

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 (κ)
- 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

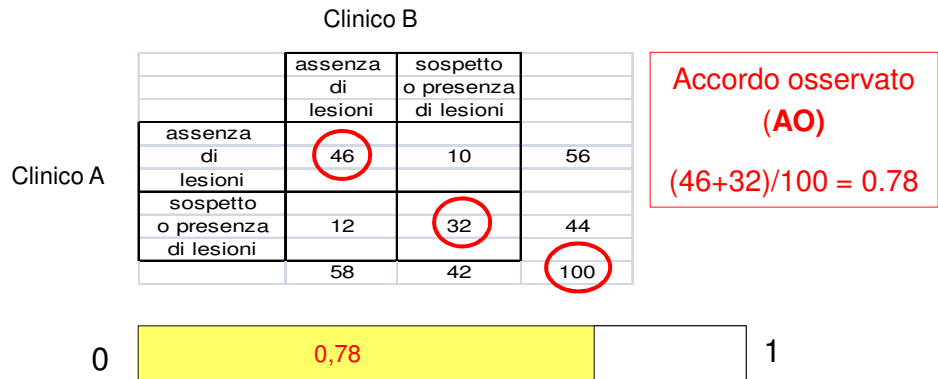
I metodi per la misura della diagnosi

Come misuriamo la riproducibilità di un test diagnostico?



Riproducibilità del test diagnostico

L'accordo tra osservatori

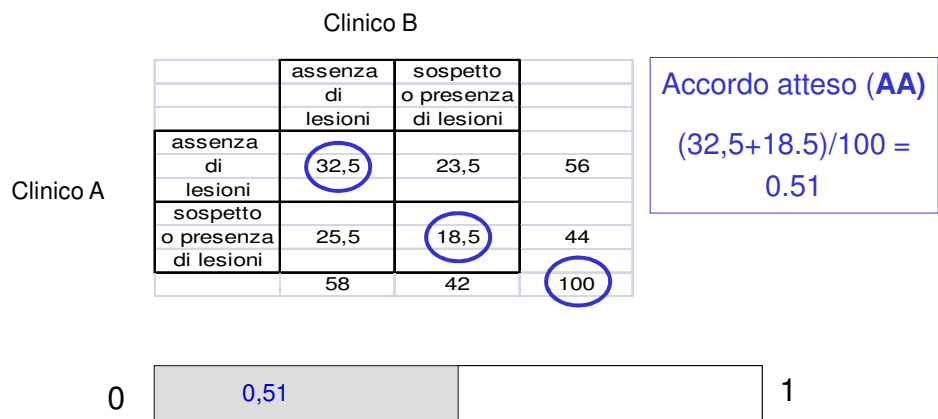


Tale misura non considera però il fatto che i due clinici potrebbero concordare nella diagnosi per effetto del caso



Riproducibilità del test diagnostico

Se la concordanza fosse solo casuale



Accordo atteso per effetto del caso



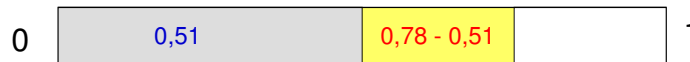
Riproducibilità del test diagnostico

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 (κ)

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 ^a	Should this woman be referred for surgical assessment?	Two radiologists examined the same 1214 mammograms	97%	91%	.67
Exercise electrocardiogram (ECG) ^b	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 ^c	Does this patient have iron deficiency?	Two pathologists examined the same films from 29 women with iron-deficiency anemia ^d	69%	49%	.39
		One pathologist twice examined the same films from 29 women with iron-deficiency anemia ^d	83%	51%	.65

^aData 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.

^bData 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.

^cData 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.

^dHemoglobin level 11.6 g/dl or less and serum iron level less than 12.5 µmol/L (70 mg/dl).

Objectives. To compare the sensitivity, specificity, and interobserver reliability of conventional cervical smear tests, monolayer cytology, and human papillomavirus testing for screening for cervical cancer.

