

Epidemiologi og biostatistik

Diagnostik og screening

Forelæsning, uge 5, Svend Juul

Hvordan stiller man en diagnose?

Symptomer

passive: patientens spontane rapport

aktive: svar på målrettede spørgsmål

Fund

Kliniske fund

passive: ernæringstilstand, cyanose

aktive: patellar-refleks

Fysiske fund

passive: vægt, højde

aktive: lungefunktionsmål

Biokemiske fund

Billeddiagnostiske fund

Pato-anatomiske fund

At nå frem til en diagnose svarer til at afprøve en række hypoteser, forkaste de fleste og acceptere én.

Den erfarne læge er effektiv i sin udvælgelse af de hypoteser der skal afprøves.

Diagnoser er ofte behæftet med usikkerhed – også den erfarne læges.

Mange patienter fejler mere end én ting.

Hvorfor er diagnoser nyttige?

Diagnoser er skabt af mennesker.

Ikke to patienter er ens – men nogle har fælles træk.

Diagnosen er et økonomisk middel i kommunikationen om patienter.

Diagnoser strukturerer vores kollektive viden.

Studier af ætiologi mhp. forebyggelse.

Studier af behandlingseffekt.

Hvornår er korrekt diagnose særligt nyttig?

Når vi kender en behandling der gør en forskel.

Diagnostiske kriterier

Centrum:

Hvad er de centrale karakteristika for sygdommen?

Periferi:

Hvordan afgrænses sygdommen over for andre sygdomme – og over for "rask"?

Diagnoser er:

- skabt af mennesker
- behæftet med usikkerhed
- en forenklet beskrivelse af patientens problem
- nyttig til kommunikation
- nyttig til at strukturere den kollektive faglige viden
- ikke evige

Diagnostisk test: Mammografi

Mammografi indførtes først som et redskab i diagnosen af brystkræft blandt kvinder hvor der allerede var begrundet mistanke.

Testen indebærer en trænet observatørs vurdering.

Konklusionen kan have formen:

- brystkræft (positiv test)
- ikke brystkræft (negativ test)
- usikkert resultat

På en stor specialafdeling undersøgtes med mammografi 1000 kvinder, henvist pga. mistanke om brystkræft. 40% af de henviste havde faktisk brystkræft.

Test	Syg		I alt	P(syg)
	+	0		
+	360	30	390	92%
0	40	570	610	7%
I alt	400	600	1000	40%

Prævalens af brystkræft: 0,40

Sensitivitet = 0,90 (90% af de syge diagnosticeres korrekt)

Specificitet = 0,95 (95% af de raske diagnosticeres korrekt)

Prædiktiv værdi af positiv test: $360/390 = 0,92$

Prædiktiv værdi af negativ test: $40/610 = 0,07$ (0,93)

Der er tale om *estimer*.

Estimeret sensitivitet: $360/400 = 90\%$ (87%; 93%)

Estimeret specificitet: $570/600 = 95\%$ (93%; 97%)

Jeg beskrev prædiktiv værdi af negativ test som:

$$PV- = FN/(FN+SN) = 7\%$$

Den beskrives ofte som

$$PV- = SN/(FN+SN) = 93\%$$

Mere generelt:

Test	Syg		I alt	Prædiktiv værdi
	+	0		
+	SP	FP	SP+FP	$SP/(SP+FP)$
0	FN	SN	FN+SN	$FN/(FN+SN)$
I alt	SP+FN	SN+FP		

Testen er ikke perfekt. Men den kan forbedres:

- ved forbedret teknologi
- ved uddannelse af bedømmerne
- ved løbende kvalitetskontrol

Men hvad med usikre bedømmelser?

- Hvis vi regner dem som test-positive, styrker vi sensitiviteten, men svækker specificiteten.
- Hvis vi regner dem som test-negative, styrker vi specificiteten, men svækker sensitiviteten.
- Valget må afhænge af konsekvenserne.

