Screening of Diseases

DR ANAND BHIDE
ASSOCIATE PROFESSOR
COMMUNITY MEDICINE
MIMER MC

Content

- Definition of Screening
- Why do we Screen
- Screening test versus Diagnostic Test
- Types of Screening
- Criteria for Introduction of Screening Programs
- Criteria for Screening Test (Including Validity and Reliability)
- Biases on Evaluation of Screening Program
- Ethics and Economics

Definition of Screening

The **presumptive** identification of unrecognized disease or defect by the application of tests, examinations or other procedures that can be applied rapidly.

-Oxford Textbook of Public Health, 4th Edition

Why do we screen?

Because of the Iceberg phenomenon!

Now, same is true for most of the diseases. The tip of the iceberg represents the symptomatic cases a physician sees. The bigger, submerged part represents the asymptomatic cases a physician misses out.

So screening is done to help us see at least some part of this hidden iceberg.

It is carried out in the hope that earlier diagnosis and subsequent treatment favorably alters the natural history of the disease in a significant proportion of those who are presumptively identified as positive.

Concept of 'Lead-Time'

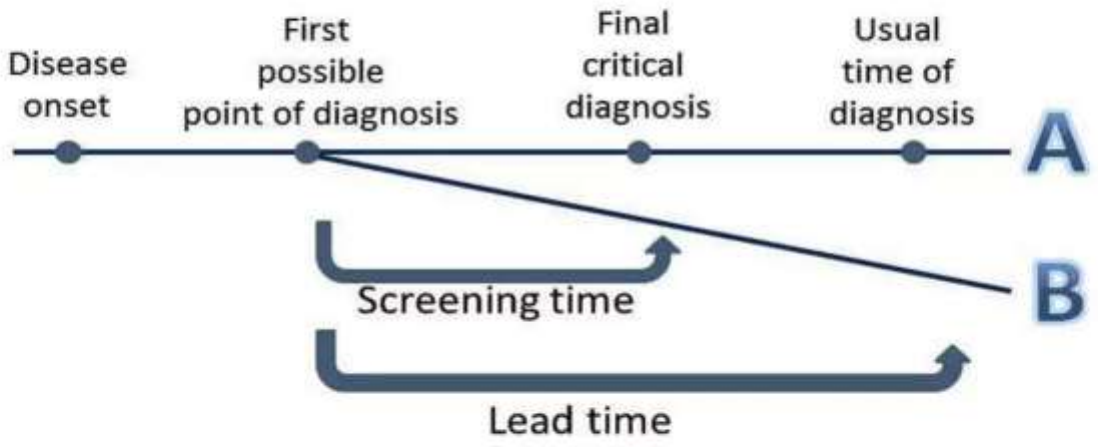


Image 1 : The Concept of Lead Time

Lead time is the advantage gained by screening. This time gives us the opportunity to intervene and change the outcome favorably from A to B.

Screening Test versus Diagnostic Test

	Screening Test	Diagnostic Test
Purpose	To detect potential disease.	To establish presence/absence of a disease.
Target Population	Large number of people who do not suspect of having a disease, but are potentially at risk individuals.	Symptomatic individuals (to establish diagnosis), or asymptomatic individuals with a positive screening test.
Test Method	Simple, acceptable to patients and staff.	Maybe invasive, expensive but justifiable as necessary to establish diagnosis.
Positive result threshold	Generally chosen towards high sensitivity.	Generally chosen towards high specificity.
Positive Result	Essentially indicates suspicion of disease that warrants confirmation.	Results provide a definite diagnosis.
Cost	Relatively cheaper.	Relatively costlier.

Some Common Screening Tests

- Random blood glucose for diabetes
- Blood pressure for hypertension
- Pap smear for cervical cancer
- Mammography for breast cancer

Types of Screening

- Mass Screening
- High Risk / Selective Screening
- Multipurpose screening
- Multiphasic screening
- Opportunistic/Case finding screening

Mass Screening – Screening done in whole population or a sub group of population, irrespective of the particular risks an individual may have of contacting the disease. Example – If we go tomorrow to a corporate office and screen every employee for Hypertension, it will be called mass screening.

High Risk Screening – Screening that is done in individuals having some risk factors for the disease in question. Example – If we go tomorrow to a corporate office and screen only those employees who smoke & drink.

Multipurpose screening – Screening by more than one test done simultaneously to detect more than one disease.

Multiphasic screening - It is the application of two or more screening tests in combination to a large number of people at one time instead of carrying out separate screening tests for single diseases.

Opportunistic screening – To detect diseases in individuals seeking healthcare for some other reasons.

Example - A corporate office of 300 workers. 100 are smokers.

Image 2- Example explaining different types of screening

Criteria for Introduction of Screening Programmes (Wilson and Jungner Criteria)

- The condition sought should be an important health problem.
- There should be an accepted treatment for patients with recognized disease.
- Facilities for diagnosis and treatment should be available.
- There should be a recognizable latent or early symptomatic stage.
- There should be a suitable test or examination.
- The test should be acceptable to the population.
- The natural history of the condition, from latent stage to declared disease, should be adequately understood.

- There should be an agreed policy on whom to treat as patients.
- The cost of case-finding (including diagnosis and treatment of patients diagnosed) should be economically balanced in relation to possible expenditure on medical care as a whole.
- Case-finding should be a continuing process and not a “once and for all” project.

Emerging screening criteria proposed over the past 40 years

- The screening program should respond to a recognized need.
- The objectives of screening should be defined at the outset.
- There should be a defined target population.
- There should be scientific evidence of screening program effectiveness.
- The program should integrate education, testing, clinical services and program management.
- There should be quality assurance, with mechanisms to minimize potential risks of screening.

- The program should ensure informed choice & confidentiality.
- The program should promote equity and access to screening for the entire target population.
- The overall benefits of screening should outweigh the harm.

Criteria for a Test to be used as a Screening Test

- Acceptable
- Easy to perform
- Rapid
- Cheap
- Valid
- Reliable

Acceptablility

- The test should be acceptable to the people, and the healthcare providers.
- Usually, the tests which are very painful, discomforting and embarrassing are not acceptable to the population.
- For example : Breast examination for lumps by health care provider, Urine examination for glucose.

Easy to perform

- The test should be simple to perform. Even better if it can be performed by Allied health professional. Example – Measuring Blood Pressure

Rapid

- The test should able to give fast results. For Example – Preference of RBS over OGTT

Low cost

- The test should be available at a reasonable price. For example - RBS

Validity

- It refers to what extent the test accurately measures what it is supposed to measure. It has two components – Sensitivity and Specificity.
- **Sensitivity** of the test is defined as the ability of the test to identify correctly those who have the disease.
- **Specificity** of the test is defined as the ability of the test to identify correctly those who do not have the disease.