Design. Cross sectional study in which the three techniques were performed simultaneously with a reference standard (colposcopy and histology).

Setting. Public university and private practices in France

Participants. 828 women referred for colposcopy because of previously detected cytological abnormalities and 1757 women attending for routine smears.

Main outcome measures. Clinical readings and optimised interpretation (two blind readings followed, if necessary, by consensus). Sensitivity, specificity, and weighted κ computed for various thresholds of abnormalities.



First interpretation	Second interpretation				
	Normal	ASCUS/AGUS	LSIL	HSIL	Invasive cancer
Cervical smear					
Normal	448	35	15	13	0
ASCUS/AGUS	13	25	14	4	0
LSIL	34	11	33	8	0
HSIL	6	4	4	86	0
Invasive cancer	0	0	0	0	8
Monolayer cytology					
Normal	442	33	14	8	0
ASCUS/AGUS	29	11	10	6	0
LSIL	48	10	28	10	0
HSIL	16	10	22	65	0
Invasive cancer	0	0	0	0	2

Interobserver agreement was good for conventional smears (weighted κ 0.69, 95% CI 0.64 to 0.74) but only moderate for monolayers (0.57, 0.52 to 0.63).



C.I. di Metodologia clinica

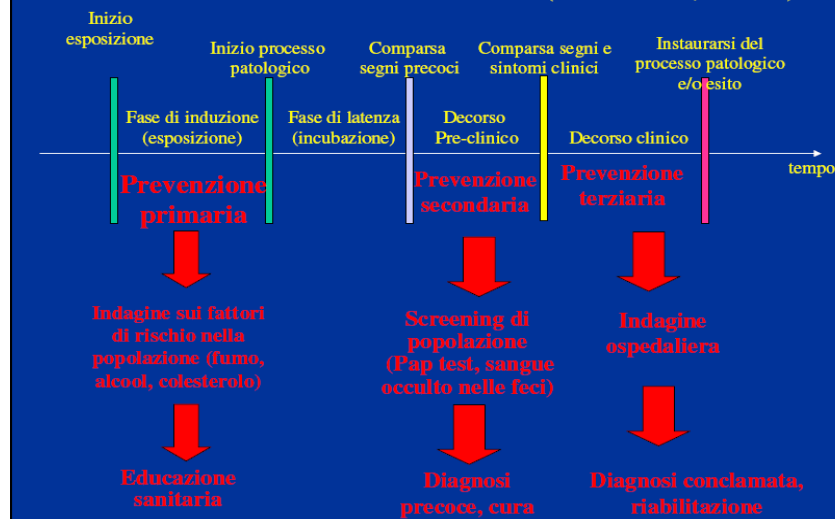
I metodi per la misura della diagnosi

Che cos'è un test di screening?



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



The NEW ENGLAND
JOURNAL of MEDICINE

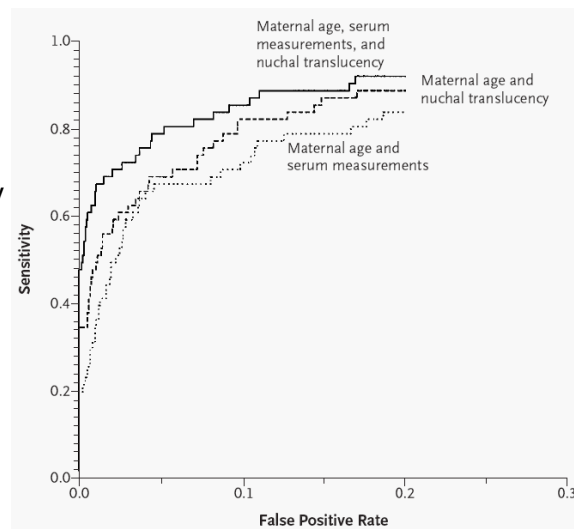
First-Trimester Screening for Trisomies 21 and 18

N Engl J Med 2003;349:1405-13.

