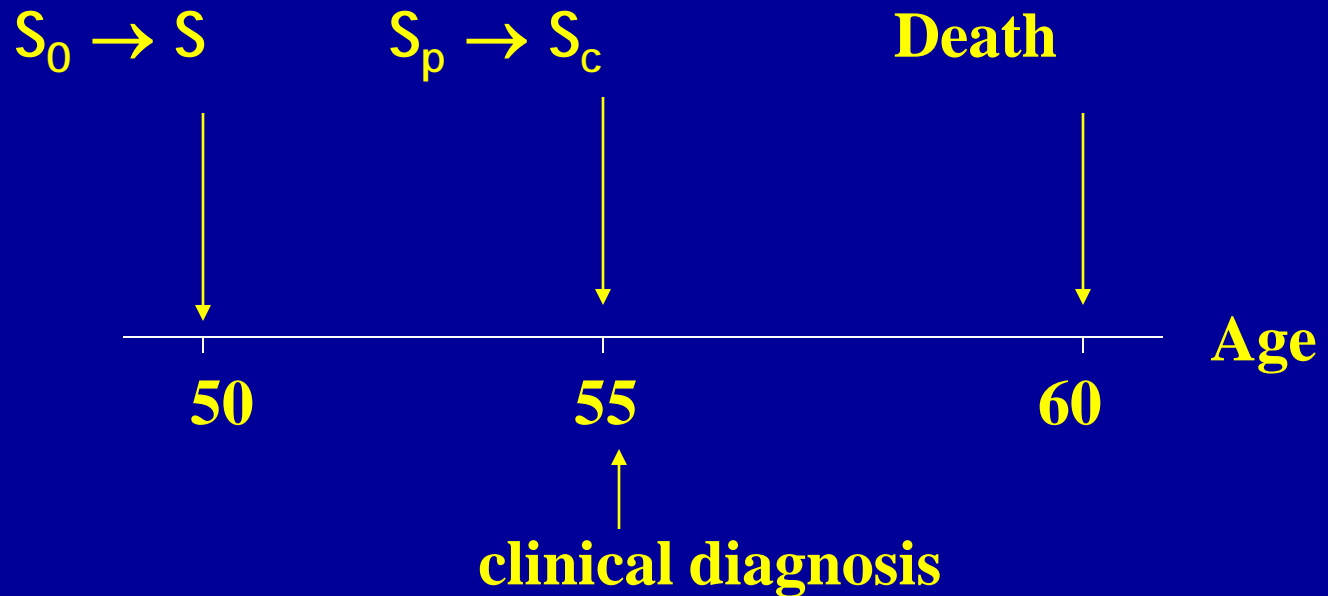
# Stage Shift and Earlier Diagnosis

- $S_0$ : Disease Free State or Cannot Be Detected
- $S_p$ : Pre-clinical State
- $S_c$ : Clinical State



**Note : the longer the mean lead time the greater the probability of diagnosing disease in an earlier stage.**

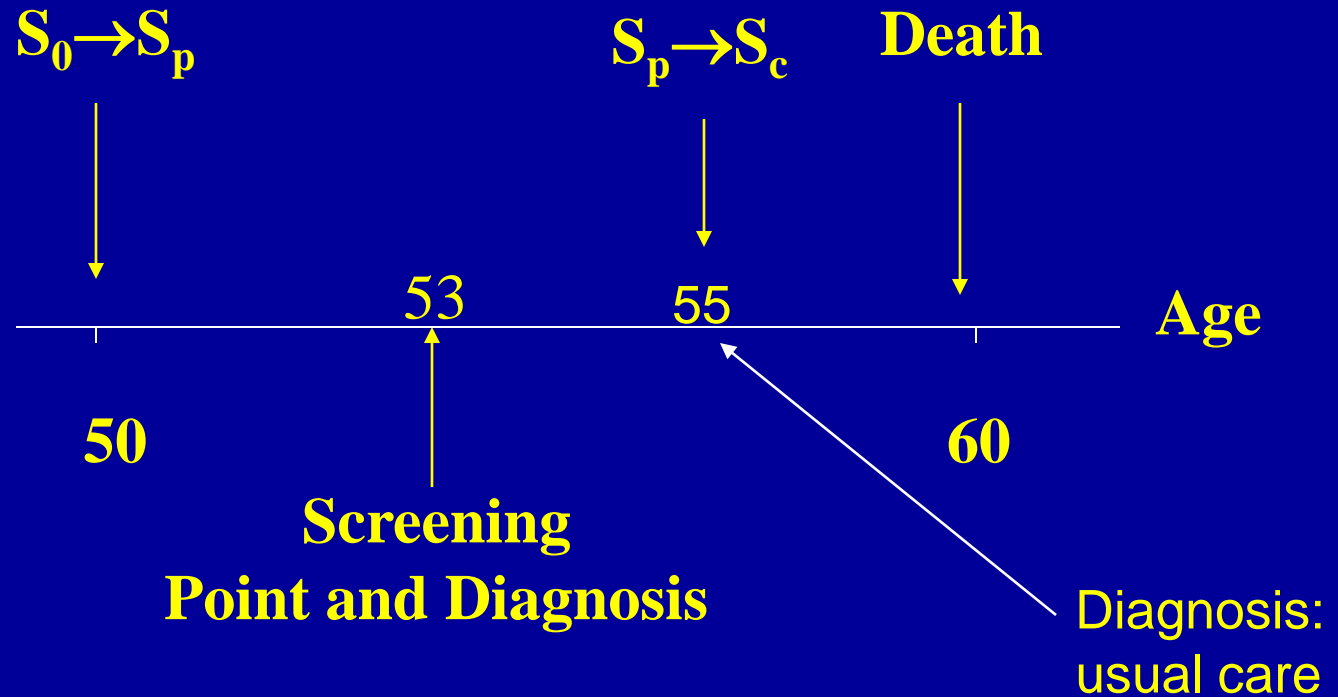
# Lead Time Bias : Usual care



**Survival from Clinical Diagnosis =  $60 - 55 = 5$  Years**

- $S_0$  = disease free state,  $S_p$  = pre-clinical state
- $S_c$  = clinical state

# Early Detection But Survival Is Not Enhanced



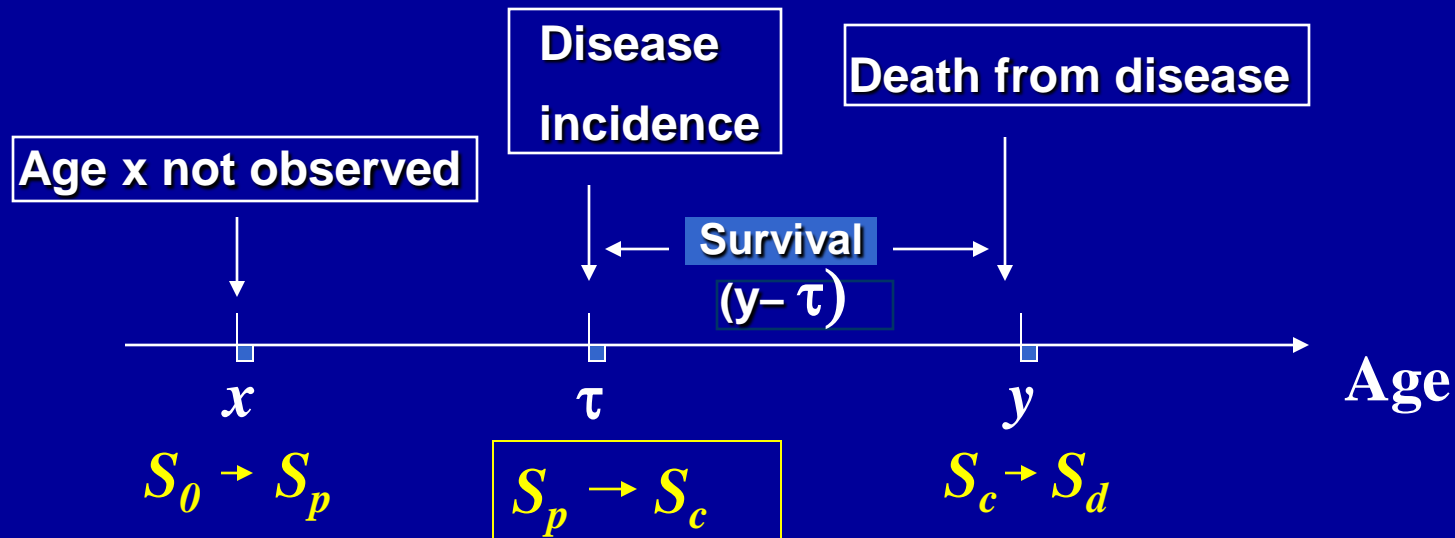
**Survival from Screening Diagnosis  $60 - 53 = 7$  Years**