To ascertain whether someone actually has a disease or not, a Gold Standard test should be there to compare.

Gold Standard

- A gold standard test is usually the diagnostic test or benchmark test that is the best available under reasonable conditions. Its results are considered definitive.

- For Example :- Culture for Infectious Diseases and Biopsy for cancers.

Example

(for dichotomous results)

Total population – 1000

Prevalence – 10%

		Disease		Total
		+	–	
Test	+	80 (TP)	100 (FP)	180
	–	20 (FN)	800 (TN)	820
Total		100	900	1000

Table No.1 - Calculation of the Sensitivity and specificity of Screening Examinations.

Sensitivity = 80/100 = 0.8 or 80 %

Specificity = 800/900 = 0.89 or 89 %

Accuracy = 880/1000 = 0.88 or 88 %

Formulas to remember

- Sensitivity = TP / TP+FN
- Specificity = TN / TN+FP
- Accuracy = TP+TN / TP+TN+FP+FN

Example

(for continuous variables)

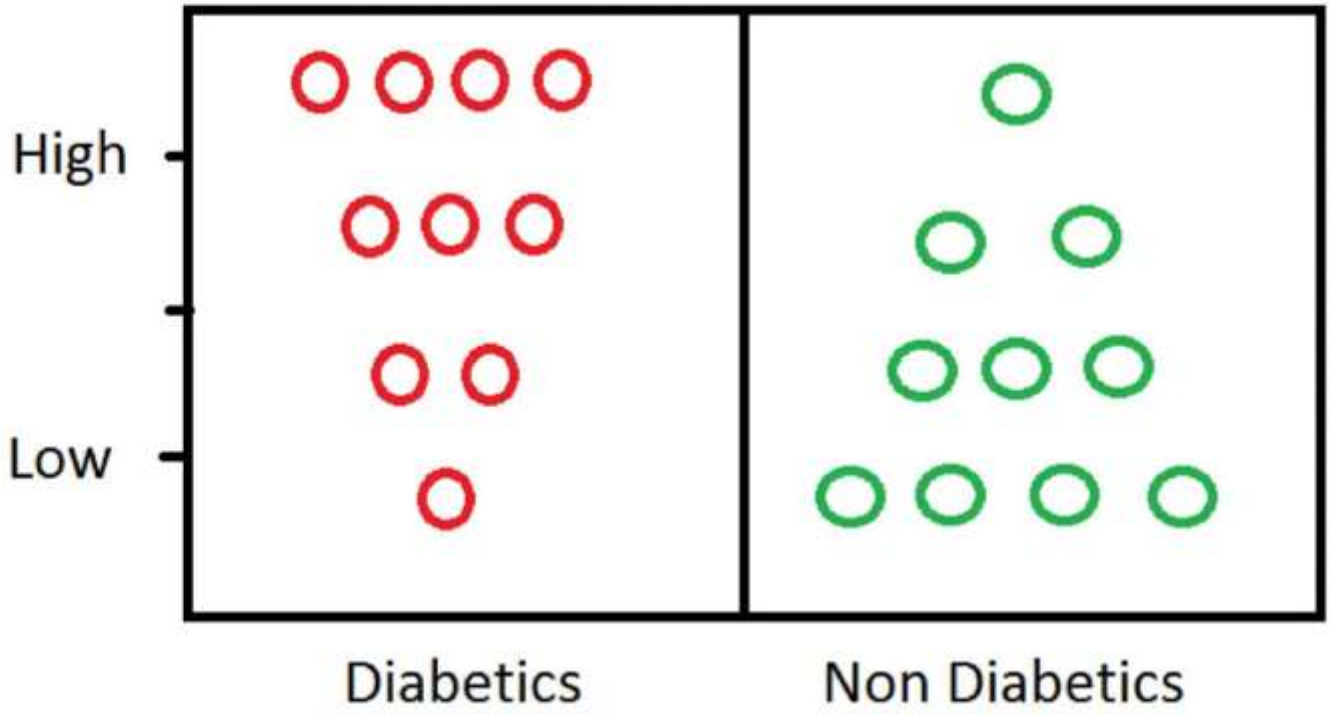
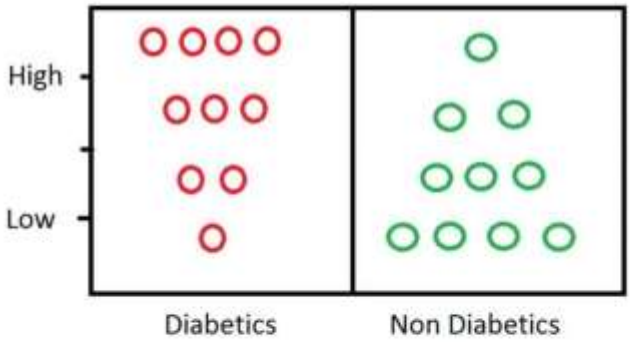


Image 3 : Diabetic and Non Diabetic people having different levels of RBS. Total Population = 20 Prevalence = 50%

If we set the cut off high



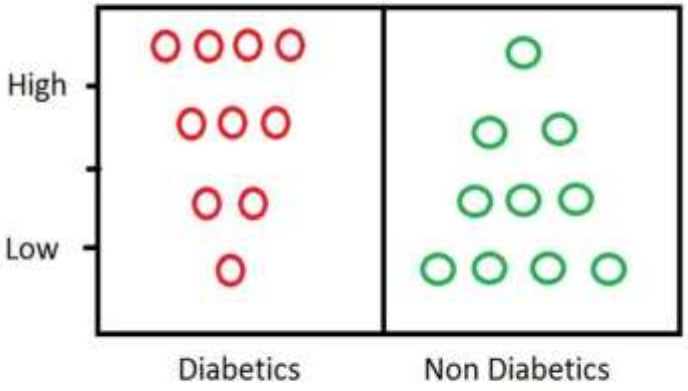
		Disease		Total
		+	-	
Test	+	4	1	5
	-	6	9	15
Total		10	10	20

Table No.2 - Calculation of the Sensitivity and specificity when cut off is high.

Many Diabetics labeled as Non Diabetics. (False Negatives)
Most of the Non Diabetics identified correctly.

Sensitivity = 4/10 = 0.4 or 40 %
Specificity = 9/10 = 0.9 or 90 %

If we set the cut off low



		Disease		Total
		+	–	
Test	+	9	6	15
	–	1	4	5
Total		10	10	20

Table No.3 - Calculation of the Sensitivity and specificity when cut off is low.