Background

Screening for aneuploid pregnancies is routinely performed after 15 weeks of gestation and has a sensitivity of approximately 65 percent, with a false positive rate of 5 percent.

First-trimester markers of aneuploidy have been developed, but their use in combination has not been adequately evaluated in clinical practice.



Clinical questions

- What is the best colorectal cancer screening option?
- What are the recommendations of professional organizations?
- What is the quality of the evidence for each screening test?
- What are the sensitivity and specificity of colorectal screening tests?
- What are the risks and discomforts of colorectal screening test?
- Do patient preference and adherence vary among the tests?
- Do the screening tests differ with respect to their availability because of either differing resources or health insurance coverage?
- How is the best screening test chosen?



Clinical questions

- Should colorectal cancer screening recommendations be modified for individuals at higher than average risk for colorectal cancer?
- In the face of conflicting recommendations, how should physicians decide about screening individuals with a family history of colorectal cancer?
- What is the current adherence and what are the barriers to screening?
- What strategies can be implemented to improve rates of colorectal cancer screening?
- Is colorectal cancer screening cost-effective?



Screening

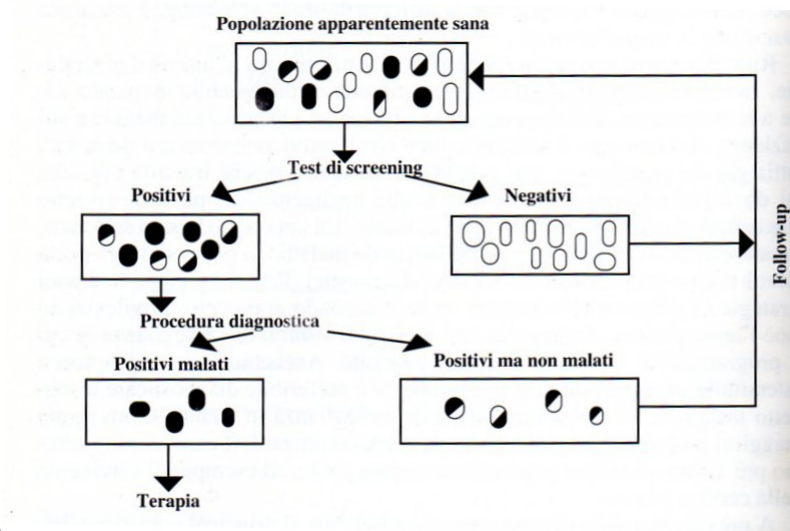
Mira a riconoscere, all'interno di una popolazione apparentemente sana, gli individui probabilmente malati, con l'intento di modificarne in senso favorevole il decorso clinico.

