

## C.I. di Metodologia clinica

### I metodi per la misura della diagnosi

#### Obiettivo

Conoscere ed utilizzare i principali strumenti per interpretare l'attendibilità e la rilevanza dei test diagnostici ai fini della decisione clinica



### I metodi per la misura della diagnosi

Alla fine di questa lezione dovrete essere in grado di:

- Descrivere cosa si intende per riproducibilità di un test diagnostico
- Calcolare e interpretare l'indice Kappa ( $\kappa$ )
- Descrivere le finalità di un programma di screening
- Valutare potenzialità e limiti di un test per la diagnosi precoce



## L'accuratezza dei test diagnostici

- Qual è la probabilità che questo particolare paziente con questo particolare risultato di questo particolare test abbia proprio quella malattia?
- Quanto sono affidabili i risultati dell'esame?
  - ✓ Accuratezza
  - ✓ Riproducibilità



## C.I. di Metodologia clinica

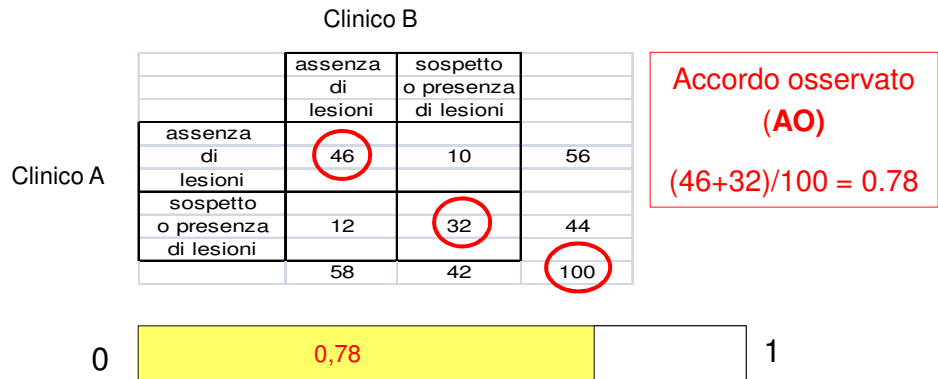
### I metodi per la misura della diagnosi

**Come misuriamo la riproducibilità di un test diagnostico?**



## Riproducibilità del test diagnostico

### L'accordo tra osservatori

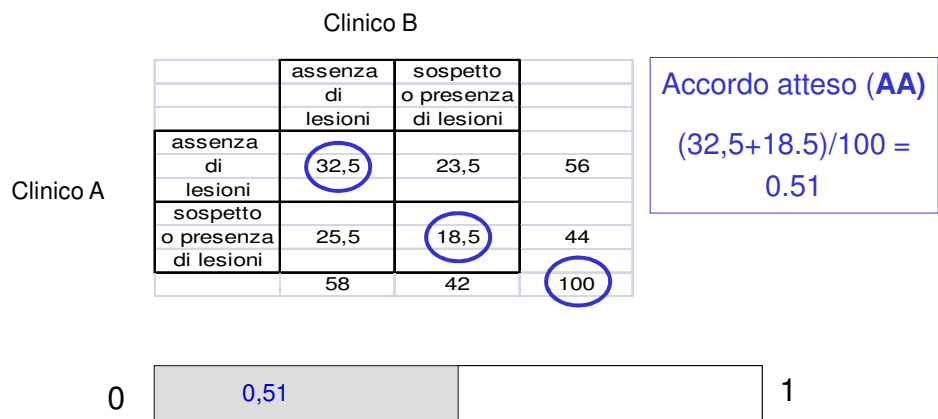


Tale misura non considera però il fatto che i due clinici potrebbero concordare nella diagnosi per effetto del caso



## Riproducibilità del test diagnostico

### Se la concordanza fosse solo casuale



Accordo atteso per effetto del caso



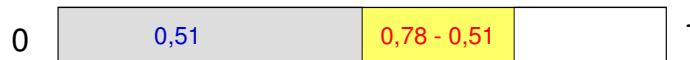
## Riproducibilità del test diagnostico

Se la concordanza totale fosse uguale ad 1 il massimo accordo potenziale ottenibile non dovuto al caso sarebbe:



$$1 - AA = (1 - 0,51) = 0,49$$

Mentre l'accordo reale osservato oltre il caso sarebbe :



$$AO - AA = (0,78 - 0,51) = 0,27$$



## L'indice Kappa ( $\kappa$ )

Il rapporto tra le differenze (AO-AA) e (1-AA) esprime la concordanza oltre il caso