**Survival (with usual care diagnosis)  $60 - 55 = 5$  years**

# Dynamics of the Natural History (1) : Usual care

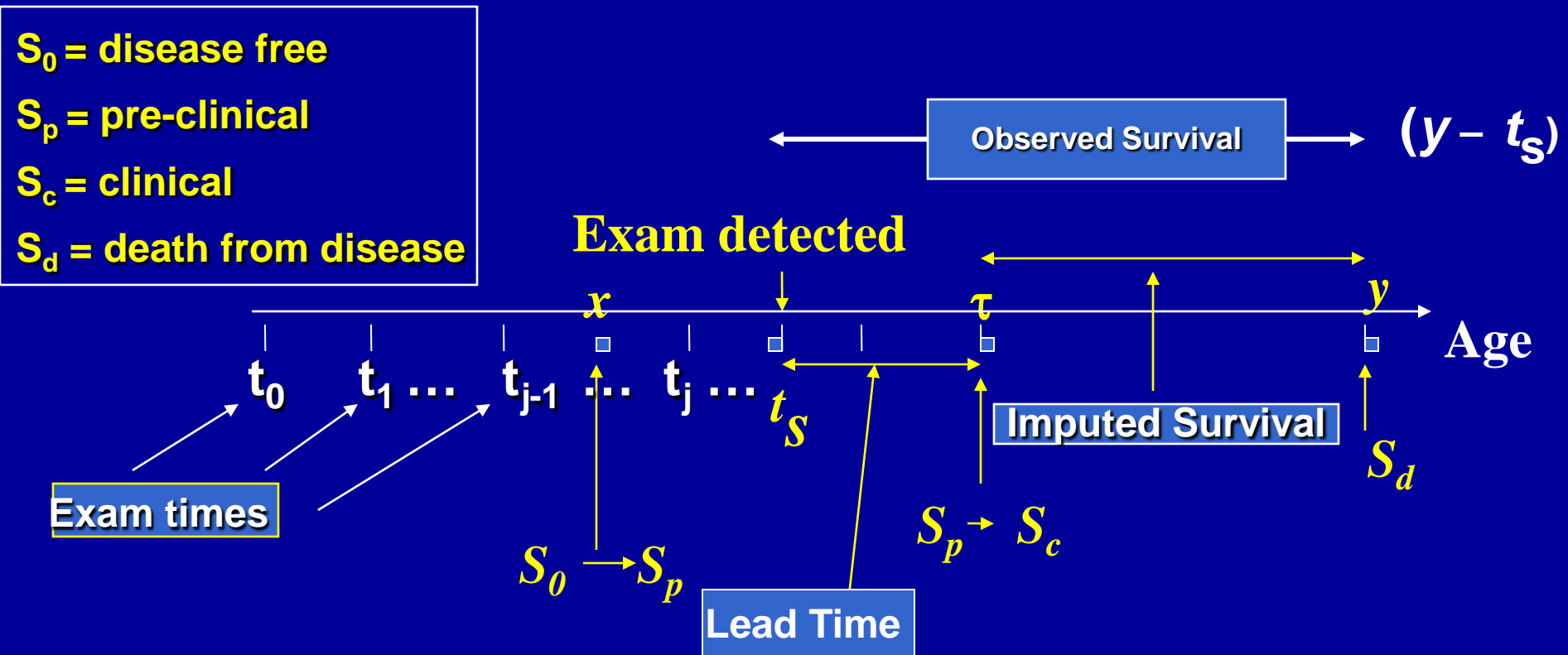
## Disease States

- $S_0$  : **Disease free state** – disease free or disease state which cannot be detected
- $S_p$  : **Pre-clinical state** - asymptomatic with no signs/symptoms
- $S_c$  : **Clinical state** – when diagnosed by routine methods
- $S_d$  : **Death state** (death due to disease)



Usual care : disease is diagnosed and treated at  $\tau$ .

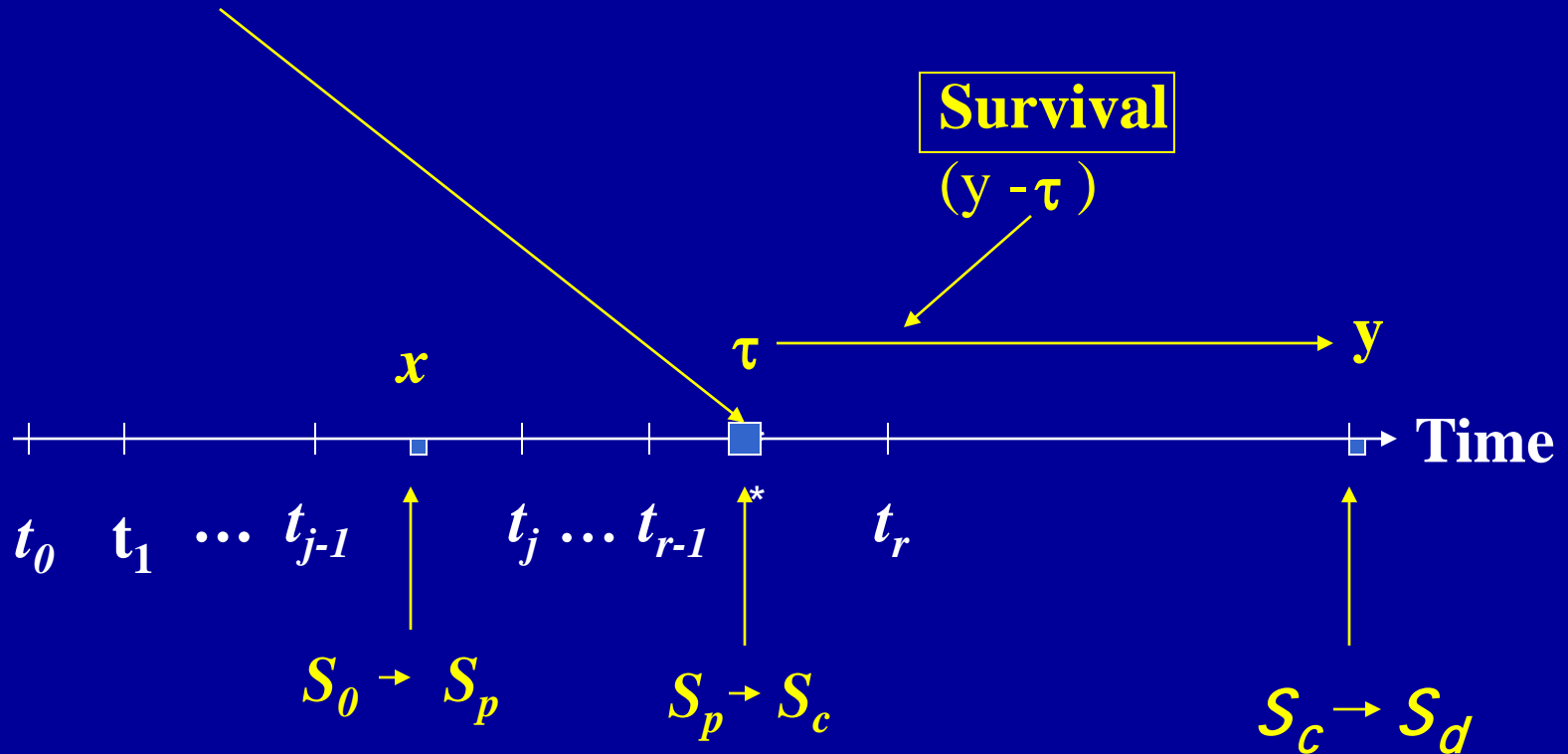
## Dynamics with Screening (2) : Exam detected case at $t_s$



- Ages  $\tau$  and  $x$  are not observed.
- Treatment begins at  $t_s$
- Observed survival time  $(y - t_s)$
- $(\tau - t_s)$  is lead time.
- Imputed survival = Survival with origin  $\tau = (\text{observed survival}) - (\text{lead time})$
- There may be a number (unknown) of false negative exams

# Dynamics with Screening (3)

Interval Case : Case diagnosed between  $t_{r-1}$  and  $t_r$



Exams at  $t_0 < t_1 < \dots < t_{r-1}$



# Applications

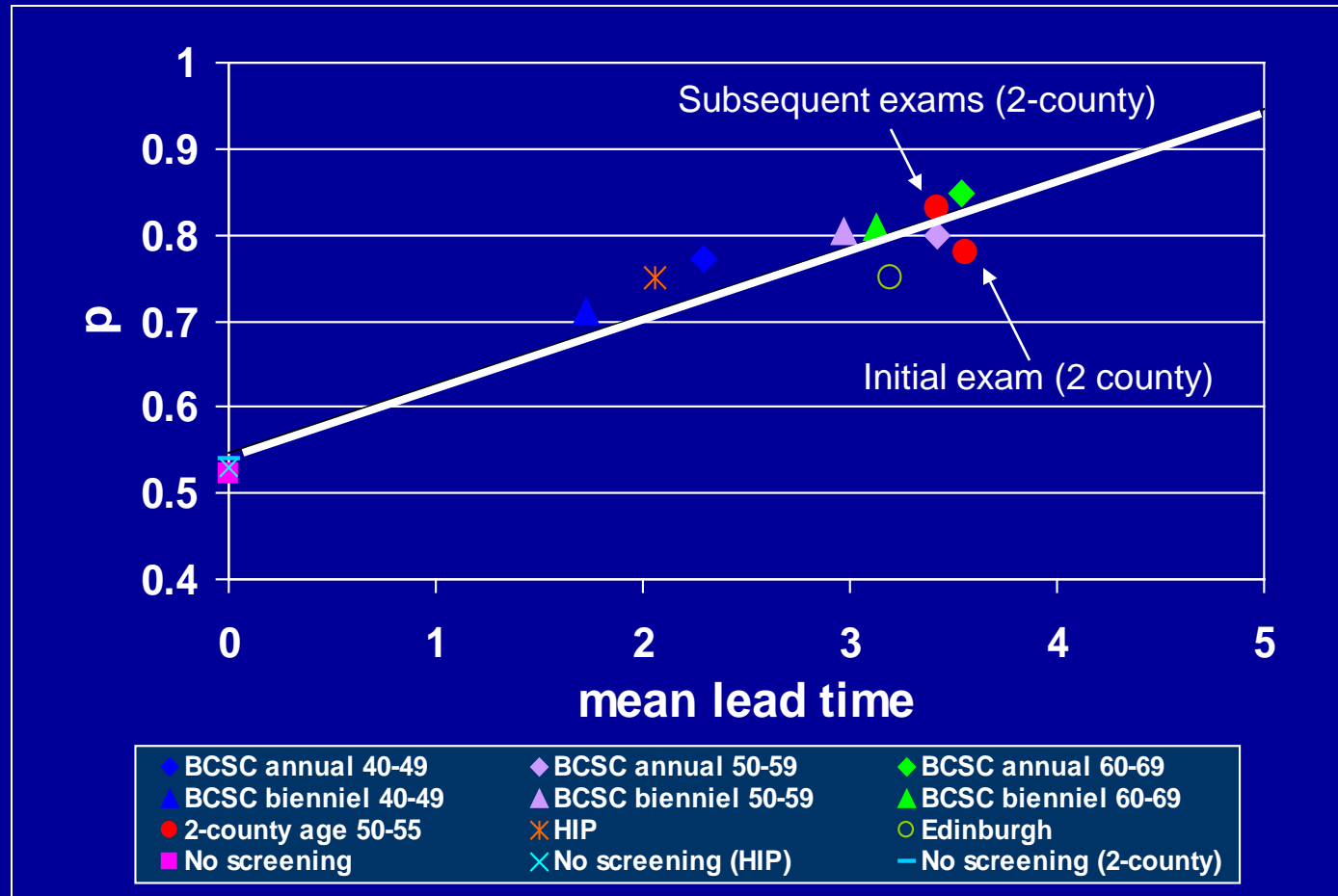
## Breast Cancer Screening (Mammography)