Reference-interval

Karakteristik af testresultaternes fordeling blandt ikke-syge.

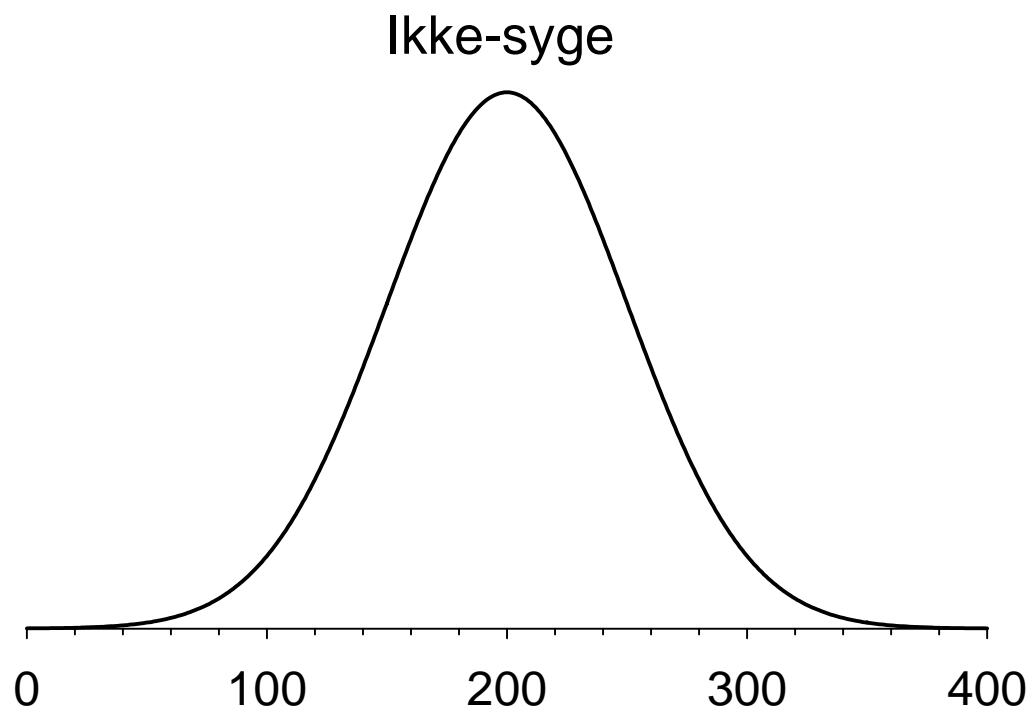
F.eks. plasma-kreatinin: Mænd: 62-133 $\mu\text{mol/l}$

Kvinder: 44-115 $\mu\text{mol/l}$

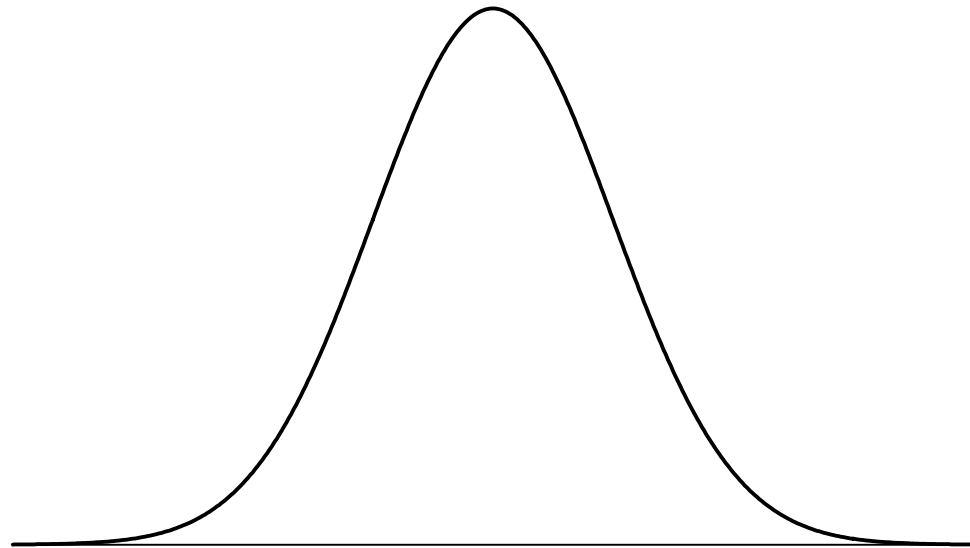
Reference-intervallet bestemmes typisk som et 95% prædiktions-interval i en rask population.