Questo rapporto è chiamato indice **Kappa**

$$\text{Kappa} = \frac{AO - AA}{1 - AA} = \frac{0,27}{0,49} = 55\%$$

La concordanza dei risultati fra gli osservatori rende ragione del 55% della concordanza potenzialmente raggiungibile, al netto dell'effetto del caso



## L'indice Kappa

### Kappa

### Entità della concordanza

< 0	scadente
0 - .20	mediocre
.21 - .40	modesta
.41 - .60	discreta
.61 - .80	buona
.81 - 1.0	ottima

La **riproducibilità** dipende dal contesto di studio (condizioni reali, osservatori competenti) e dalla prevalenza della malattia



## Esempi

### Per valutare la riproducibilità

- > qual è il test sottoposto a valutazione?
- > qual è la domanda clinica cui ci si riferisce?
- > qual è il contesto di valutazione?

Table 2-11. Clinical disagreement in interpreting diagnostic materials

Diagnostic material	Clinical question	Clinicians	Observed	Agreement expected by chance	Kappa
Mammogram*	Should this woman be referred for surgical assessment?	Two radiologists examined the same 1214 mammograms	97%	91%	.67
Exercise electrocardiogram (ECG) <sup>b</sup>	Is the ST-T response normal, borderline, or abnormal?	Two cardiologists examined the same ECGs from 38 patients	57%	39%	.30
		One cardiologist twice examined the same ECGs from 38 patients	74%	33%	.61
Peripheral blood film <sup>c</sup>	Does this patient have iron deficiency?	Two pathologists examined the same films from 29 women with iron-deficiency anemia <sup>d</sup>	69%	49%	.39
		One pathologist twice examined the same films from 29 women with iron-deficiency anemia <sup>d</sup>	83%	51%	.65

\*Data modified from J. Chamberlin, S. Ginks, P. Rogers, et al. Validity of clinical examination and mammography as screening tests for breast cancer. *Lancet* 2:1026, 1975.

<sup>b</sup>Data modified from H. Blackburn, G. Blomqvist, A. Freiman, et al. The exercise electrocardiogram: Differences in interpolation. Report of a technical group on exercise electrocardiography. *Am. J. Cardiol.* 21:871, 1968.

<sup>c</sup>Data modified from V. F. Fairbanks. Is the peripheral blood film reliable for the diagnosis of iron-deficiency anaemia? *Am. J. Clin. Pathol.* 55:447, 1971.

<sup>d</sup>Hemoglobin level 11.6 g/dl or less and serum iron level less than 12.5 µmol/L (70 mg/dl).

### Obiettivo

Confrontare sensibilità, specificità e riproducibilità di tre test per lo screening del cancro della cervice uterina (Pap-test, citologia superficiale, test HPV) in confronto con lo standard (colposcopia e istologia)

La riproducibilità fra osservatori era buona per lo striscio convenzionale (Kappa pesato 0.69, 95% CI 0.64 - 0.74), ma solo moderato per la citologia superficiale (Kappa pesato 0.57, 95% CI 0.52 to 0.63)



## C.I. di Metodologia clinica

### I metodi per la misura della diagnosi

Che cos'è un test di screening?