Many Non Diabetics labeled as Diabetics. (False positives)
Most of the Diabetics identified correctly.

Sensitivity = 9/10 = 0.9 or 90 %
Specificity = 4/10 = 0.4 or 40 %

Lessons Learnt

- Different cut off points will yield different sensitivities and specificities.
- The cut off point determines how many individuals will be labeled as diseased.
- A cut off point that will identify more True Positives, will also identify more False Positives.
- A cut off point that will identify more True Negatives, will also identify more False Negatives.

What is the problem with False Negatives and False Positives?

False Negatives

1. False reassurance
2. Ignoring of disease signs and symptoms
3. Postponement of treatment
4. Comparable to death sentence if the disease is treatable only in early stages

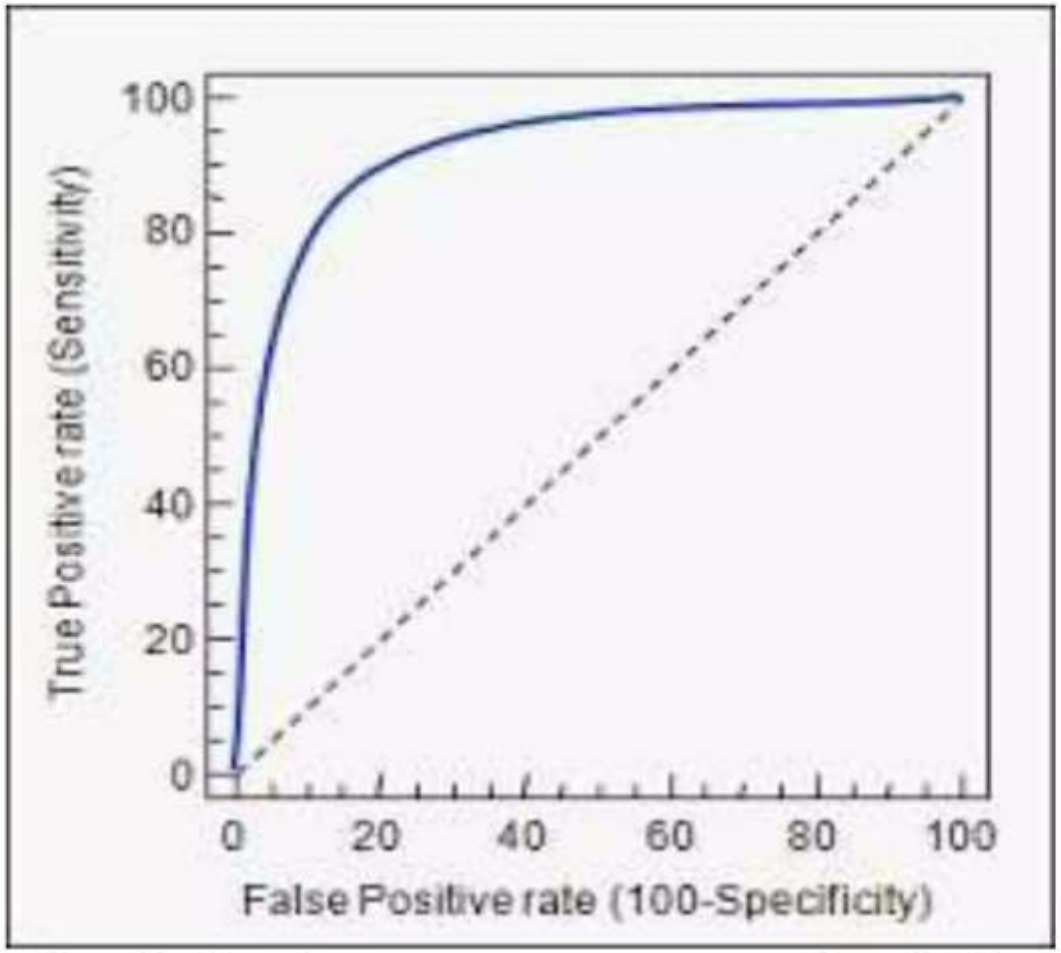
False Positives

1. Discomfort, inconvenience, anxiety , stress
2. Burden on health facilities.
3. One who gets the label, finds it difficult to remove the label.
4. Suicide if the disease is an embarrassment .

So what should be the cut off?

- The choice is always yours. ROCC (Receiver Operating Characteristic Curve) can help.
- In this curve, Sensitivity (True Positive Rate) is plotted on the vertical axis and 1-Specificity (False Positive Rate) is plotted on the horizontal axis. After this, all possible combinations are plotted and a graph is obtained.
- The point farthest from the 45 degree diagonal is the optimal point.

Sn=1, Sp=1 Sn=1, Sp=0



Sn=0, Sp=1 Sn=0, Sp=1

Image 4 : ROC Curve

- But ROC Curve gives equal weightage to Sensitivity and Specificity, which may not be desirable in many cases.

Minimize False Positives (False Negatives can be afforded),	Minimize False Negatives (False Positives can be afforded)
The disease is not serious	The disease is serious
Diagnostic Test is costly	Diagnostic Test is inexpensive
The disease does not spread fast	The disease spreads fast
The disease is easily treatable in later stages	The disease is easily treatable in earlier stages only.

Use of Multiple tests

Sequential Testing

- In sequential testing, a less expensive and less invasive test is generally performed first and those who test positive are tested further with a 2nd test which is expensive, but generally having more validity.
- Only person tested positive on both the tests is considered positive. Example :

Total population – 1000

Sn of Test 1 = 80%

Sp Test 1 = 80%

Prevalence – 10%

Sn of test 2 = 90%

Sp Test 2 = 90%

		Disease		Total
		+	-	
Test 1	+	80	180	260(Positives)
	-	20	720	740
Total		100	900	1000

Table No.4 – Results of Test 1 on a Sequential Screening Examination. Sn=0.8, Sp=0.8

		Disease		Total
		+	-	
Test 2	+	72	18	90
	-	8	162	170
Total		80	180	260

Table No.5 – Results of Test 2 on a Sequential Screening Examination. Sn=0.9, Sp=0.9

		Disease		Total
		+	-	
Test Combine	+	72	18	90
	-	28	882	910
Total		100	900	1000

Table No.6 – Combined Results of a Sequential Screening Examination.

- Combined Sensitivity = 72/100 = 0.72 or 72%
- Combined Specificity = 882/900 = 0.98 or 98%

- So in sequential testing, Sensitivity decreases while Specificity increases.

