

Kursus i Epidemiologi og Biostatistik, forår 2003.

Epidemiologiske mål Studiedesign

Svend Juul, 3.2.2003

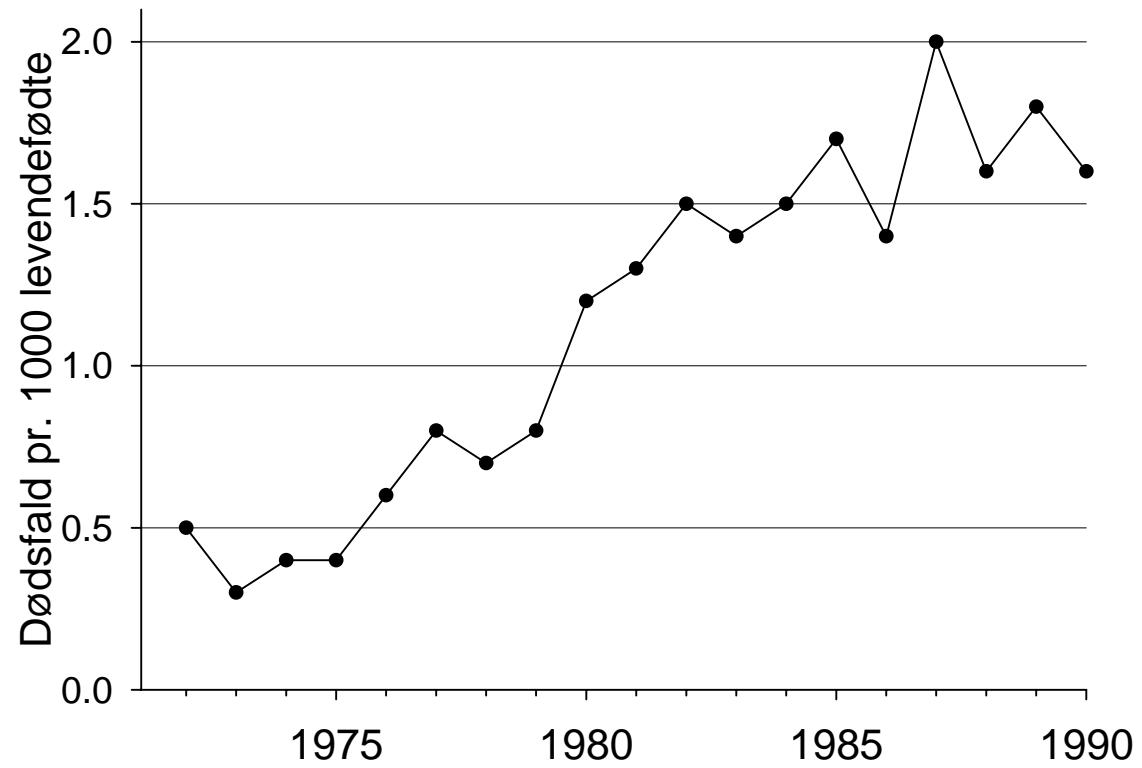
Pludselig uventet spædbarnsdød (vuggedød, Sudden Infant Death Syndrome, SIDS)

Uventet dødsfald hos et rask spædbarn (8 dage - 1 år).
Obduktion o.a. giver ingen forklaring.

Hyppigheden næsten 4-doblet fra først i 1970'erne til sidst i 1980'erne; ca. 100 dødsfald årligt.

Samme udvikling i en række andre lande.

Vuggedødsfald i Danmark 1972-1990



Januar 1991:

Case-kontrol undersøgelse fra New Zealand: Forældre til vuggedøde børn blev spurgt om mange ting; en af dem var barnets sædvanlige sovestilling.

En 'tilfældig' gruppe forældre til raske børn blev spurgt om det samme.

Undersøgelsen tydede på en 5-6 gange øget risiko for vuggedød blandt børn, der sov i maveleje.
(OR = 5,7; 95% CI: 3,3; 10,1).