Reference-intervallets grænser er IKKE de værdier (cut-points) der adskiller de syge fra de raske.

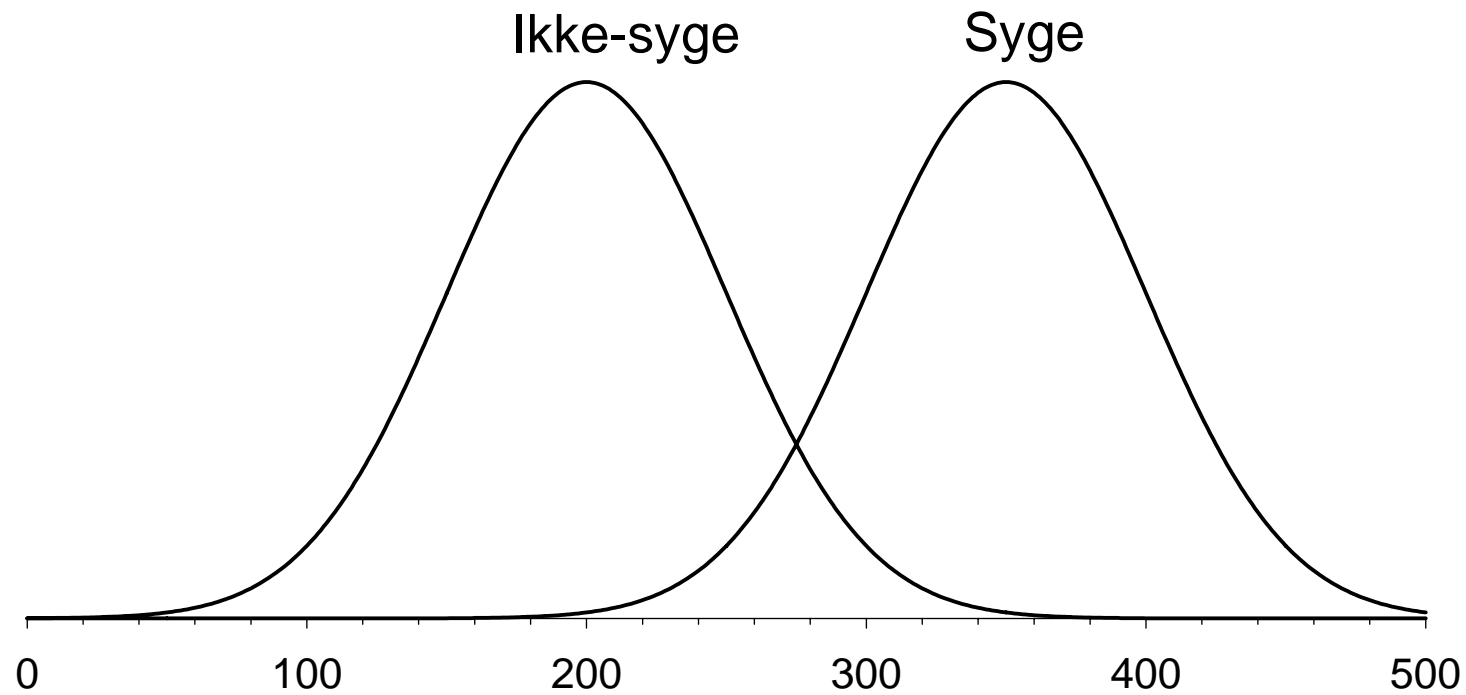
Betegnelsen "normal-værdier" er uheldig.