### CLINICAL PRACTICE

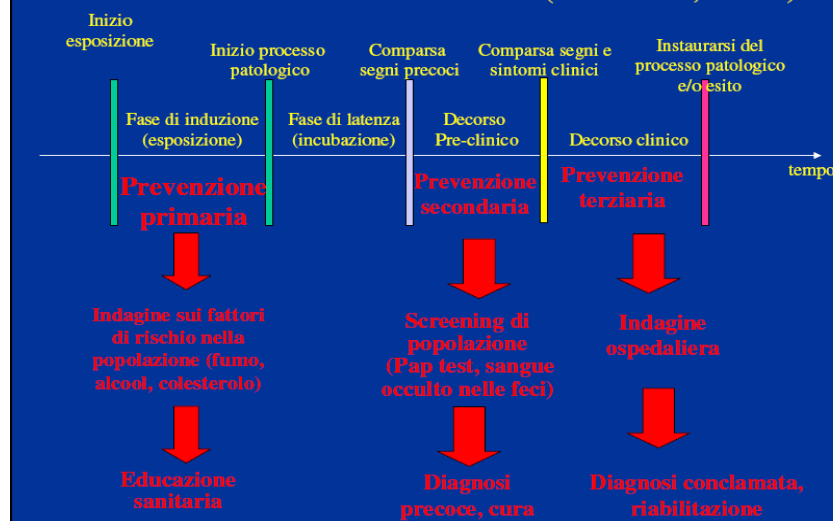
Una donna di 60 anni che ha smesso di fumare 20 anni prima viene per una visita di controllo. In passato ha fumato un pacchetto di sigarette al giorno per 10 anni. L'anamnesi non presenta nulla di rilevante. La donna si sente bene e fa regolarmente attività sportiva. Suo marito ha fumato un pacchetto di sigarette al giorno per almeno 30 anni, ma ha smesso 10 anni fa.

La donna chiede se è utile che lei e il marito si sottopongano alla TAC per lo screening del cancro del polmone. Che consiglio le dai?



### Che intendiamo per prevenzione?

#### Storia naturale di una malattia (Rothman, 1981)



## Strategie di intervento

- Diagnosi clinica precoce
- Case-finding
- Screening
  - ❑ isolato
  - ❑ periodico

**N.B.** Diverso dalle indagini predittive (es. genetiche) che valutano il 'rischio' di andare incontro ad una malattia



## Screening

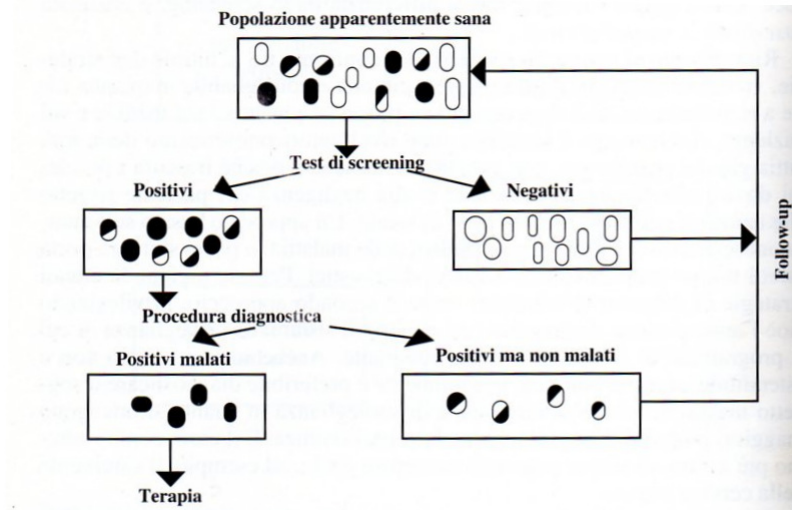
**Mira a riconoscere, all'interno di una popolazione apparentemente sana, gli individui probabilmente malati, con l'intento di modificarne in senso favorevole il decorso clinico.**

- quanto più precoce è la diagnosi tanto più efficace è il trattamento
- lo screening mira ad anticipare la diagnosi alla fase preclinica, prima dell'insorgenza dei sintomi
- conseguentemente, il trattamento è più efficace