Maj 1991:

Follow-up undersøgelse fra Australien: Ved 1 månedes alderen blev forældrene til mange børn spurgt om sædvanlig sovestilling.

Børn, der sov på maven, havde 4-5 gange øget risiko for senere vuggedød.

(OR = 4,5; 95% CI: 1,3; 15).

Uddrag af Sundhedsstyrelsens vejledninger

1961	Læg barnet skiftevis på højre og venstre side og på ryggen
1970	Læg barnet skiftevis på højre og venstre side, men ikke på ryggen
1980	Små børn skal ligge på maven eller siden, aldrig på ryggen
1990	Læg barnet skiftevis på maven og på siden, så evt. gylp kan løbe ud

Evidens:

1. To undersøgelser viser sammenhæng mellem sovestilling og dødelighed
2. Stigningen i dødelighed følger ændrede anbefalinger om sovestilling

Der er tale om *evidens*, men ikke *bevis* for en årsagssammenhæng.

December 1991:

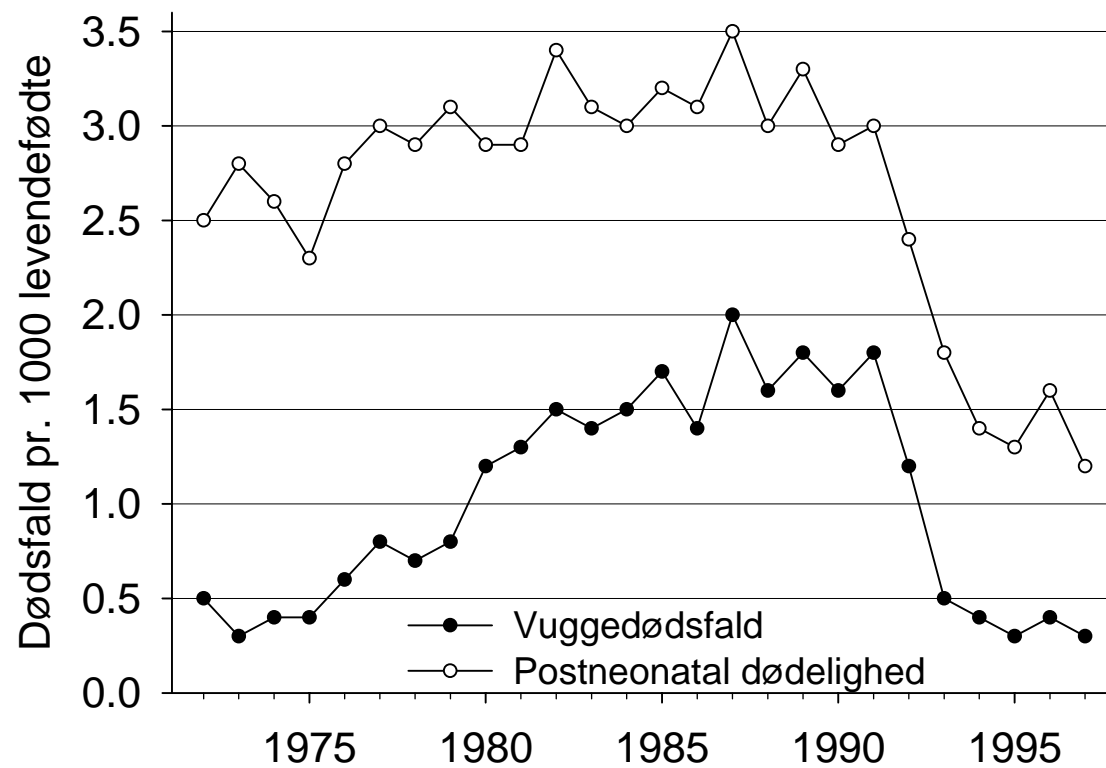
Sundhedsstyrelsen ændrer anbefalinger:

1992	Små børn, som ikke kan vende sig selv, må aldrig ligge på maven
------	---

Hvis:

1. der virkelig er tale om en årsagssammenhæng,
 2. de ændrede anbefalinger medfører ændret adfærd,
- skulle man herefter se et fald i dødeligheden.