- **Public Health Programs (Breast cancer )**
- **Relation between mean lead time and proportion of negative nodes.**
  
- **Comparison of U.S., U.K. and Nordic countries**
- **Choosing screening intervals according to risk.**
- **Maximum Potential benefit**
- **High Risk Women**
- **Do women under 40 and/or over 65 benefit from screening ?**
- **Over diagnosis : breast and prostate**

# Relation of Mean Lead Time and Proportion of Negative Nodes

- Most important prognostic variable is whether disease has spread to lymph nodes ; nodes are negative if has not spread , positive otherwise.
- Distribution of Lead Time derived theoretically – function of exam schedule, risk status and sensitivity.
- Proportion of negative node diagnosed patients can be found from clinical trials and data bases.

## proportion of negative nodes vs mean lead time (2-county, BCSC, HIP, Edinburgh)

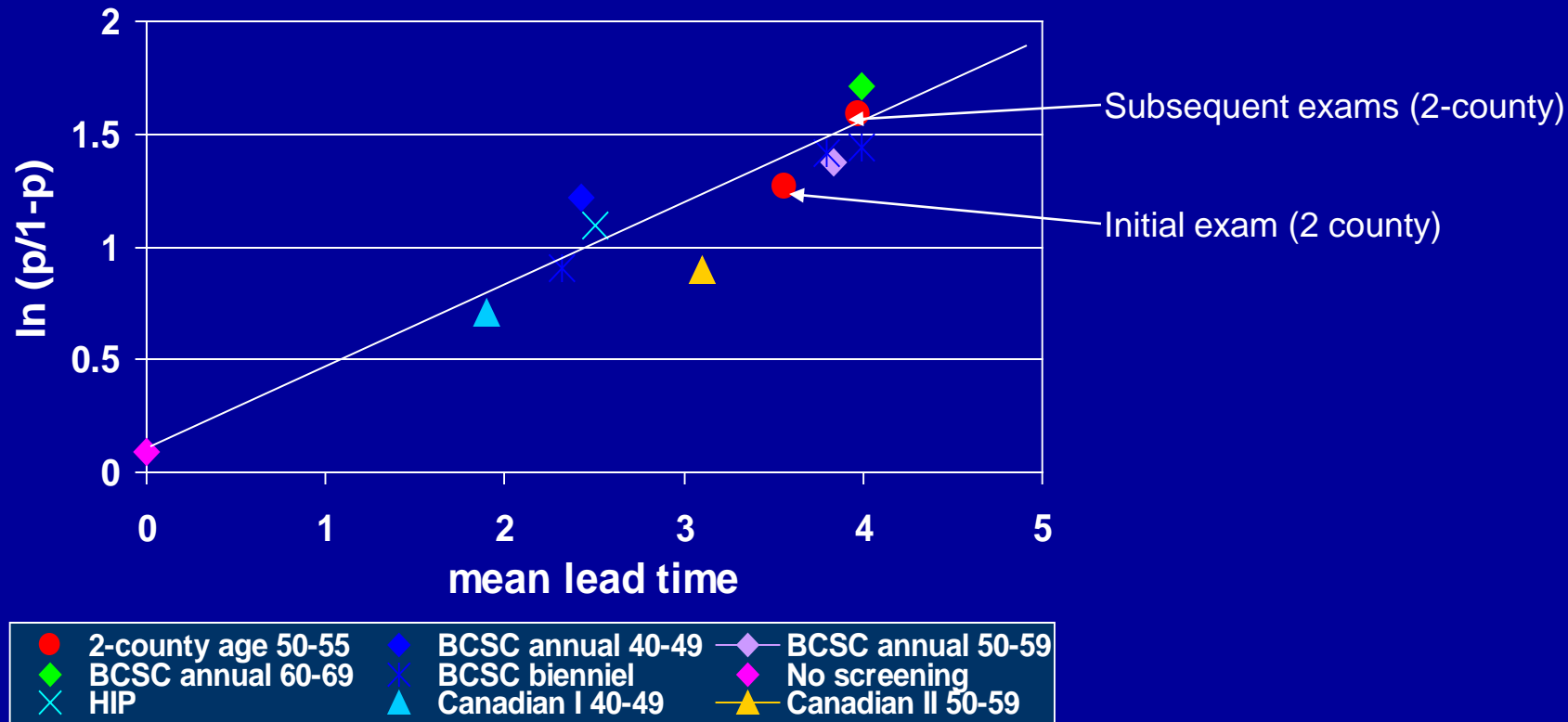


$p = (.080 \times \text{mean lead time}) + 0.546$ ,  $p = \text{proportion of negative nodes}$

Mean sojourn time varies by age from 2 years at age 40 to 4 years at age 50 and above.

Sensitivity varies from 0.7 at age 40 to 0.9 at age 50 and above.

## Ln (p/1-p) vs mean lead time (2-county, BCSC, HIP, Canadian I & II)



$$\text{Ln}(p/1-p) = (.35 \times \text{mean lead time}) + 0.111, \quad p = \text{proportion of negative nodes}$$

Mean sojourn time varies by age from 2 years at age 40 to 4 years at age 50 and above. Sensitivity varies from 0.7 at age 40 to 0.9 at age 50 and above.

# Value of plot : proportion of negative nodes vs. mean lead time

- **Clinical trials are necessary to evaluate benefit – mortality reduction.**
- **If stage distribution is known at diagnosis it is possible to estimate mortality reduction as a function of follow up time.**
- **Given : screening schedule, age distribution of population and sensitivity, the mean lead time may be estimated using theoretical distribution. . Hence from plot , the proportion of cases in each stage can be estimated.**
- **This allows an estimate of mortality reduction . – No need for clinical trial.**
- **Side issues : how does mortality reduction change with changes in sensitivity and scheduling.**
- **Methods are very general and may be used for other diseases which satisfy basic assumptions – natural history is progressive and benefit derives form stage shift.**

# Does length biased sampling of cases account for earlier stage disease ?

Compare stage of interval diagnosed cancers with control group. –Why?

The length biased process of early detection selects cases who tend to sojourn longer in pre-clinical state. If longer sojourn times are associated with earlier stage disease, then the interval cancers will tend to have shorter sojourn times and have more advanced staging.

If stage of interval cancers is same as control (no screening group), then length bias does not account for stage shift; i.e. stage shift is due to diagnosing the disease earlier before it transits to a more advanced stage.

# Summary on Stage Shift : Breast cancer

<u>Trial</u>	<u>Stage</u>	<u>Control</u>	<u>Study</u>	<u>Screen Detected</u>	<u>Interval Detected</u>
HIP	N+	52%	41%	30%	50%
2-County	II-IV	46%	30%	19%	45%
Malmo I *	“	59%	39%	23%	56%
Edinburgh	“	87%	64%	50%	72%

-----  
 \* II-IV : Stages II - IV ( AJCC)

- **The interval cancers tend to have the same stage as a control group.**
- **If stage shift was due to length biased sampling, the interval cancers would tend to have more advanced stage than a control group. This is not true.**