Syge



Tests med kontinuert udfald



Ved fastlæggelse af cutpoint i en konkret diagnostisk situation må man tage hensyn til konsekvenserne af falsk negative og falsk positive konklusioner.

Alvorlig konsekvens af en falsk negativ konklusion:

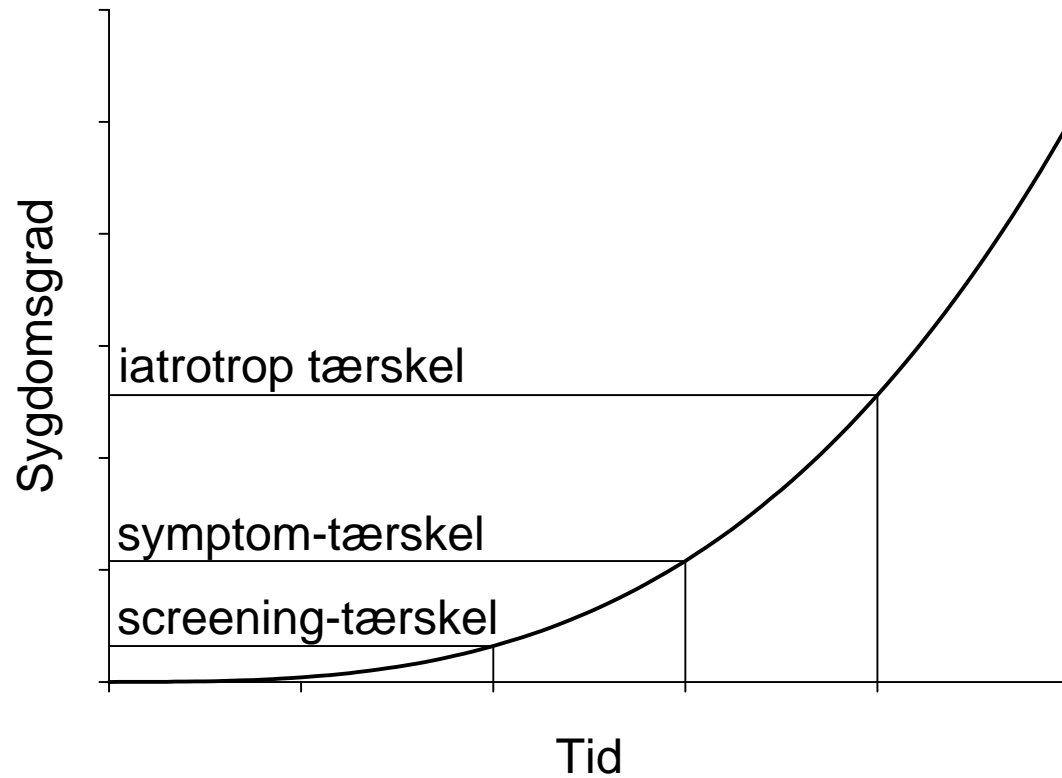
- Hvis en behandlelig sygdom med alvorlig prognose overses

Alvorlig konsekvens af en falsk positiv konklusion:

- Hvis der iværksættes en belastende behandling uden sundhedsgevinst.