Use of Multiple tests

Simultaneous/Parallel Testing

- In simultaneous testing, two tests are applied to the individuals at the same time and persons testing positive on either one of them are considered positive.
- Example :

Total population – 1000

Sn of Test 1 = 80%

Sp of test 1 = 80%

Prevalence – 10%

Sn Test 2 = 90%

Sp Test 2 = 90%

		Disease		Total
		+	-	
Test 1	+	80	180	260
	-	20	720	740
Total		100	900	1000

Table No.7 – Results of Test 1 on a Simultaneous Screening Examination. Sn=0.8, Sp=0.8

		Disease		Total
		+	-	
Test 2	+	90	90	180
	-	10	810	820
Total		100	900	1000

Table No.8 – Results of Test 2 on a Simultaneous Screening Examination. Sn=0.9,Sp=0.9

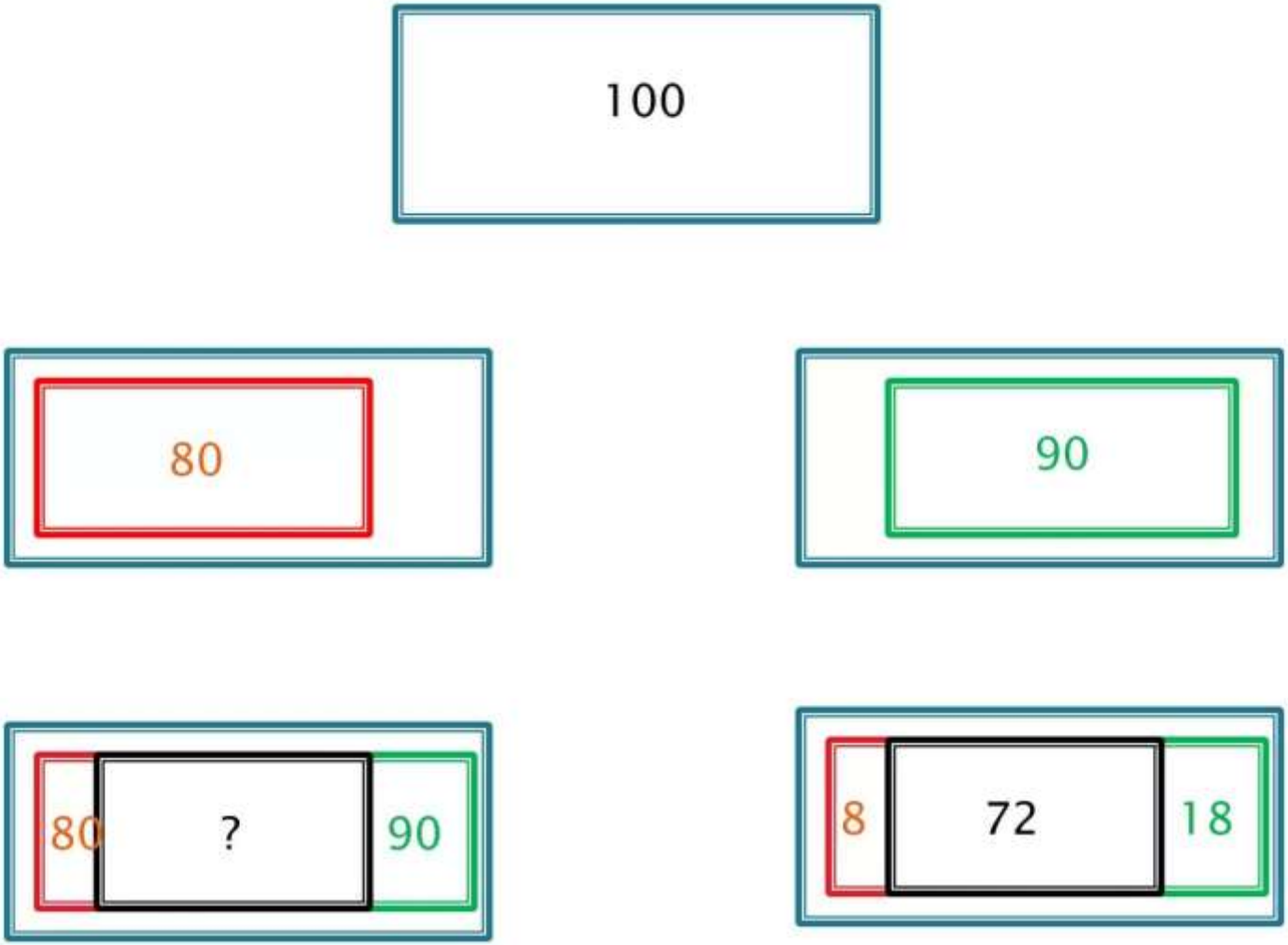


Image 4 : Know how many of the diseased tested positive on both the tests and how many tested positive on any one of the test?