# THRESHOLD METHOD :

## Application to breast cancer

### Background

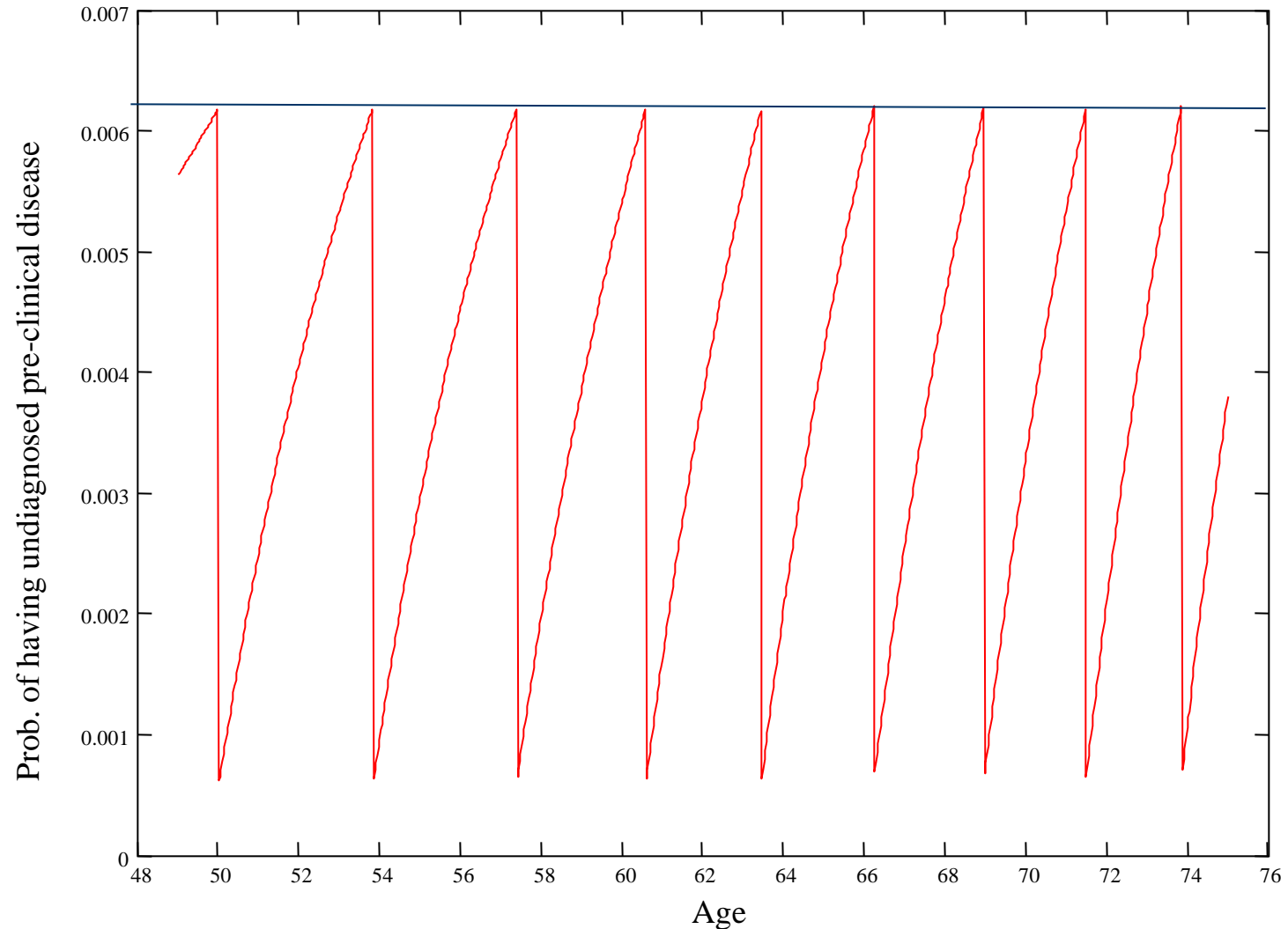
- There is general agreement that mammographic screening should begin at 50 years of age.
- Prevalence at 50 = 1.8 cases/ 1000
- Prevalence increases with age

First exam at age 50 . Choose subsequent exams so that independent of age, the risk will never be larger than at 50.--- 50, 54, 57, 60, 63, 66, 69, 72, 75, 77



# $S(t) = \text{Prob}(\text{Undiagnosed Pre-Clinical Disease})$

## Exam Schedule with $S(50) = .0062$ (Threshold Value)



# Mammogram Exam Schedules for Ages [50, 79]

	Threshold (staggered)	Annual	Every 2 yrs Nordic and Task Force	Every 3 yrs U.K.
<b>No. of Exams (threshold)*</b>	11 (.0062)	30	16	11
<b>Mortality Reduction (%)</b>	26%	37%	30%	26%

\*Threshold exam schedule corresponds to exam schedule whenever risk is the same as at age 50.

- **Annual: U.S.: ACS/NCI Recommendation**
- **Every 2 Years: U.S. Preventive Services Task Force and Nordic Recommendation**
- **Every 3 Years: U.K. Recommendation**
- \* **Mortality Reduction =**  

$$\frac{[\text{Mortality (controls)} - \text{Mortality (screened)}]}{[\text{Mortality Controls}]}$$

# Risk Based Screening

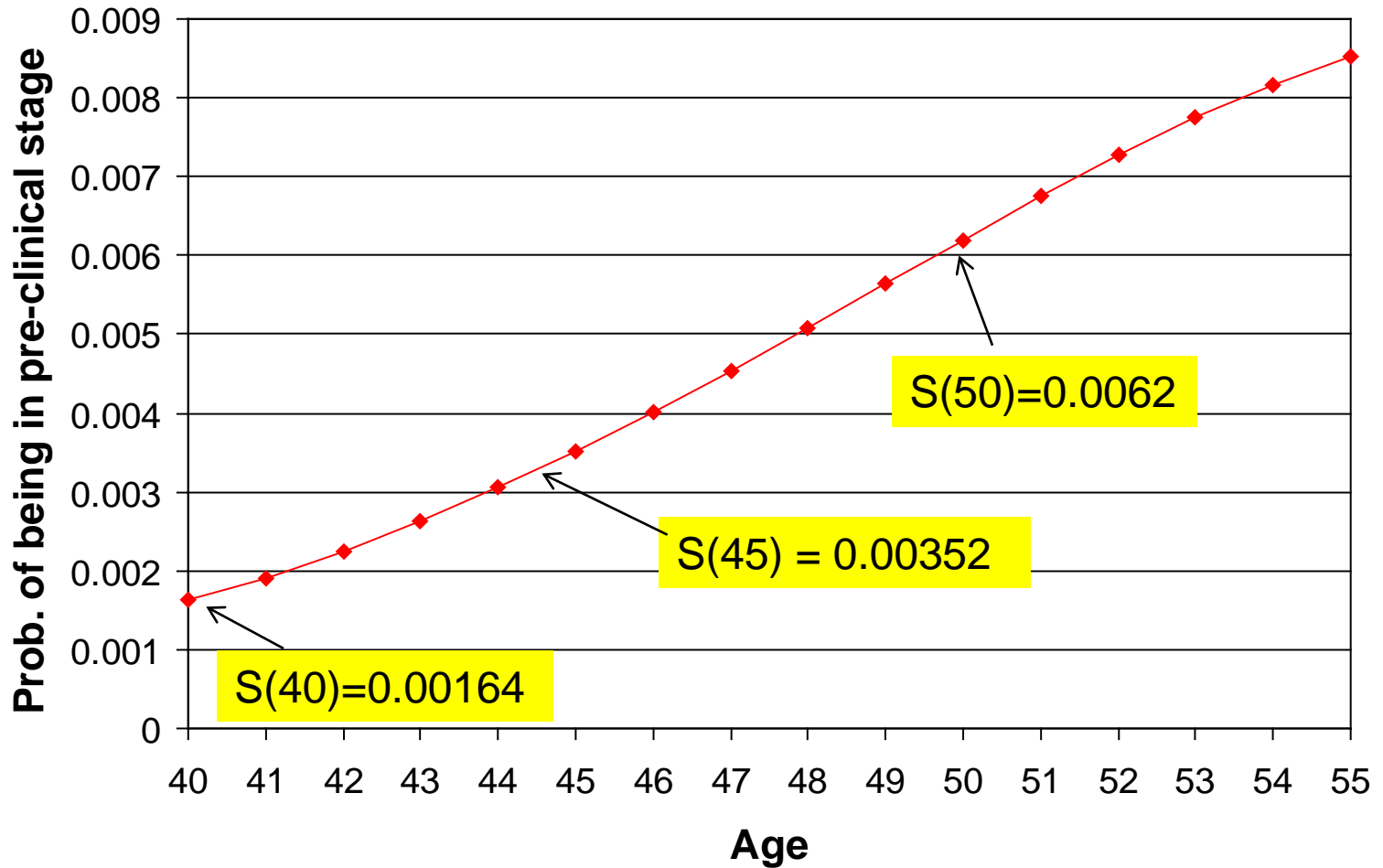
- Current recommendations for screening are based on time and initial age to initiate screening program.
- Since goal is to diagnose individual in pre-clinical state – define

**Risk = Probability individual is in pre-clinical state.**

# Choosing initial age

- Incidence of many chronic diseases increases with age. Hence risk increases with age.
- The choice of the initial age to begin screening assumes that the risk for lower ages is not high enough to warrant the beginning of a screening program.