Screening

Model for udvikling af en kræftsygdom



Formål med screening:

1. at reducere dødelighed
2. at reducere lidelse og funktionstab
 - ved helbredelse
 - ved mindre indgreb
- (3. at opnå tryghed)

Midler:

1. at finde sygdomstilfælde i tidligere stadier end ellers
2. at behandle de fundne tilfælde

Forudsætninger for sundhedsgevinst:

1. at vi kan finde tilfælde tidligere end ellers (testens sensitivitet)
2. at tidlig behandling bedrer prognosen
3. Netto-sundhedsgevinst: Positive sundhedseffekter overstiger negative.
 - a) risiko ved test
 - b) overbehandling
 - c) bivirkninger ved behandling

Andre velfærdseffekter:

1. Tryghed
2. Utryghed

Økonomiske hensyn:

Omkostningerne skal være "rimelige" i forhold til sundhedsgevinsten.

Meromkostningerne skal vurderes: Flere udgifter til testning, færre til behandling.

Cost-effectiveness analyse:

Sundhedsgevinst/meromkostninger

F.eks.: 50.000 kr. pr. vundet leveår

Problemer ved implementering

1. Acceptabilitet: god deltagelse
2. Kriterier for opfølgning og behandling skal være velbeskrevne – og accepteret af aktørerne
3. Kapacitet til opfølgning og behandling
4. Administrative problemer: Sikre, at opfølgning faktisk sker

Mammografi i klinisk diagnostik.

Test	Syg		I alt	P(syg)
	+	0		
+	360	30	390	92%
0	40	570	610	7%
I alt	400	600	1000	40%

Prævalens af brystkræft: 0,40

Sensitivitet = 0,90 (90% af de syge diagnosticeres korrekt)

Specificitet = 0,95 (95% af de raske diagnosticeres korrekt)

Prædiktiv værdi af positiv test: $360/390 = 0,92$

Prædiktiv værdi af negativ test: $40/610 = 0,07$ (0,93)

En ikke så god test til screening

Test	Syg		I alt	P(syg)
	+	0		
+	90	4.995	5.085	1,8%
0	10	94.905	94.915	0,01%
I alt	100	99.900	100.000	0,1%

Prævalens af brystkræft: 0,001

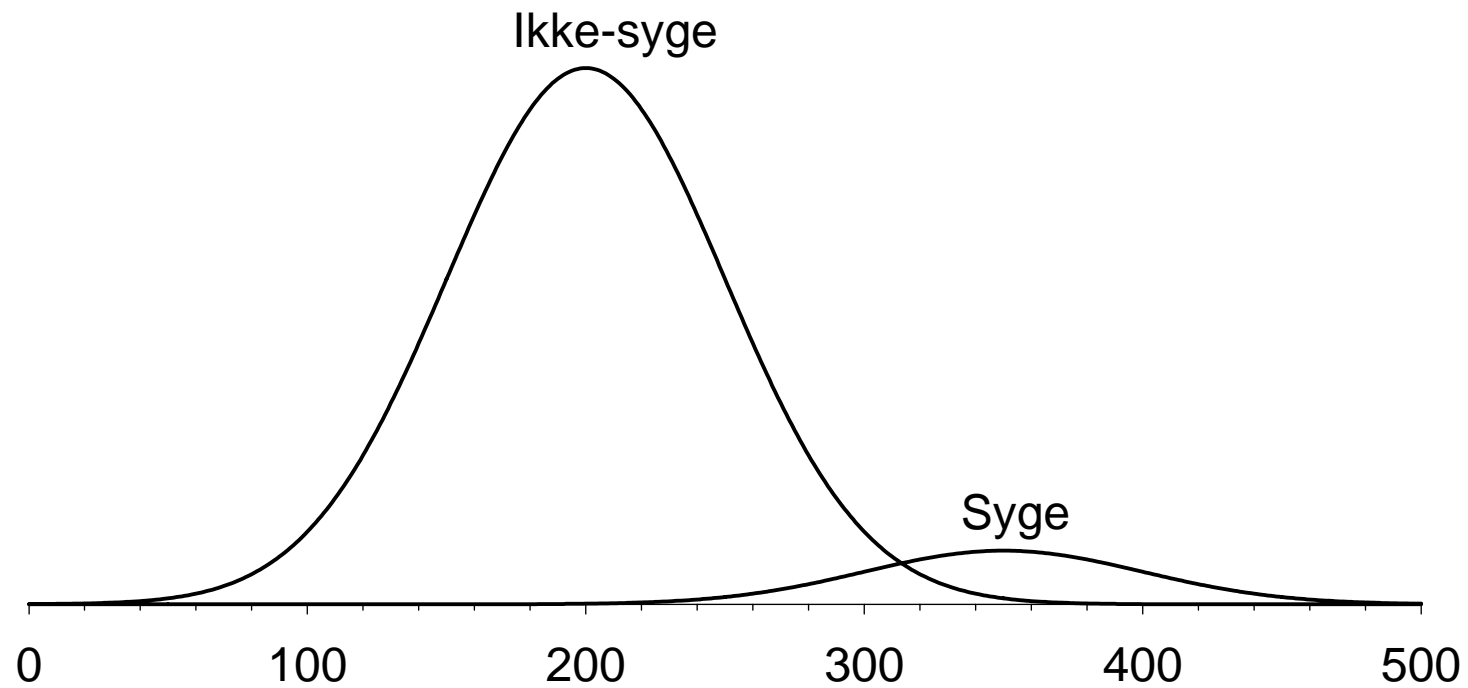
Sensitivitet = 0,90 (90% af de syge diagnosticeres korrekt)