- 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



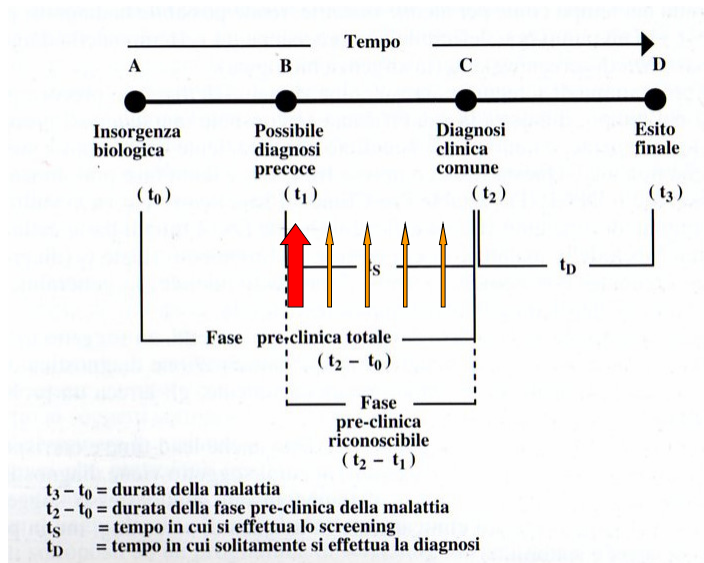
Obiettivo dello screening

Attività: Anticipare la diagnosi

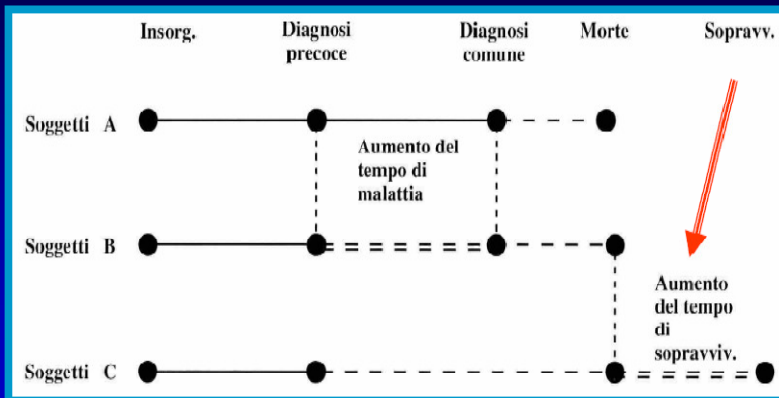
Efficacia: Ridurre la mortalità



Screening e storia naturale della malattia



Valutazione dell'efficacia di un programma di screening



Obiettivo dello screening

Attività: Anticipare la diagnosi

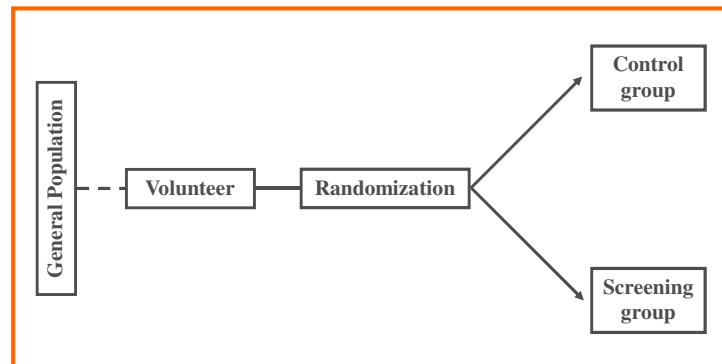
Efficacia: Ridurre la mortalità

Problemi

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

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

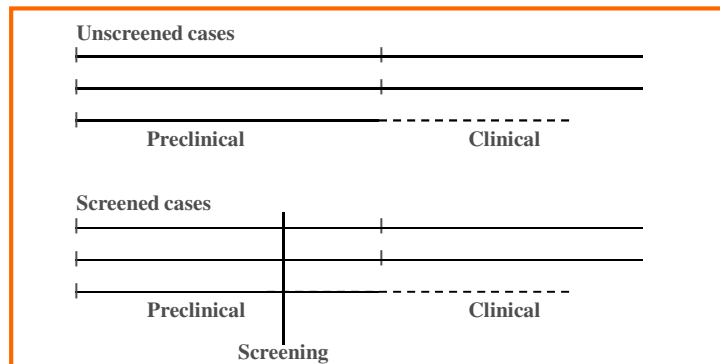
Problemi

- 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

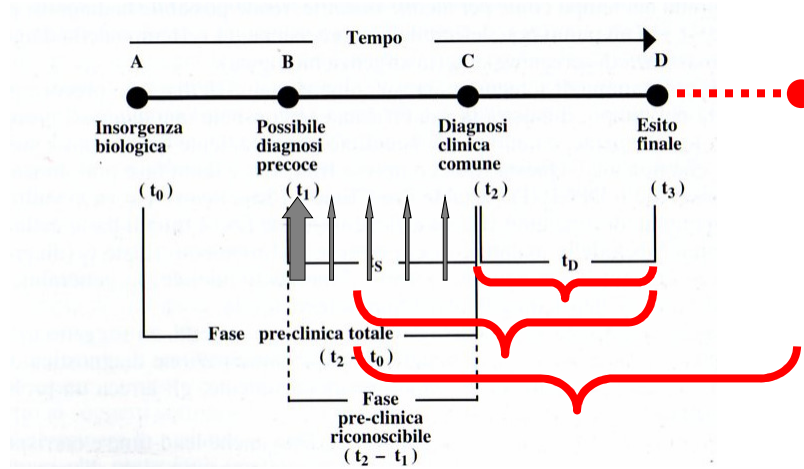
Efficacia: Ridurre la mortalità

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

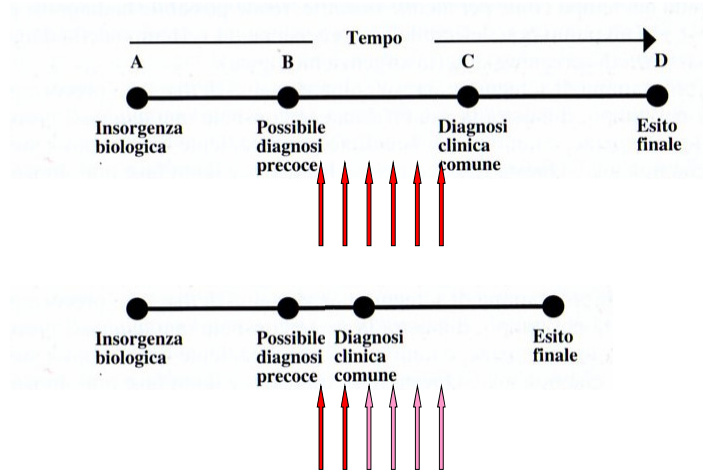
Problemi

- 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



Componenti di un programma di screening

- **Caratteristiche della malattia**
 - ❑ Caratteristiche della fase preclinica riconoscibile
 - ❑ Prevalenza della malattia nella popolazione
 - ❑ Evoluzione della malattia
 - ❑ Identificabilità della popolazione obiettivo
- Caratteristiche delle procedure di screening
- Efficacia degli interventi terapeutici



Componenti di un programma di screening

- Caratteristiche della malattia
- **Caratteristiche delle procedure di screening**