# S(t): Probability of having undiagnosed pre-clinical disease at age t ( breast cancer)

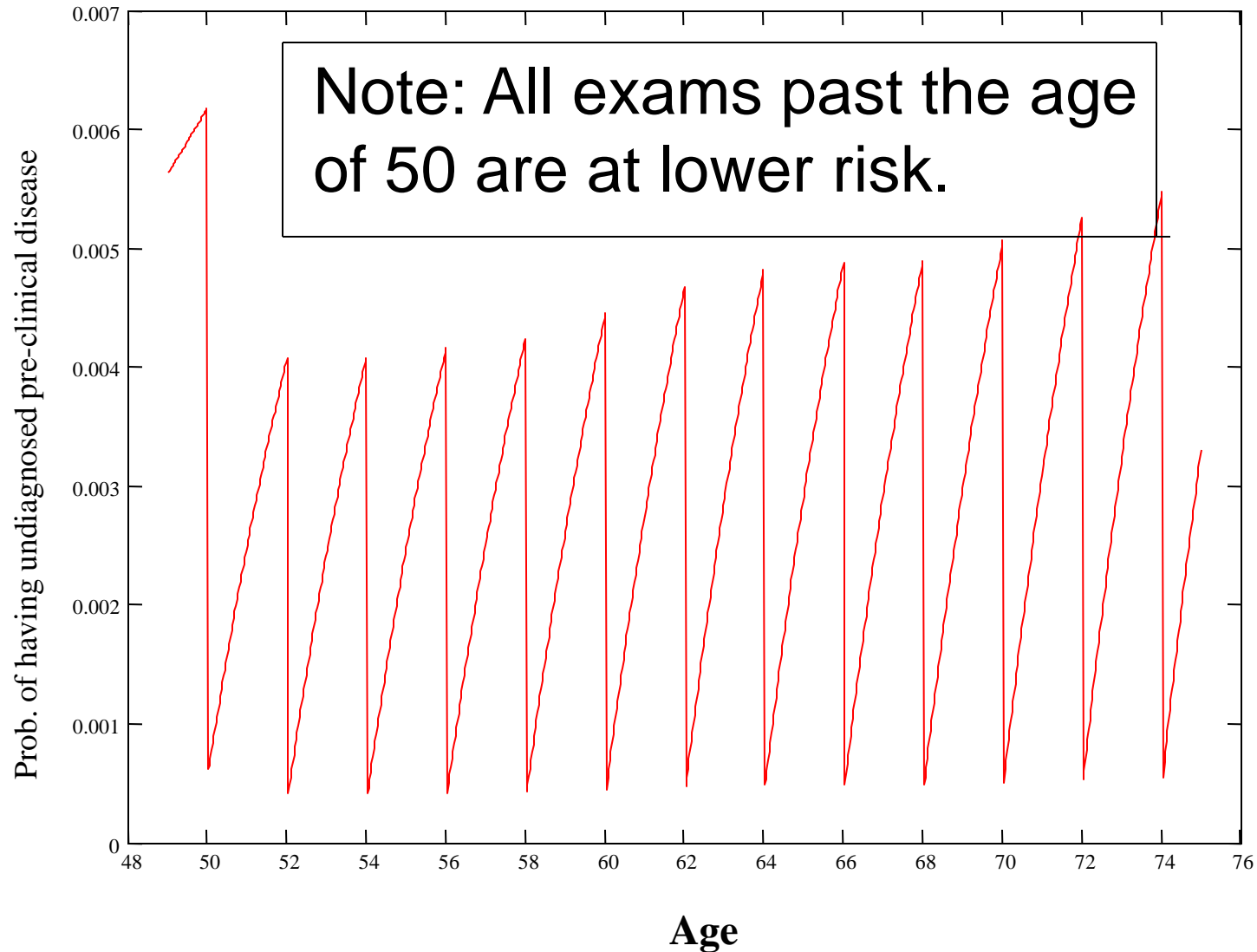


# U.S. Preventive Services Task Force (USPSTF) Recommendations for Breast Cancer Screening

- **November 2009** : Recommends biennial screening mammography only for women ages 50—74.
- **Resulted in much controversy.** ACS had long recommended annual screening beginning at age 40. **ACS opposed recommendation.**
- **December 2009** : USPSTF modified recommendations by stating decision to “start regular biennial screening before the age of 50 should be an individual one and take into account patient’s values regarding individual specific benefits and harms”.

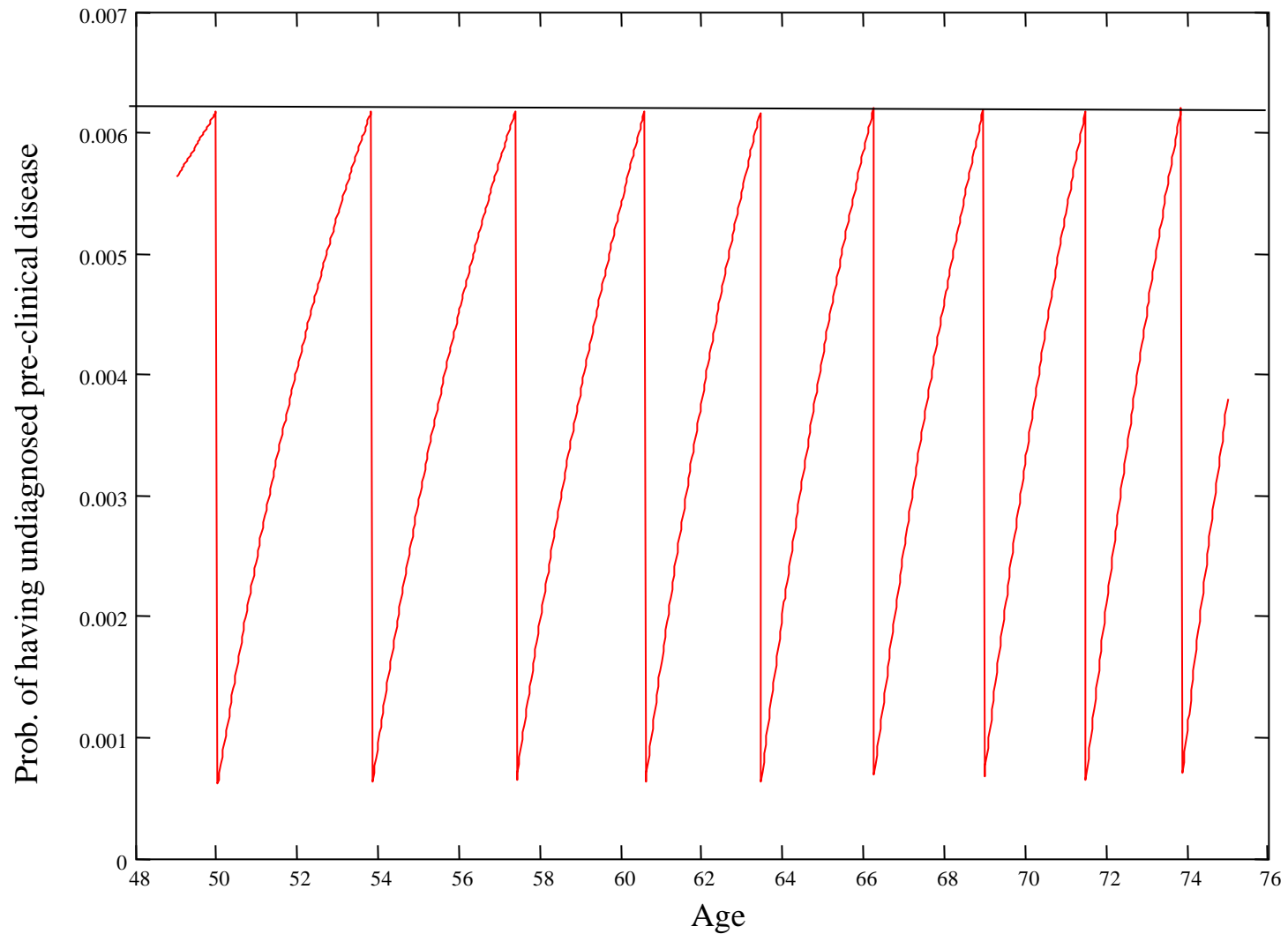
# $S(t) = \text{Prob}(\text{Undiagnosed Pre-clinical Disease})$

Following Biennial Exam Schedule with  $\beta(t) = .90$  for [50,74]



# **$S(t)=\text{Prob}(\text{Undiagnosed Pre-Clinical Disease})$**

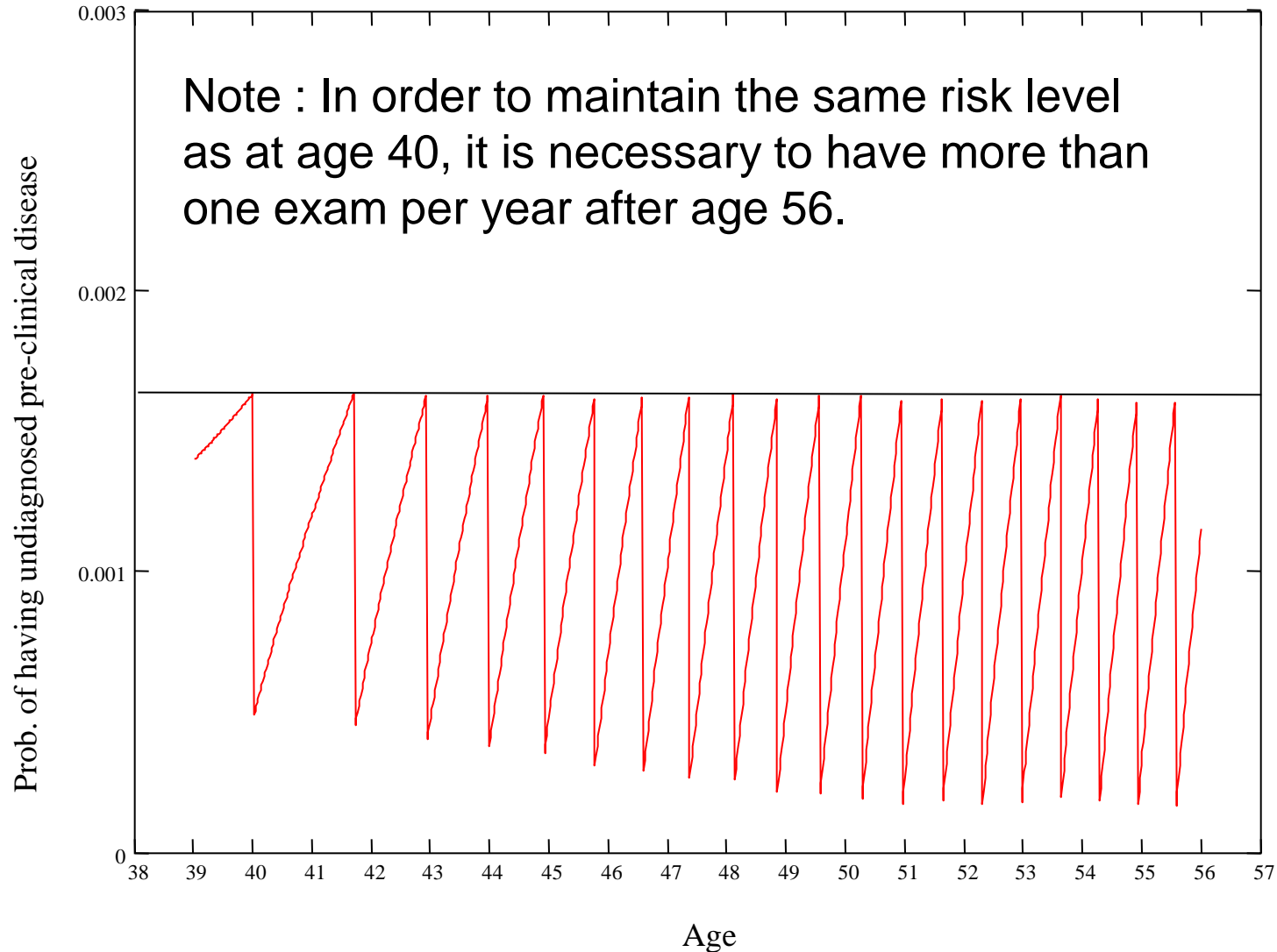
## **Exam Schedule with $S(50)=.0062$ (Threshold Value)**