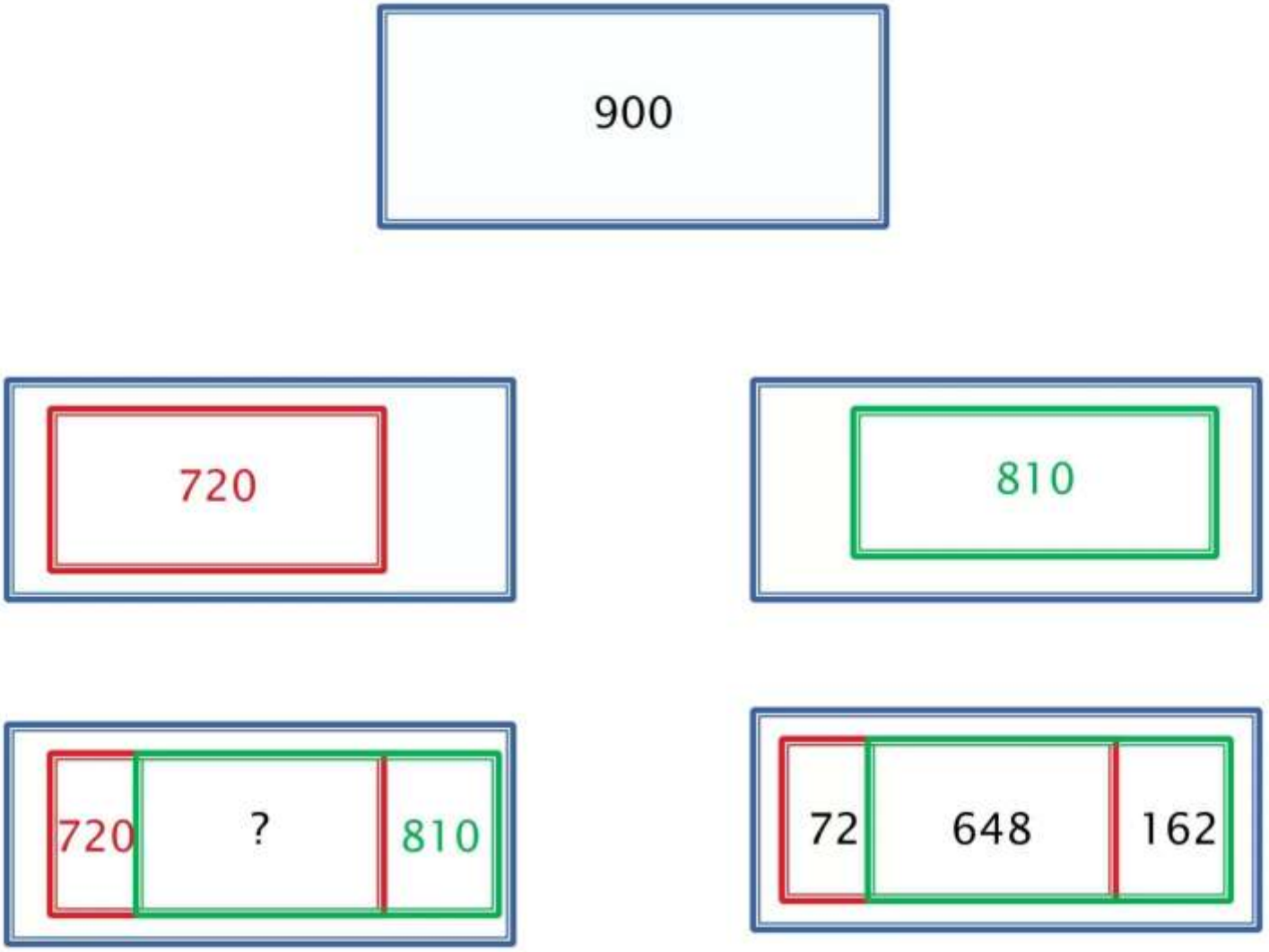


Image 5 : Know how many among the non diseased tested negative on both the tests?

		Disease		Total
		+	-	
Test Combine	+	98	252	180
	-	2	648	820
Total		100	900	1000

Table No.9 – Combined Results of a Simultaneous Screening Examination.

- Combined Sensitivity = 98/100 = 0.98 or 98%
- Combined Specificity = 648/900 = 0.72 or 72%

- So in simultaneous testing, Sensitivity increases while Specificity decreases.

Formulas

Sequential Testing

Sensitivity = Sn1 X Sn2

Specificity = (Sp1 + Sp2) – (Sp1 X Sp2)

Simultaneous Testing

Sensitivity = (Sn1 + Sn2) – (Sn1 X Sn2)

Specificity = (Sp1 X Sp2)

Positive Predictive Value

- How many people among the total tested positive are actually having the disease.
- It is given by TP/TP+FP

Example
Total population – 1000
Prevalence – 10%
Sn = 50%
Sp = 50%

		Disease		Total
		+	-	
Test	+	50 (TP)	450 (FP)	500
	-	50 (FN)	450 (TN)	500
Total		100	900	1000

Table No.10 – Measuring the Positive Predictive Value of a test. Sn=0.5,Sp=0.5

PPV = 50/500 = 0.1 or 10 %

A test with low PPV usually looses its reputation among physicians.

PPV depends on Prevalence, Sensitivity and Specificity

Total population – 1000 Prevalence – 10% Sn = 50% Sp = 50%

		Disease		Total
		+	-	
Test	+	50 (TP)	450 (FP)	500
	-	50 (FN)	450 (TN)	500
Total		100	900	1000

Total population – 1000 Prevalence – 20% Sn = 50% Sp = 50%

		Disease		Total
		+	-	
Test	+	100 (TP)	400 (FP)	500
	-	100 (FN)	400 (TN)	500
Total		200	800	1000

Table 11 & 12 Analysing the dependence of PPV on Prevalence.

PPV changes from 10% to 20 %

Total population – 1000 Prevalence – 10% Sn = 50% Sp = 50%

		Disease		Total
		+	-	
Test	+	50 (TP)	450 (FP)	500
	-	50 (FN)	450 (TN)	500
Total		100	900	1000

Total population – 1000 Prevalence – 10% Sn = 90% Sp = 50%

		Disease		Total
		+	-	
Test	+	90 (TP)	450 (FP)	540
	-	10 (FN)	450 (TN)	560
Total		100	900	1000

Table 13 & 14 : Analysing the dependence of PPV on Sensitivity.

PPV changes from 10% to 16.67 %

Total population – 1000 Prevalence – 10% Sn = 50% Sp = 50%

		Disease		Total
		+	-	
Test	+	50 (TP)	450 (FP)	500
	-	50 (FN)	450 (TN)	500
Total		100	900	1000

Total population – 1000 Prevalence – 10% Sn = 50% Sp = 90%

		Disease		Total
		+	-	
Test	+	50 (TP)	90 (FP)	140
	-	50 (FN)	810 (TN)	860
Total		100	900	1000

Table 15 & 16 : Analysing the dependence of PPV on Specificity.

PPV changes from 10% to 35.7%

Negative Predictive Value

- How many people among the total tested negative are actually free of disease.
- It is given by TN/TN+FN

Example
Total population – 1000
Prevalence – 10%
Sn = 50%
Sp = 50%

		Disease		Total
		+	-	
Test	+	50 (TP)	450 (FP)	500
	-	50 (FN)	450 (TN)	500
Total		100	900	1000

Table No.17 – Measuring the Negative Predictive Value of a test.

NPV = 450/500 = 0.9 or 90 %

Reliability/Repeatability

- It is the ability of the test to reproduce results if the test is repeated.
- Factors that contribute to this are

1. Intrasubject variation
2. Intraobserver variation
3. Interobserver variation

Intrasubject variation

- Occurs because the values of a test may be different on an individual at different times.
- Example :

1. Blood Pressure of an individual fluctuates over the day
2. Blood Glucose levels of an individual fluctuates over the day.

Intraobserver variation

- It occurs because an observer may give two different results at different times even when the finding is same.
- For example – A radiologist giving two different opinions on the same X-ray film at two different times.

Interobserver variation

- Two different examiners often do not give exactly same results, even if they are examining the subject on the same time.
- For example :

1. Two interns giving different Blood Pressure readings for the same patient in Aliganj OPD
2. Two Radiologists giving different opinions about X-ray films.

		Obs 2		Total
		+	-	
Obs 1	+	A	B	A+B
	-	C	D	C+D
Total		A+C	B+D	A+B+C+D

Table No.18 – Measuring the percent agreement between two observers.