Specificitet = 0,95 (95% af de raske diagnosticeres korrekt)

Prædiktiv værdi af positiv test: $90/5.085 = 0,018$

Prædiktiv værdi af negativ test: $10/94.915 = 0,0001$

Valg af cut-point ved screening



En tests egenskaber kan beskrives ved sensitivitet og specificitet.

Fortolkningen af et testresultat forudsætter à priori viden eller formodning: sandsynligheden for sygdom.

I screening-situationen er à priori sandsynligheden for sygdom lav.

Derfor mange falsk positive testresultater.

– med mindre specificiteten er meget fin.

Kan man overføre erfaringerne om en tests sensitivitet og specificitet direkte fra klinikken til screening-situationen?

Ved screening for cancer stiles efter at sygdomsprocessen erkendes tidligt, mens sygdomsmanifestationerne er beskedne.

En test vil være mindre sensitiv mht. tidlige sygdomsstadier.

Der indføres et screening-program for en cancerform.

De tilfælde, der findes ved screening, har længere gennemsnitlig overlevelse end de tilfælde, der findes på sædvanlig måde.

- Det virker altså . . ?

De lever længere **efter at diagnosen er stillet**, men er der opnået:

- en sundhedsgevinst?
- en velfærdsgevinst?

Selv om behandlingen er helt uden indflydelse på overlevelsen, vil tidlig diagnose bevirke længere overlevelse efter diagnosetidspunktet.

Et screeningprograms sundhedseffekt kan ikke vurderes alene på forløbet af de diagnosticerede tilfælde.

Effekten må vurderes på befolkningsniveau.

Ikke-eksperimentelle metoder:

Geografiske sammenligninger mellem områder med og uden screeningprogrammer.

Forudsætning: Hvis der ikke var screening, ville incidensen være ens.

Tidsrække-analyser.

Forudsætning: Hvis der ikke var screening, ville incidensen være konstant.

Effekt-kriteriet er reduktion i **befolkningens** dødelighed (mortalitet), ikke blot prognosen for fundne tilfælde (letalitet).

Eksperimentelle metoder:

Det kontrollerede, randomiserede forsøg, hvor nogle tilbydes, andre ikke tilbydes screening.

Mortaliteten i interventions- og kontrolgruppe sammenlignes.