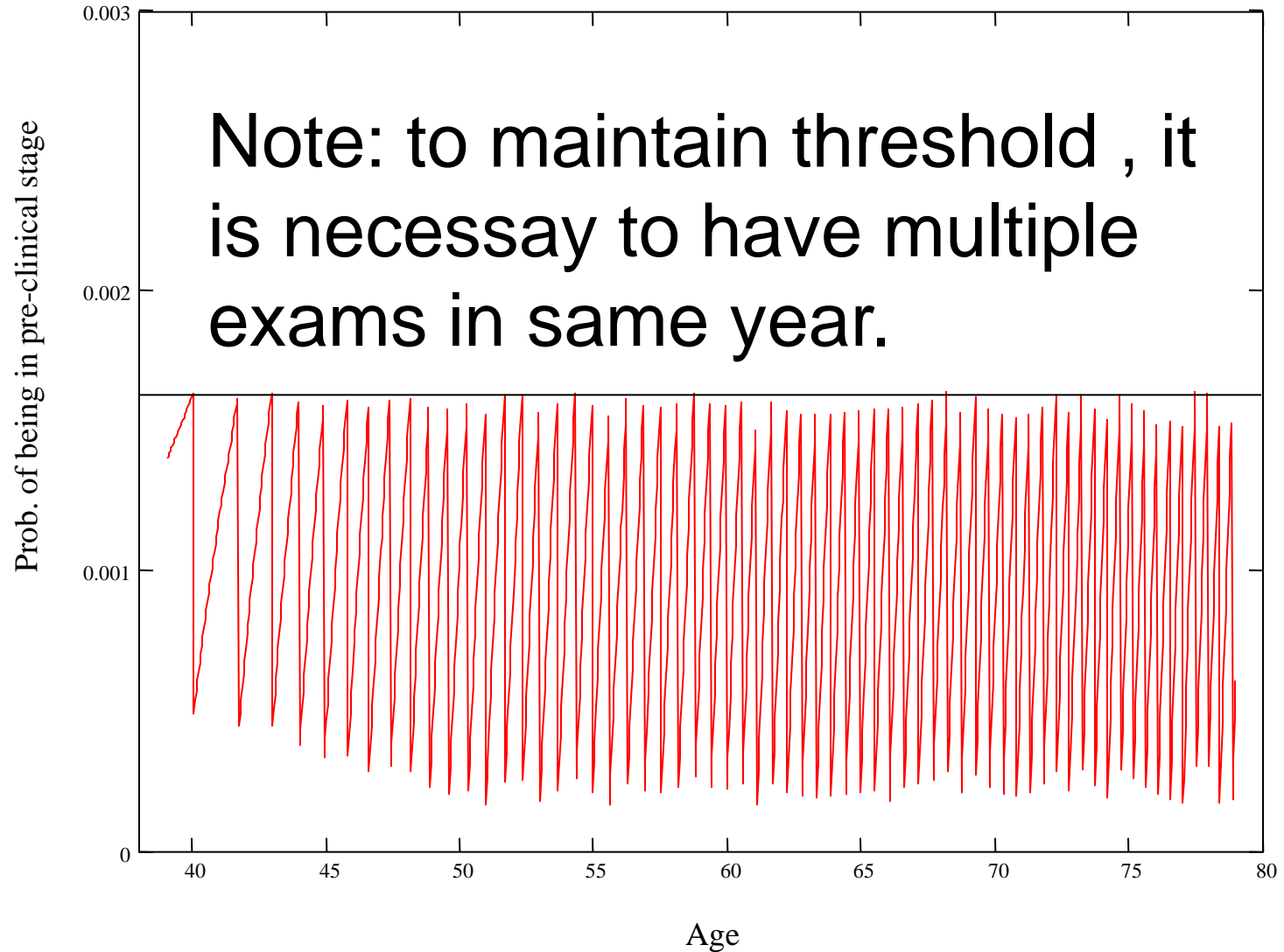


# $S(t) = \text{Prob}(\text{Undiagnosed Pre-Clinical Disease})$

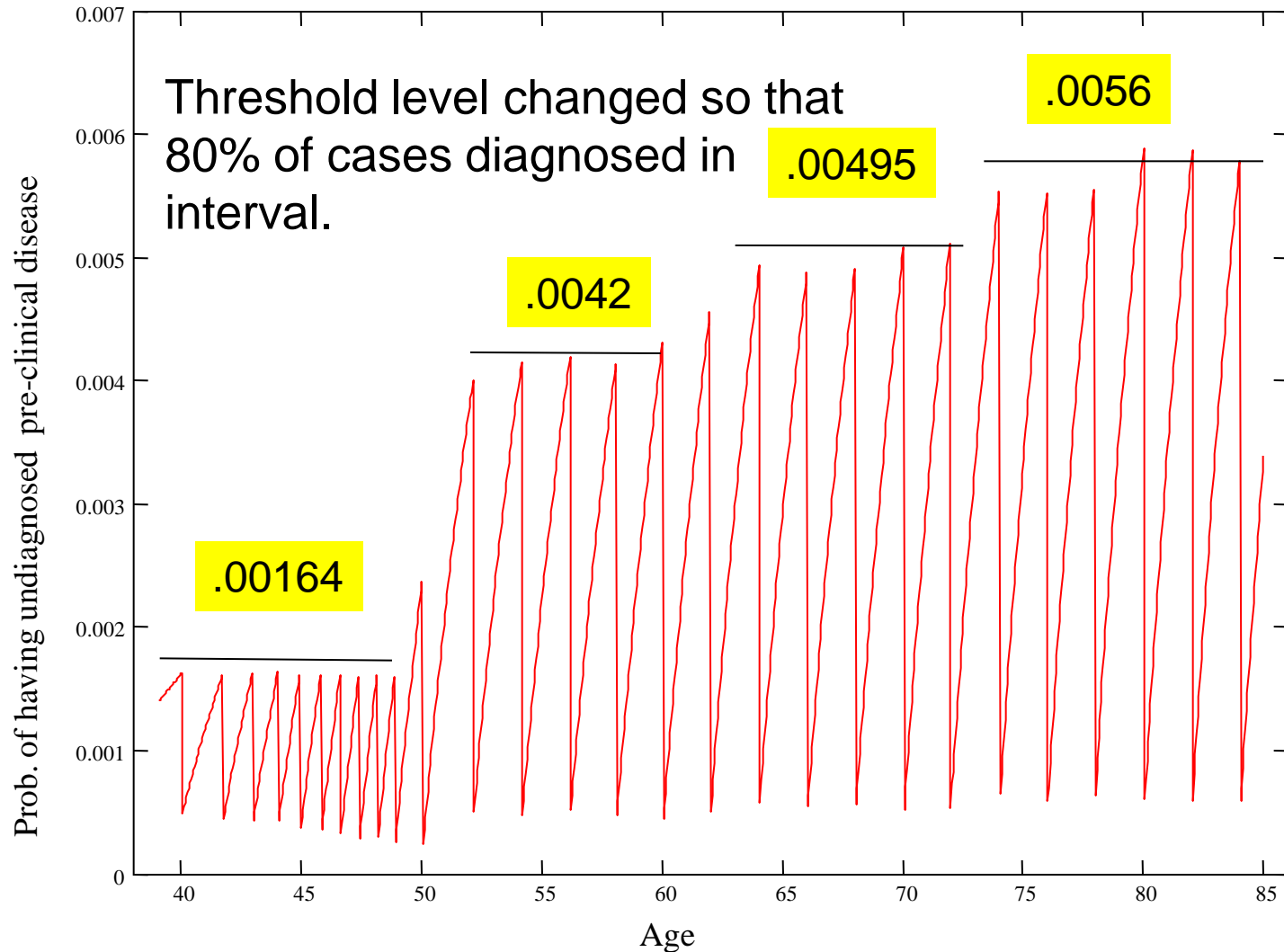
## Exam Schedule with $S(40) = 0.00164$ (Threshold Value)



# S(t) with Threshold Method starting at age 40



# S(t) with Threshold Method Starting at Age 40 Using Non-Constant Threshold Values



# Summary

<u>Recommendations</u>	<u>No. of exams</u>	<u>Mortality Reduction***</u>
USPSTF*	13	21.8%
Threshold*	9	19.0%
Mixed threshold**	28	27.9%
ACS** (annual)	35	26.8%
Threshold**	55	31.2%

\* Screening ages : 50-74

\* \* Screening ages : 40-74

\*\*\* Deaths counted in age range 40-85

# Conclusions

- Schedule recommendations based on fixed time periods are illogical as some exams are ultimately given at lower risk levels than the initial exam risk.
- Risk based schedules are superior to fixed time schedules and result in lower costs and/or higher benefit.