Percent Agreement = A+D/A+B+C+D

But if the D is very large (which usually is),

Percent agreement = A/A+B+C

Kappa Statistic

- There will always be some agreement between two people. This is called chance agreement.
- But we want to know to what extent do the readings agree beyond what is expected by chance. An approach to answer this question is Kappa Statistic.

Kappa = Percent agreement – Chance agreement / 100 – Chance agreement

		Dr. Rahul		Total
		+	-	
Dr. Kiran	+	60	40	100
	-	20	80	100
Total		80	120	200

Percent agreement = (60+80) / 200 = 0.7 or 70%

		Dr. Rahul		Total
		+	-	
Dr. Kiran	+	40	60	100
	-	40	60	100
Total		80	120	200

Table No.19 & 20 – Measuring the Kappa Statistic for two observers.

Chance agreement = (60+40) / 200 = 0.5 or 50%

- Kappa Statistic = (70-50) / (100-50)
  = 0.4 or 40%

- Landis & Koch suggested that :

1. K <0.4 = Poor agreement
2. K 0.4 to 0.75 = Intermediate to good agreement
3. K > 0.75 = Excellent agreement

- Kappa tells us about the quality of healthcare services being offered. If the kappa is low, there is room for improvement.

Validity and Reliability

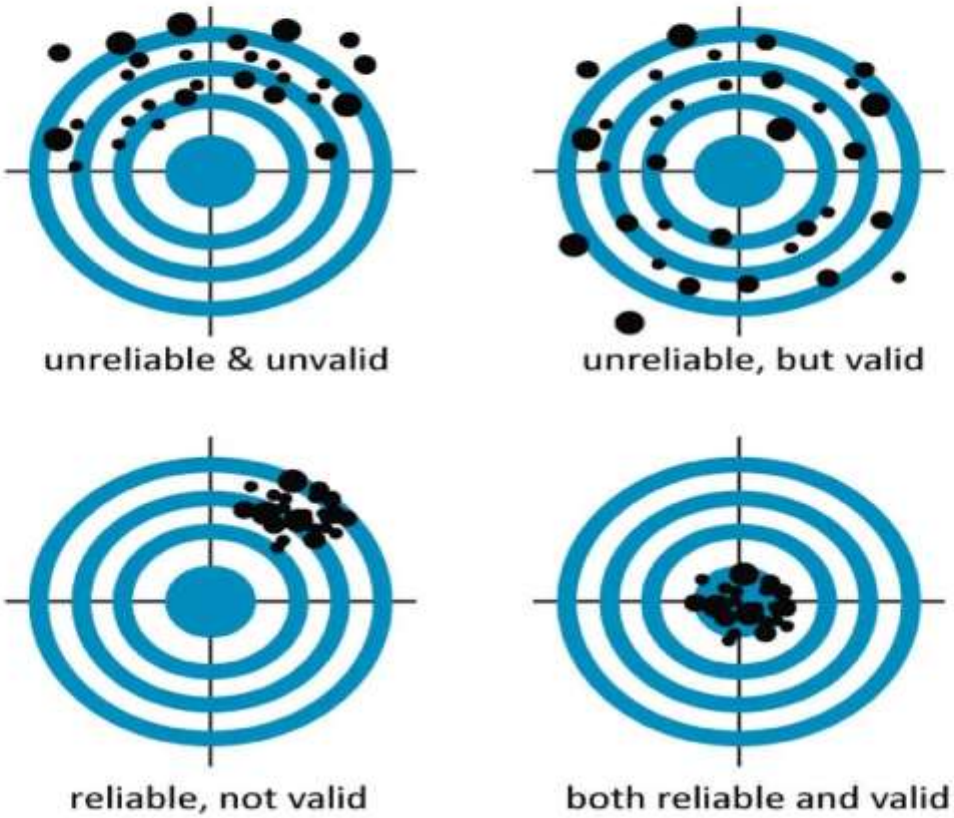


Image 6 : Different possible situations for validity and reliability.

Yield

- The amount of disease that was diagnosed as a result of screening effort.

		Disease		Total
		+	-	
Screening Test	+	270	70	340
	-	30	630	660
Total		300	700	1000

Table No.21 – Measuring the Yield of a screening test. Prevalence=30%, Sn=0.9, Sp=0.9

After the screening test, diagnostic test is applied and the 270 individuals with disease are identified.

Yield = 270/1000 = 27%

Biases on Evaluation of a Screening Programme

Biases giving good image to Screening of Diseases :