 - ❑ Proprietà (sensibilità, specificità, predittività, riproducibilità)
 - ❑ Fattibilità (periodicità, semplicità, innocuità, economicità)
 - ❑ Disponibilità degli esami diagnostici di secondo livello
- Efficacia degli interventi terapeutici



Componenti di un programma di screening

- Caratteristiche della malattia
- Caratteristiche delle procedure di screening
- **Efficacia degli interventi terapeutici**
 - **Endpoint primari (di effetto)**
 - Aumento della sopravvivenza in studi sperimentali
 - Riduzione della mortalità specifica
 - Riduzione dell'incidenza di forma avanzate
 - **Endpoint secondari (di processo)**
 - Entità e omogeneità di partecipazione al programma
 - Proporzione di casi screenati sul totale
 - Distribuzione dello stadio di malattia alla diagnosi
 - Proporzione di terapie invalidanti sul totale



CLINICAL PRACTICE

A 60-year-old woman who quit smoking 20 years earlier comes for a routine visit. She previously smoked one pack of cigarettes a day for 10 years. Her medical history is otherwise unremarkable. She feels well and exercises regularly. Her husband smoked one pack of cigarettes per day for at least 30 years but stopped smoking a decade ago.

She asks whether she and her husband should undergo computed tomographic (CT) scanning to screen for lung cancer. What do you advise?



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.