# Maximum Benefit from Screening

Maximum benefit occurs if 100% of cases are diagnosed by early detection exam.

## Schedule Sensitivity

= Proportion of cases diagnosed by exam

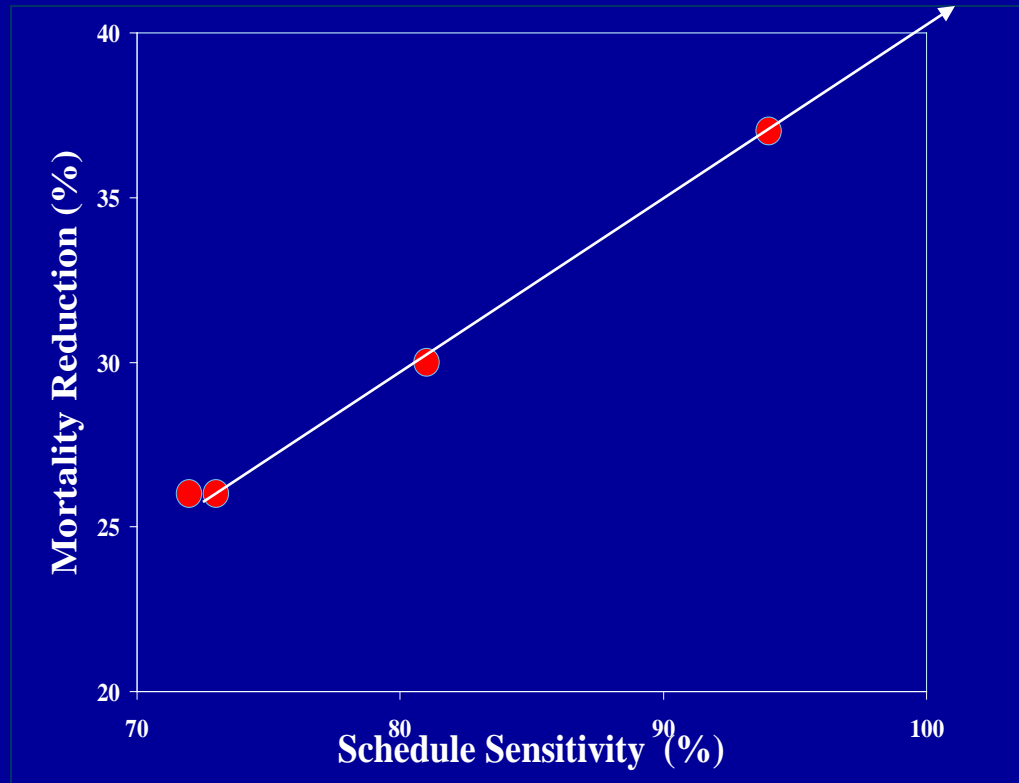
$$= E / E + I + B$$

**E** = Exam diagnosed cases

**I** = Interval cases

**B** = Cases diagnosed before first exam

# Mortality Reduction vs. Schedule Sensitivity : Ages 50-79



**Maximum reduction in mortality is 43.5% if  
100% of cases diagnosed by exam.**

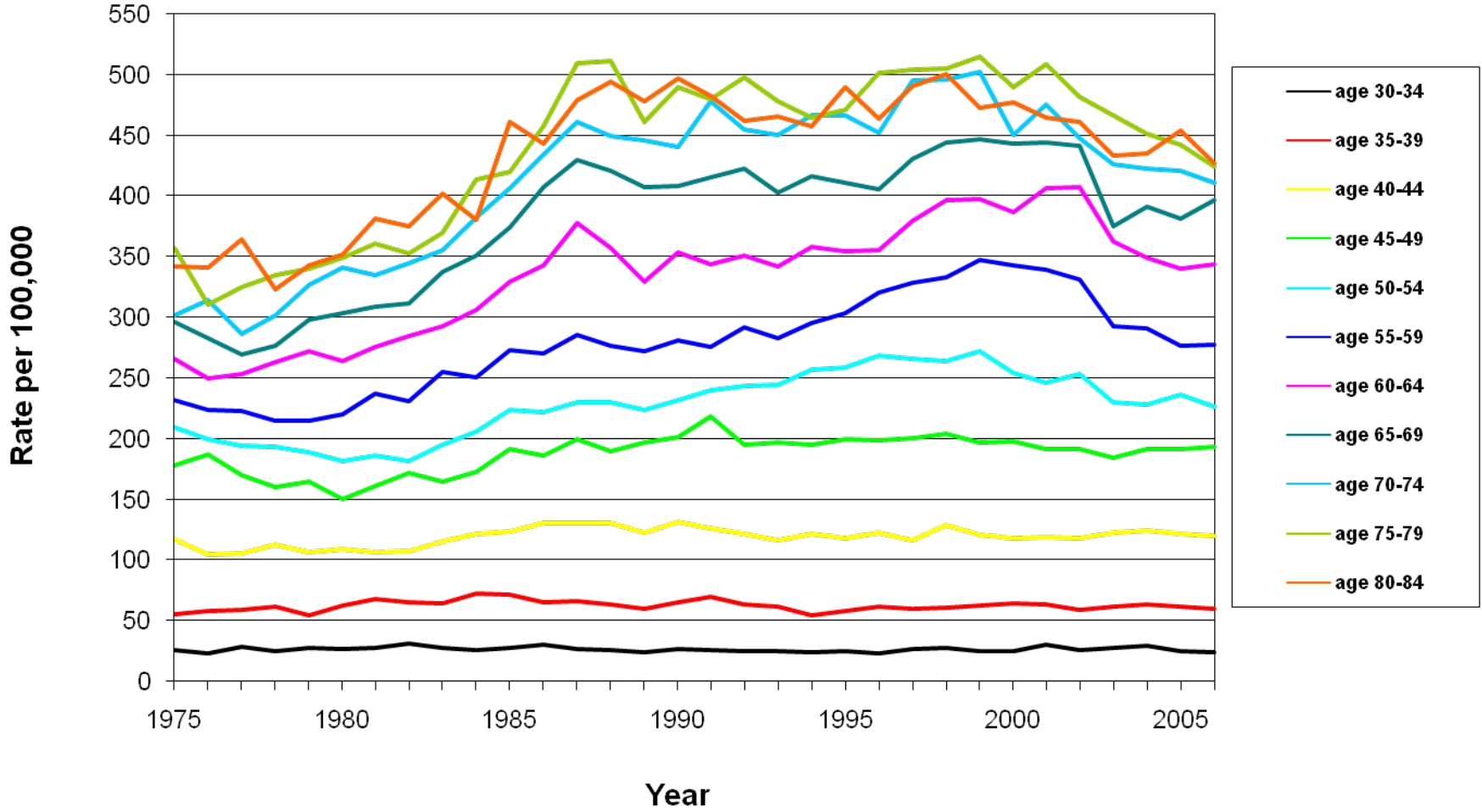
# Screening Younger Women [40, 49]

- **Screening women in [40, 49]**
  - **Relatively low chance of developing breast cancer**
  - **Difficulties in detecting breast cancer**
  - **Relatively high cost**
- **1997 NIH Consensus Development Panel**
  - **Review of data from 8 clinical trials**
  - **“The available data did not warrant a single recommendation for all women in their forties.”**
- **Involvement of US Congress**
- **Recommendation of ACS and NCI**



# U.S. Breast Cancer Incidence Rates 1975-2006

## Low Incidence Rates for Younger Women



# Strategy in Evaluating Benefit for Women Aged 40-49

- **Clinical trials and recent data indicate a stage shift with early detection for this age group. Node negative ~77%% (screening) vs. 53% (usual care).**
- **Compare the mortality of a screened group ( screening only for women in the age group 40-49 ) with a control group (not receiving special exams). Only count deaths for subjects who were in the pre-clinical state during ages 40-9. Note that these subjects may die of disease past the age of 49. The population who were in the pre-clinical state in their 40's is the target population who can benefit.**

# Mortality reduction : Screening in 40's

	I	II	II	IV	V
<b>Idealized mortality reduction*</b>	<b>24%</b>	<b>21%</b>	<b>18%</b>	<b>15%</b>	<b>12%</b>
<b>mortality reduction**</b>	<b>5.3%</b>	<b>4.3%</b>	<b>3.7%</b>	<b>3.2%</b>	<b>2.5%</b>

\* Counts only deaths for subjects in pre-clinical state in ages 40-9.

\*\* Counts all deaths for ages 40-79.

**I. 40, 41, 42, 43, 44, 45, 46, 47, 48, 49 (annual)**

**II. 40, 42, 44, 46, 48, 49**

**III. 40, 43, 46, 49**

**IV. 40, 45, 49**

**V. 40, 49**

# Conclusion

- Screening women in their 40's does result in some benefit.
- However the contribution to the overall mortality reduction is 5.3% for 10 annual exams over a 40 year follow-up period.
- This is too small to pick up in a clinical trial unless there is a huge number of individuals in trial.