- Lead Time Bias
- Length Time Bias
- Overdiagnosis Bias
- Selection Bias

Lead Time Bias

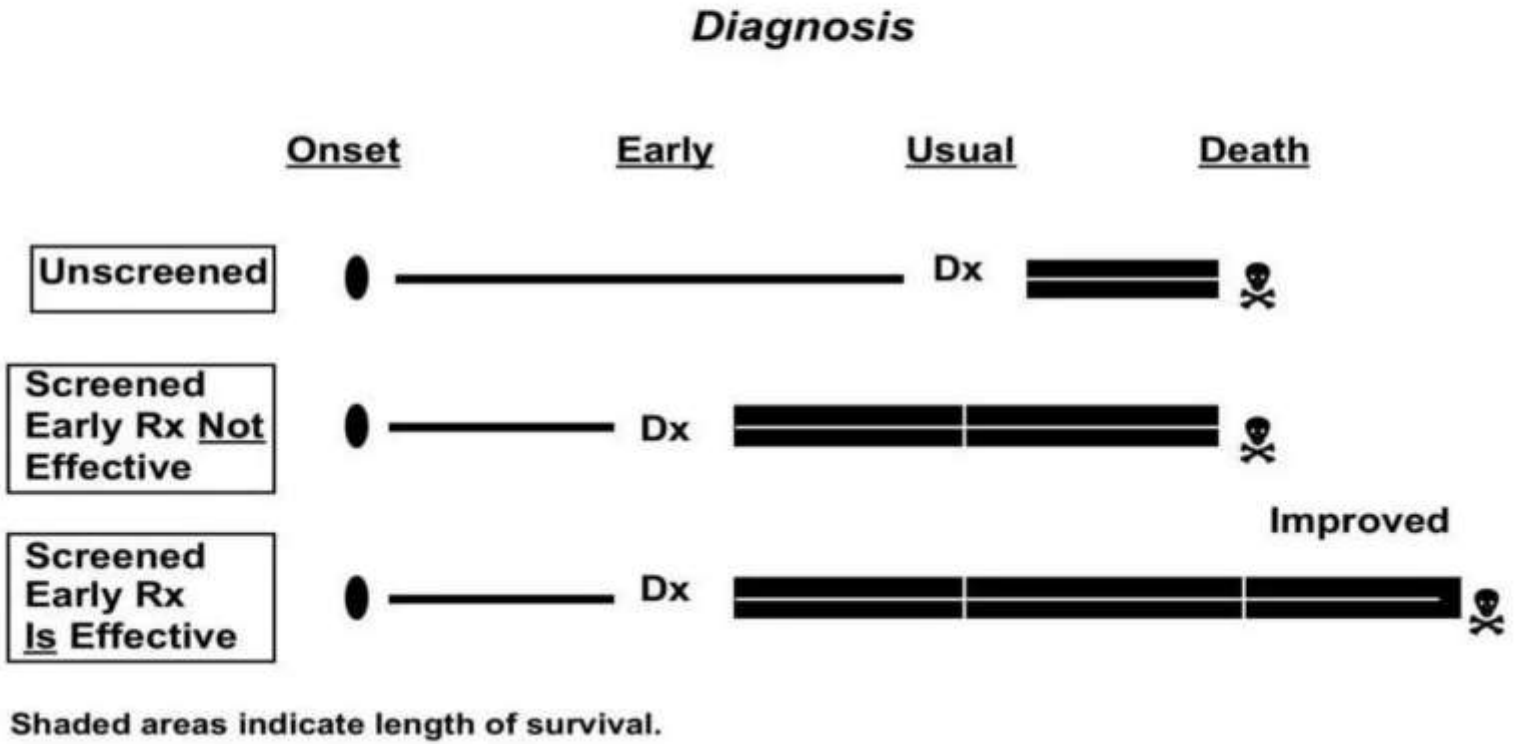


Image 7 : Lead Time Bias explained

In the 2nd situation, we may falsely believe that screening has increased the survival, but it is not so.

Length Time Bias

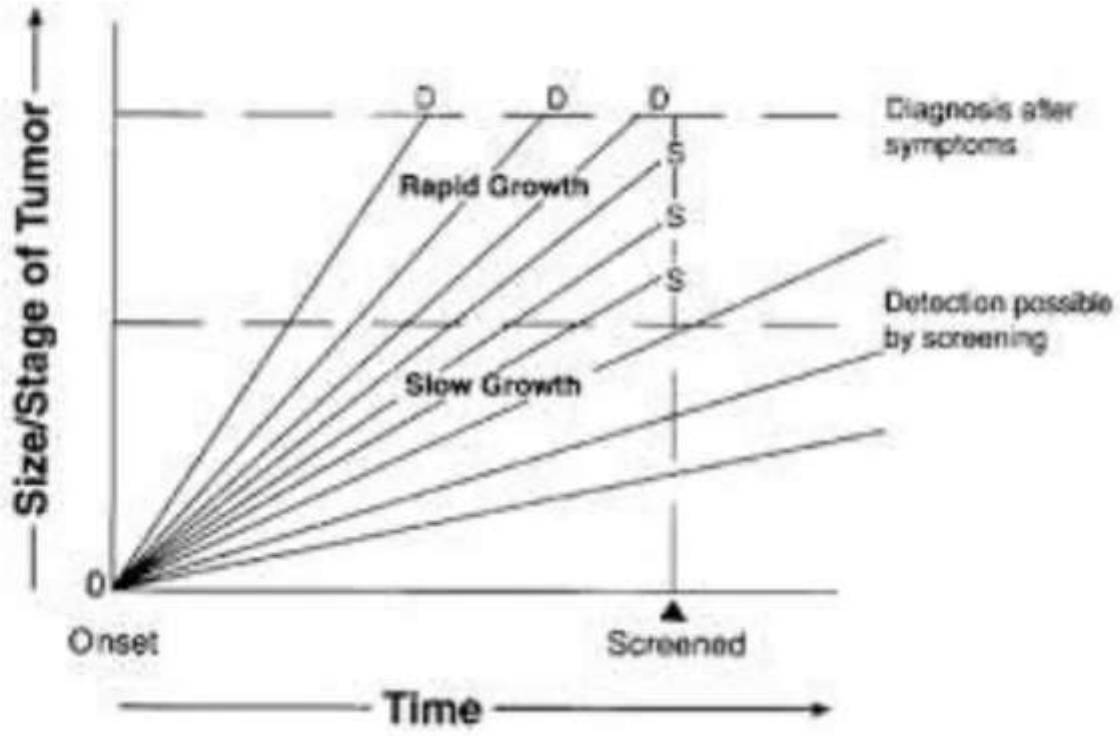


Image 8: Explanation for Length Time Bias.

1. Rapidly evolving cases (poor prognosis) are more likely to end up at the hospital with symptoms.
2. Slowly evolving cases (good prognosis) are more likely to be detected in the screening.