**E' veramente così ?**



## Programma di screening



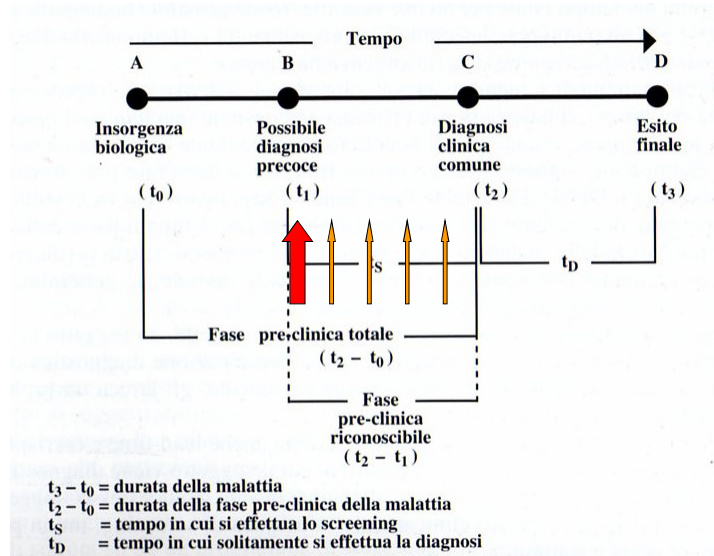
## Obiettivo dello screening

**Attività:** Anticipare la diagnosi

**Efficacia:** Ridurre la mortalità



## Screening e storia naturale della malattia



## Obiettivo dello screening

**Attività:** Anticipare la diagnosi

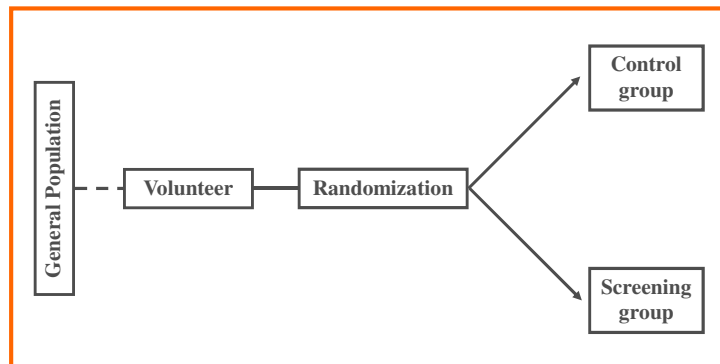
**Efficacia:** Ridurre la mortalità

### Problemi

- Autoselezione ('selection bias')
- Eccesso diagnostico ('overdiagnosis')
- Tempo di anticipazione diagnostica ('lead time')
- Selezione di casi a prognosi favorevole ('length-biased sampling')

## Selection Bias

I soggetti che effettuano lo screening sono diversi dagli altri soggetti e questo può influenzare la sopravvivenza



## Obiettivo dello screening

**Attività:** Anticipare la diagnosi

**Efficacia:** Ridurre la mortalità

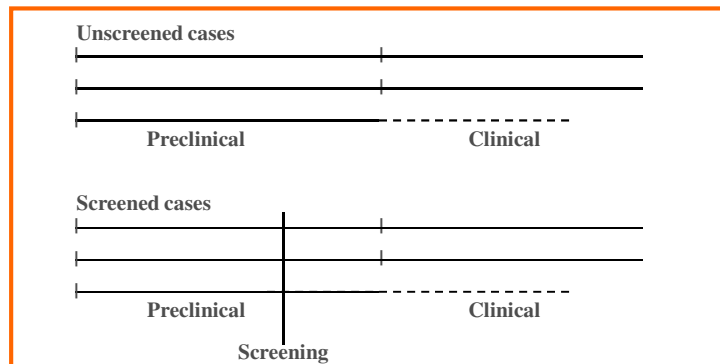
### Problemi

- Autoselezione ('selection bias')
- Eccesso diagnostico ('overdiagnosis')
- Tempo di anticipazione diagnostica ('lead time')
- Selezione di casi a prognosi favorevole ('length-biased sampling')



## Overdiagnosis Bias

Lo screening può rilevare malattie che non sarebbero mai state diagnosticate in assenza di screening



## Obiettivo dello screening

**Attività:** Anticipare la diagnosi

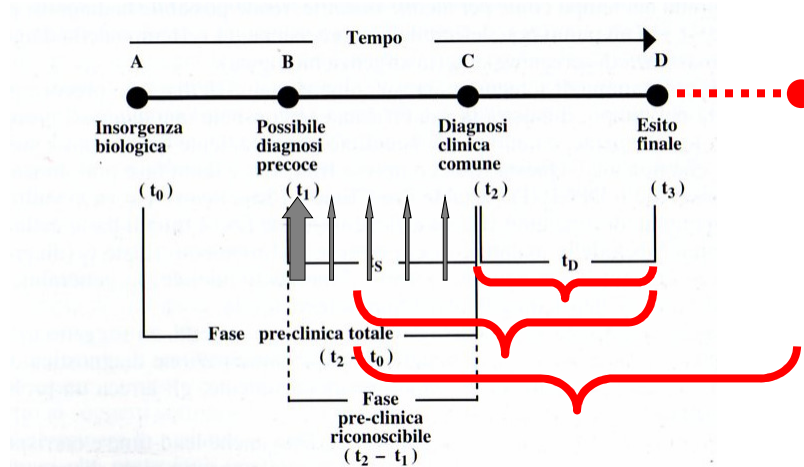
**Efficacia:** Ridurre la mortalità

### Problemi

- Autoselezione ('selection bias')
- Eccesso diagnostico ('overdiagnosis')
- Tempo di anticipazione diagnostica ('lead time')
- Selezione di casi a prognosi favorevole ('length-biased sampling')