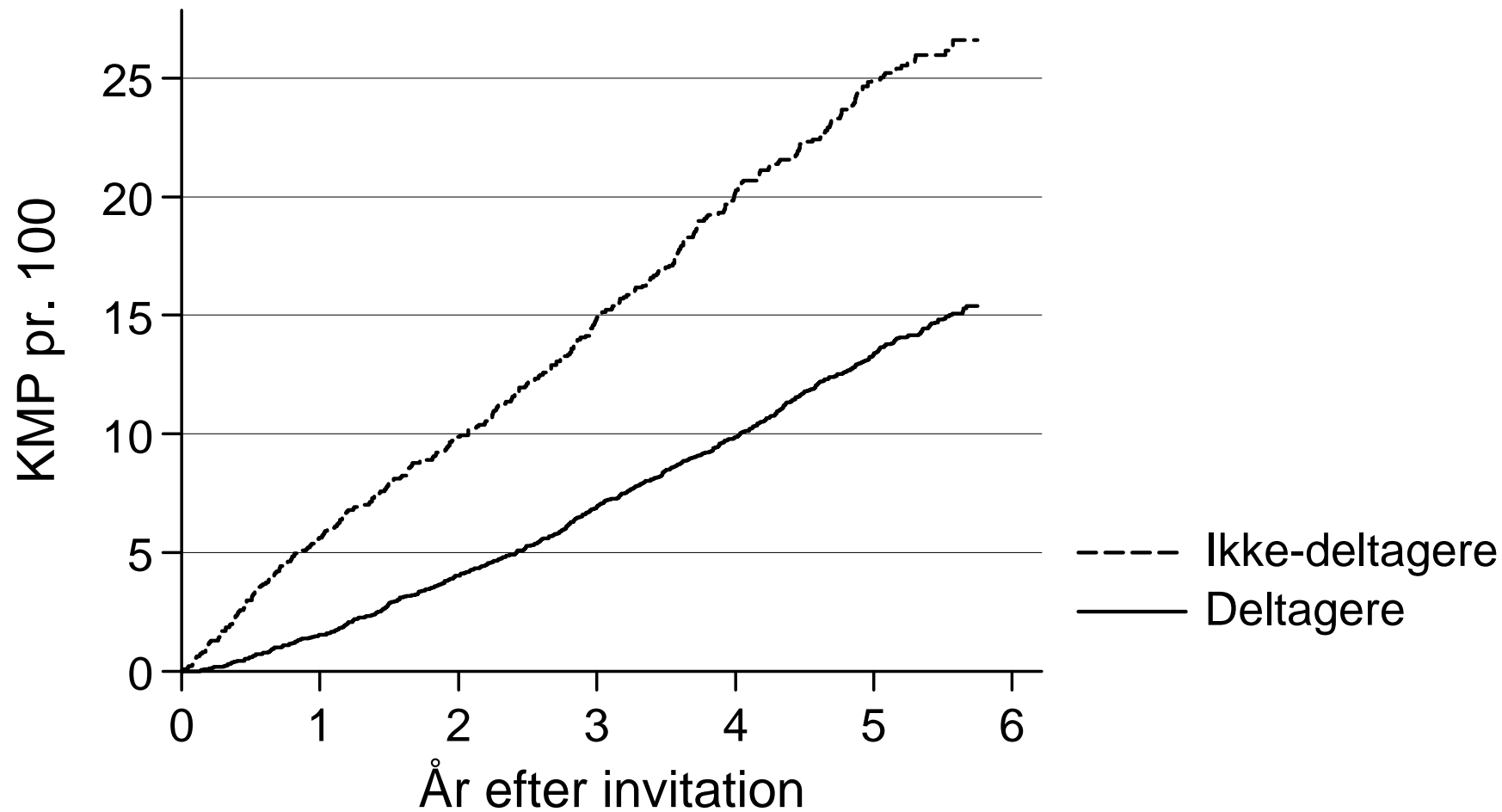
Der forudsættes typisk mange personer og lang tidshorisont.

Da programmet er så synligt, kan det være svært at opnå accept af randomisering.