# Stopping Screening by Age

- Should screening be terminated for older women?
- Suppose breast cancer annual screening is initiated at age 50 and mortality is observed to age 79.
- Screening is terminated at ages : 64, 69 , 74, 79
- **Question.** How is mortality reduction affected by termination of screening? Deaths are counted in the age interval (50-79)

## Mortality reduction vs. Age at which screening is stopped (deaths are counted to age 85)

Annual Screening from Age 50 to Age	No. of Exams	Schedule Sensitivity	Mortality Reduction
<b>64</b>	15	48%	<b>23.5%</b>
<b>69</b>	20	64%	<b>30.7%</b>
<b>74</b>	25	79%	<b>35.8%</b>
<b>79</b>	30	87%	<b>36.9%</b>

# Screening Schedules for High Risk Population

- **Subpopulations may be at Elevated Risk**
- **No Clear Guidelines for Choosing Screening Schedules Based on Risk Status**
- **Risk Factors for Breast Cancer**
  - **Previous breast cancer, mutations on BRCA1 or BRCA2, First degree relatives with breast cancer, 2+ breast biopsies, 75%+ dense tissue on MM**
  - **Age at menarche, age at first live birth, no of previous biopsies**

# Initial Age to Begin Screening

- Women at normal risk have a probability of 0.0062 of being in pre-clinical state at age 50. Can we use this as a guideline to initiate screening of high risk population at earlier age?
- If recommendation of normal population is to begin screening at age 50, then using the same probability, a high risk population should begin screening at ages:

<b>RR</b>	<b>1</b>	<b>1.5</b>	<b>2</b>	<b>3</b>
<b>Age</b>	<b>50</b>	<b>47</b>	<b>44</b>	<b>41.5</b>

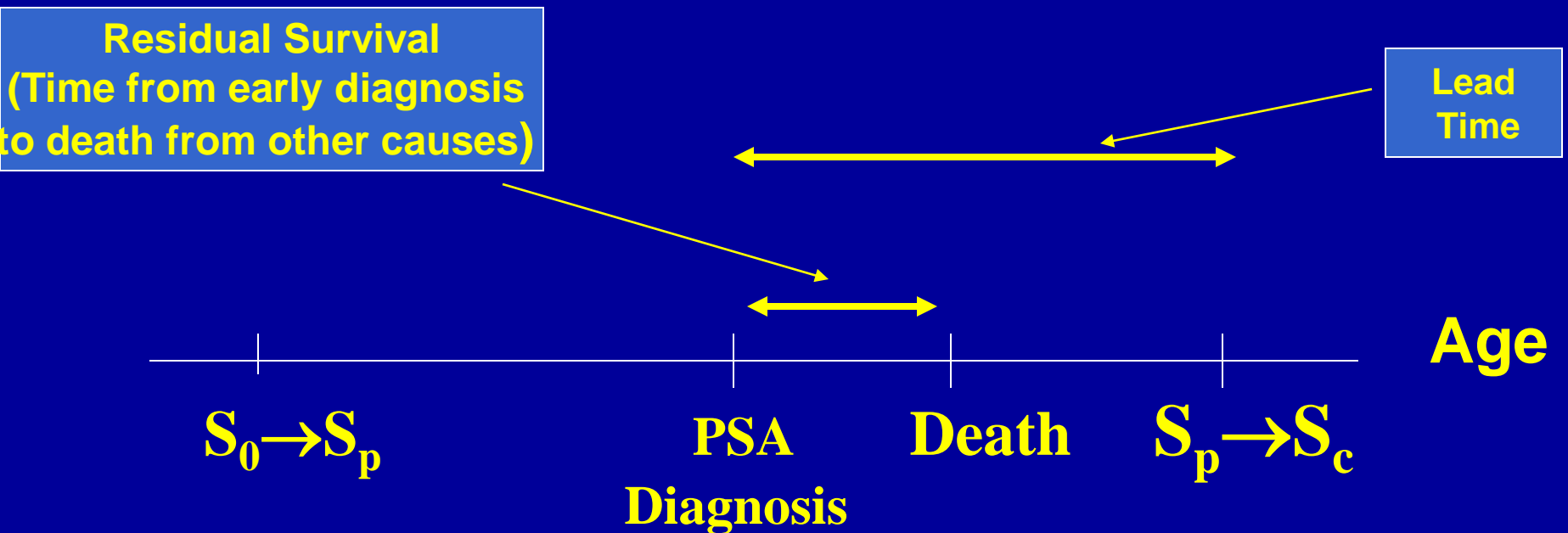


# Overdiagnosis: Prostate Cancer

Background: Prostate Specific Antigen (PSA) test is widely used to diagnose prostate cancer. A positive result triggers a biopsy. Nearly all diagnosed cases by PSA are asymptomatic.

Question: Would the prostate cancer exhibit clinical symptoms during a man's lifetime? If not --- PSA diagnosis is an overdiagnosis

**Over diagnosis: Lead Time > Residual Survival**



# Comparisons of Probability of Over Diagnosis

(Prostate cancer)

<u>Age</u>	<u>Model</u>	<u>Draisma</u>
55	0.26	0.27
60	.33	.38
65	.40	.47
70	.48	.53
75	.57	.56
80	.65	

Draisma et al. jnci (2003)—Based on European Trial; (sojourn time 15.4 yr, relevant cases)

Populations : model (U.S. ,all races), Others are white.

# Probability of Over-diagnosis Conditional on Age at Detection: Breast Cancer

**Age**

40	50	60	70	80	90
.00	.02	.05	.10	.21	.42

**Probability**

**Conclusion . Over diagnosis not a serious problem**

# Conclusions

- 1. Models are necessary in order to answer important issues arising in the screening of chronic diseases.**
- 2. Recommendations for screening schedules should be risk based rather than time based.**

**THANK YOU FOR COMING**

- The following seven slides may not be used.

# Initial Age to Begin Screening : High Risk Women

- **Probability of being in the pre-clinical state for the non-high risk population (RR =1), by age, is:**

Age	40	45	50
Probability	.0016	.0035	.0062

**If the high risk women (RR >1 ) were to have the same threshold probability as the non-high risk population, when would screening start?**

# Age to Start Screening

$$P(40) = .0016$$

RR	1	1.5	2	3
Age	40	37.5	35.5	31.5

$$P(45) = .0035$$

RR	1	1.5	2	3
Age	45	42.5	40.5	38

All high risk women have the same threshold probability as normal population when screening begins



# Over Diagnosis

- **It is possible for some diseases to be diagnosed early which would never have clinical symptoms in a person's lifetime.**
- **Ordinarily the disease is treated when diagnosed; it is not known whether the disease may exhibit clinical symptoms during a person's lifetime.**

# Summary on Stage Shift : Breast cancer

<u>Trial</u>	<u>Stage</u>	<u>Control</u>	<u>Study</u>	<u>Screen Detected</u>	<u>Interval Detected</u>
		↓			↓
HIP	N+	52%	41%	30%	50%
2-County	II-IV	46%	30%	19%	45%
Malmo I *	“	59%	39%	23%	56%
Edinburgh	“	87%	64%	50%	72%

-----  
 II-IV : Stages II - IV ( AJCC)

- **The interval cancers tend to have the same stage as a control group.**
- **If stage shift was due to length biased sampling, the interval cancers would tend to have more advanced stage than a control group. This is not true.**

# Notes on Modeling

- **Survival begins at point of clinical diagnosis for usual care group (control).**
- **In order to make comparisons with control group, all cases in screened group (early diagnosis, interval) must have survival beginning at point of “clinical diagnosis”. This is true for interval cases, but not true for screened diagnosed cases.**
- **It is necessary for model to subtract lead time (random variable, not observed) from survival for screened cases so that survival is measured from point of imputed clinical diagnosis (not observed).**
- **Screened cases are subject to length biased sampling. This feature must be incorporated in the model.**

# Stage Shift and Earlier Diagnosis

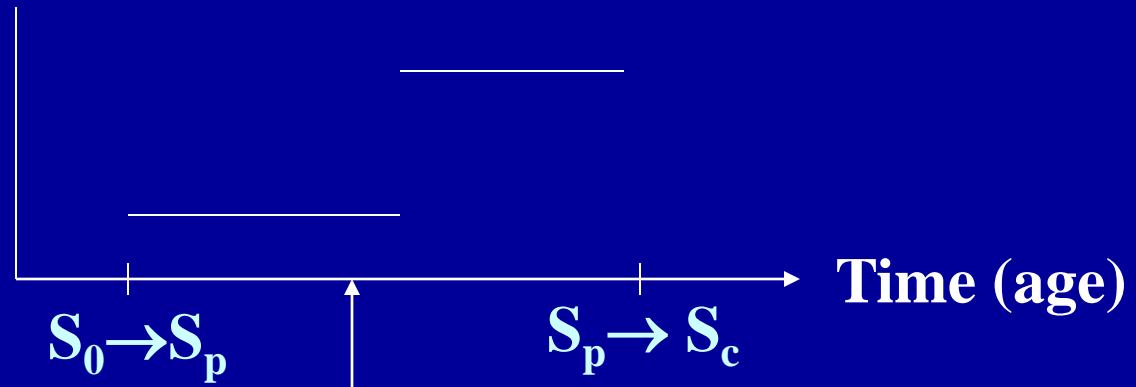
$S_0$ : Disease Free State or Cannot Be Detected

$S_p$ : Pre-clinical State

$S_c$ : Clinical State

Stage II

Stage I



Early Diagnosis