Overdiagnosis bias

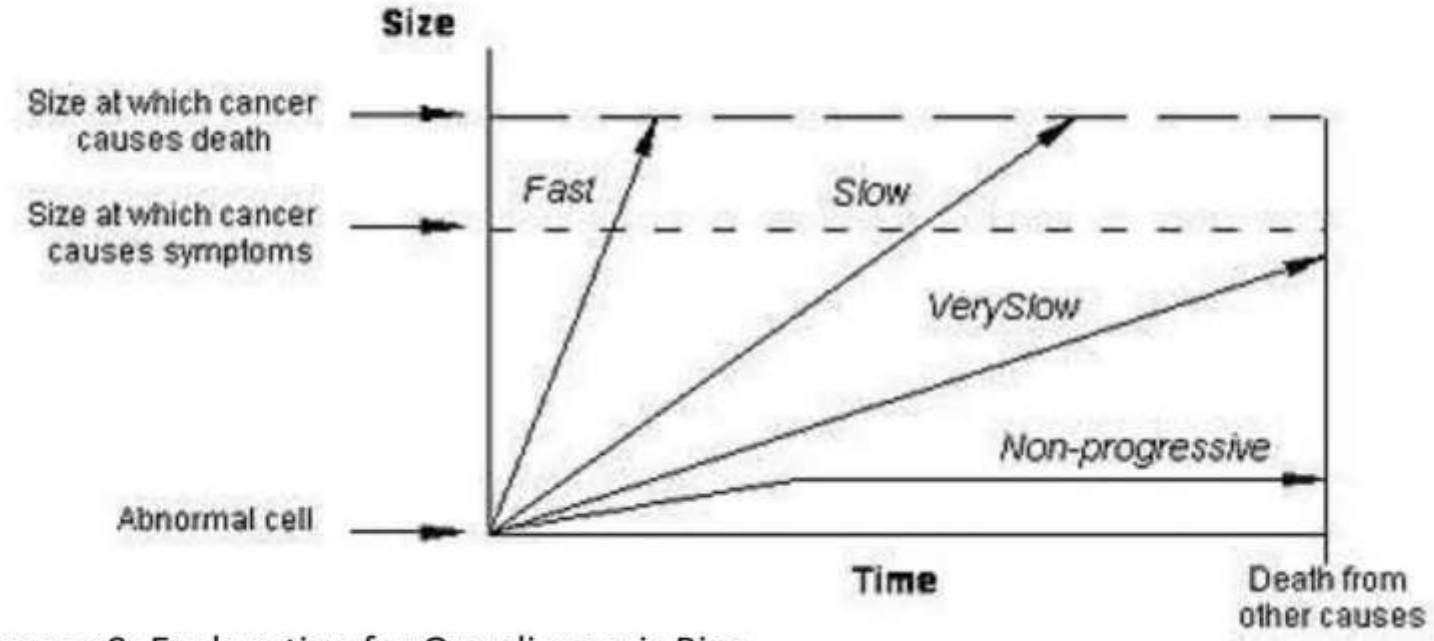


Image 9: Explanation for Overdiagnosis Bias.

- It is an extreme of length time bias.
- It occurs because we diagnose many individual with diseases with screening, who were never going to get any symptoms of that disease in their lifetime.

Volunteer Bias

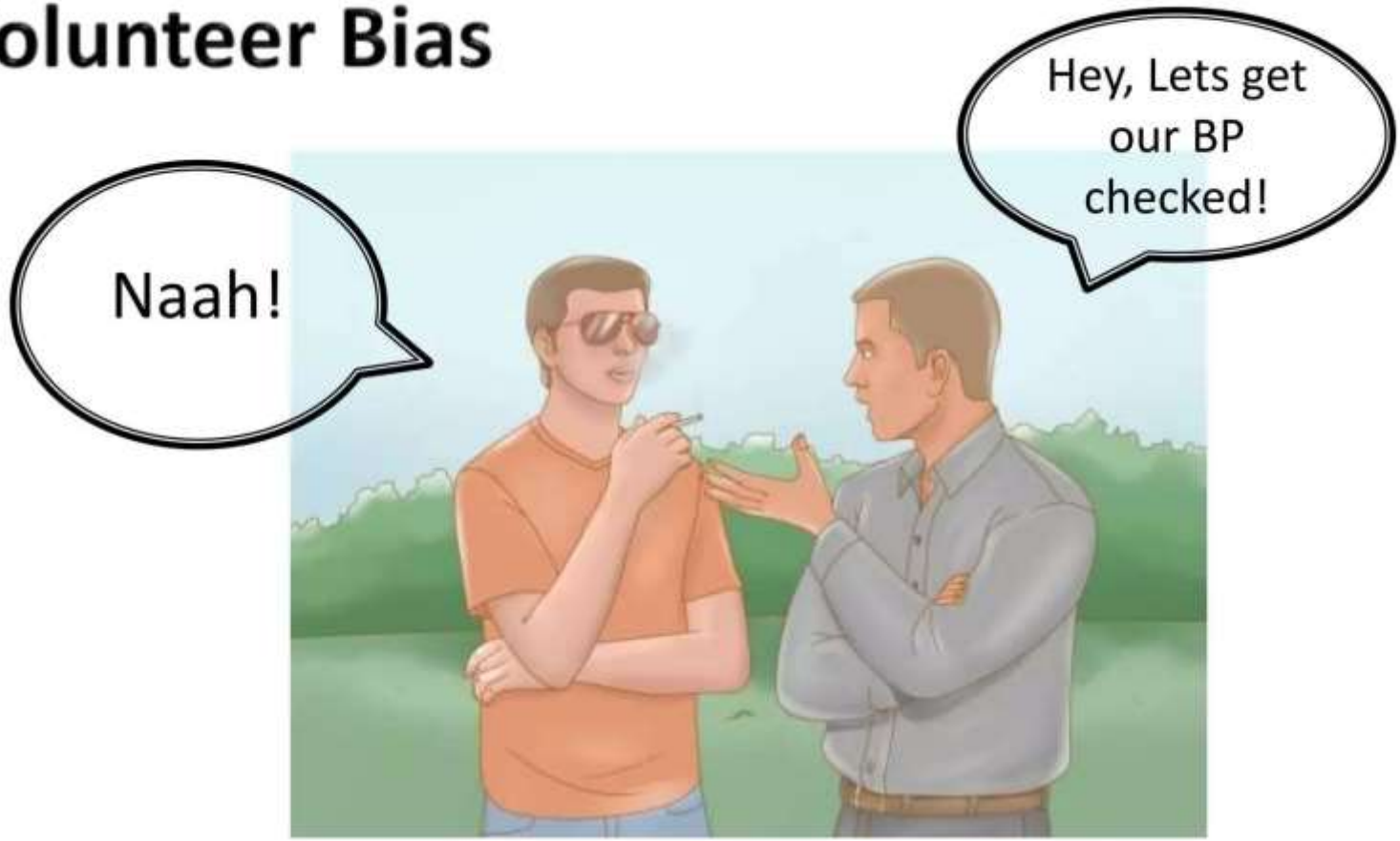


Image 10: Explanation for Volunteer Bias.

- It occurs because the people opting for screening are more health conscious. So they would be having a healthier lifestyle and thus a better prognosis.

Ethics

- In Screening, the physician or Public Health Expert initiates the process of healthcare, so he/she bears the onus of responsibility to be certain that benefit will follow.
- He must ensure that proper facilities for diagnosis and treatment are in place before screening, as otherwise , screening effort would be a waste and a cause of anxiety for the patient.

Economics

- Public Health Expert has an obligation towards the community than towards individuals, so he/she should be concerned how limited resources are equitably distributed across.
- The Cost-Benefit ratio of screening must be evaluated. However, this may turn out to be a quite complex issue, as measuring the economic value of additional life years gained is a difficult task. Also, the costs of screening are often incurred early, while benefits flow later. So the issue becomes more complex.

Refrences

- Oxford textbook of Public Health – 4th Edition
- Epidemiology by Leon Gordis – 6th Edition
- Clinical Epidemiology : The Essentials by Fletcher – 5th Edition
- Textbook of Preventive and Social Medicine by K. Park – 25th Edition
- Basic Epidemiology by R. Bonita – 2nd Edition
- https://www.who.int/bulletin/volumes/86/4/07-050112/en/


THANK
YOU!