Intention-to-treat princippet i randomiserede forsøg.

Randomiseringen skal sikre, at de to grupper er ens i udgangspunktet.

Interventionsgruppe	Kontrolgruppe
Deltagere	
Ikke-deltagere	



Screening for prostatakkræft

1. Sygdommen er ofte længe om at udvikle sig. Der kan altså være god mulighed for at finde den "i tide".
2. Prostata-specifikt antigen (PSA) er forhøjet ved prostatacancer (blodprøve).
3. Det er vist at radikal prostatektomi ved tilfælde af prostatacancer som er diagnosticeret "i tide" bedrer overlevelsen.

Holmberg et al.: 8 års dødelighed blandt 695 patienter med prostatakkræft, randomiseret til radikal prostatektomi eller afventende strategi.

	Operation (N=347)		Afvente (N=348)		RD pr. 100 (95% CI)
	Antal	KMP 8 år	Antal	KMP 8 år	
p.c. dødsf.	16	7,1%	31	13,6%	6,6 (2,1; 11,1)

NNT: 15 (9; 48)

Det må da være en god idé at screene!

Problemer

1. PSA-testens sensitivitet er estimeret til 70% og specificiteten til 50%. Det giver mange falsk positive fund.
2. Udredningen af positive fund sker ved transrektal biopsi. Risiko: Blødning, infektion.
3. Mange af de sandt positive tilfælde vil forløbe så langsomt at de ikke vil give symptomer i personens levetid. Ved autopsi af mænd > 60 år: 30% har prostatacancer.

4. Behandlingen (radikal prostatektomi, strålebehandling) har væsentlige bivirkninger:

Til 376 svenske deltagere, som var i live 3-4 år efter randomiseringen, sendtes et spørgeskema. Svar fra 86% hhv. 88%.

	Operation (166)	Afvente (160)	RR (95% CI)
Erektile dysfunktion	80%	45%	1,8 (1,5; 2,2)
Samleje < 1/md.	80%	59%	1,4 (1,2; 1,6)
Ked af sex-problemer	56%	40%	1,4 (1,1; 1,8)
Urine: obstruction	32%	38%	0,8 (0,6; 1,1)
Urine: leakage	49%	21%	2,3 (1,6; 3,2)

1000 65 årige mænd screenes:

PSA	Sygdom		I alt PV
	Ja	Nej	
> 4 $\mu\text{mol/l}$	140	400	540 26%
\leq 4 $\mu\text{mol/l}$	60	400	460 13%
Alle	200	800	1000 20%

Sens: 70% Spec: 50%

540 biopseres. Ca. 1% (5-6) får alvorlig blødning eller infektion.

175 får tilbudt radikal operation.

175 får tilbudt radikal operation. Det vil skønsmæssigt medføre:

	%	Antal
Potens-problemer	35%	61
Urin-inkontinens	28%	49

Hvor mange liv redder vi?

Vi ved det ikke. Holmbergs undersøgelse vedrørte tilfælde diagnosticeret pga. symptomer. Dødsrisikoen blandt før-symptomatiske tilfælde er givetvis lavere. Vi kan slet ikke overføre Holmberg's NNT på 15 til screening-situationen.

Men vi ved at vi skaber problemer for mange.

American Cancer Society og American Urological Association anbefaler:

Alle mænd fra 50 års alderen tilbydes årlig screening med rektaleksploration og PSA-test. Men da evidensen er mangelfuld:

"Information should be provided to men about the benefits and limitations of testing so that an informed decision about testing can be made with the clinician's assistance"

Politisk-etiske problemer:

Er screening ikke blot endnu et forsøg fra det medico-industrielle kompleks på at øge sit marked?

Ved "normale" screening-programmer kommer et positivt fund personen til gode.

Hvad med fosterdiagnostik, hvor et positivt fund fører til abortering?

Screening af arbejdstagere:

Til glæde for hvem?