## Lead Time Bias (anticipazione diagnostica)



I casi diagnosticati allo screening hanno una sopravvivenza più elevata anche in assenza di qualsiasi efficacia dell'intervento



## Obiettivo dello screening

**Attività:** Anticipare la diagnosi

**Efficacia:** Ridurre la mortalità

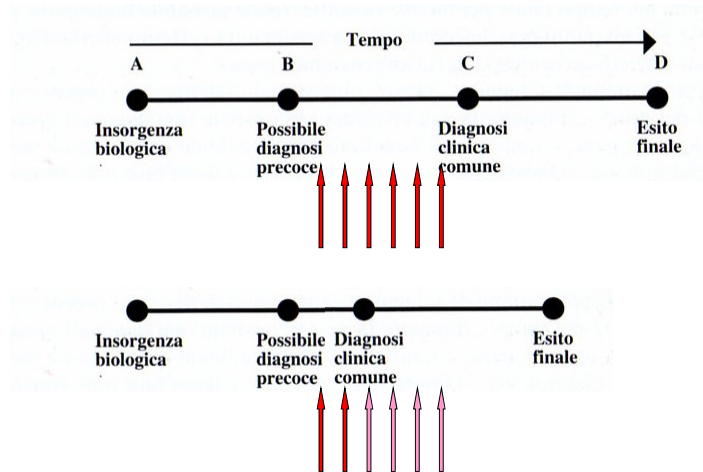
### Problemi

- Autoselezione ('selection bias')
- Eccesso diagnostico ('overdiagnosis')
- Tempo di anticipazione diagnostica ('lead time')
- Selezione di casi a prognosi favorevole ('length-biased sampling')



## Length-biased sampling

I casi con una progressione più lenta della malattia hanno una maggiore probabilità di essere diagnosticati da un intervento di screening



## Componenti di un programma di screening

- Caratteristiche della malattia
- Caratteristiche delle procedure di screening
- Efficacia degli interventi terapeutici



**Table 2. Points about Screening for Lung Cancer to Share with Patients.**

No data are available from randomized trials, which are ongoing; results are expected in four to five years.

Results from observational studies of CT screening among high-risk patients (i.e., those with a history of heavy smoking) indicate a high rate of diagnosis of lung cancer in stage 1 (a relatively curable stage).

CT screening reveals many noncalcified nodules, only a fraction of which will be found to be lung cancer.

Costly invasive procedures that are associated with serious risks may be required to evaluate some nodules.

A diagnostic workup should be done by physicians experienced in such evaluation.

The selection of a facility with physicians who are experienced and credentialed in multidisciplinary fields (including thoracic surgery, pathology, and pulmonology) is critical to an optimal outcome.

The most effective way for smokers to improve their health is to stop smoking.

There is an increased risk of subsequent lung cancers after curative resection of lung cancer, so ongoing surveillance is essential.

Screening-management trials are available for the evaluation of CT screening.



**Table 1.** Summary of Clinical Trials for Fecal Occult Blood Testing

Study Characteristics	Mandel et al <sup>12,13</sup> (United States)	Hardcastle et al <sup>14</sup> (United Kingdom)	Kronborg et al <sup>16,17</sup> (Denmark)
No. of study participants	46 551	150 251	61 933
Follow-up, y	18	7.8	13
Relative risk mortality with annual FOBT (95% CI)	.67 (0.51-0.83)	Not studied	Not studied
Relative risk mortality with biennial FOBT (95% CI)	.79 (0.62-0.97)	.85 (0.74-0.98)	.82 (0.69-0.97)
Absolute risk reduction for CRC death per 1000 subjects	4.6 (annual) 2.9 (biennial)	0.8	1.8
No. of subjects needed to screen to prevent CRC death	217 (annual) 344 (biennial)	1250	555

Abbreviations: CI, confidence interval; CRC, colorectal cancer; FOBT, fecal occult blood testing; RR, relative risk.