Vuggedødsfald i Danmark 1972-1997.

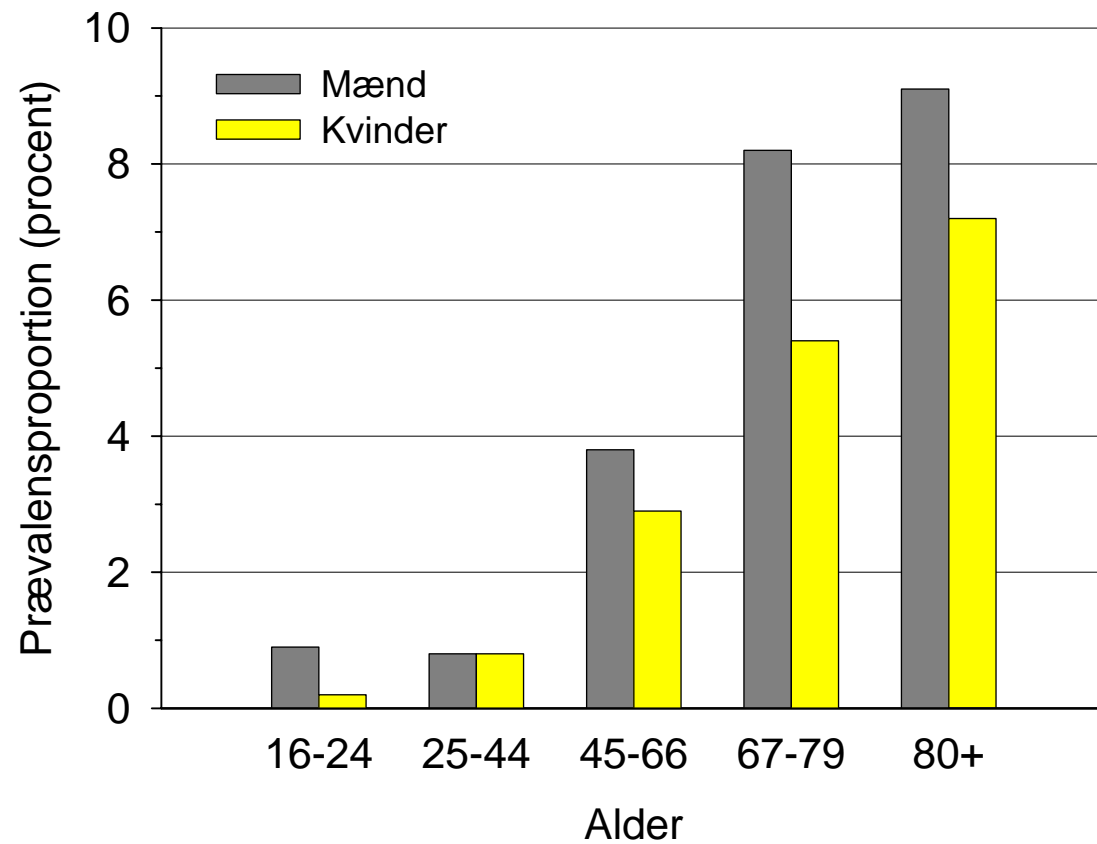


Mål for sygdomshyppighed

Incidens: hyppigheden af *begivenheder*

Prævalens: hyppigheden af en *tilstand*

Prævalens af selvrapporeret diabetes, Danmark 2000



Prævalensproportion:

$$PP = a/n = \frac{\text{Bestand af syge på et givet tidspunkt}}{\text{Studiepopulation på samme tidspunkt}}$$

Blandt 67-69 årige mænd var prævalensen af diabetes 8,2% (72 af 879 interviewede).

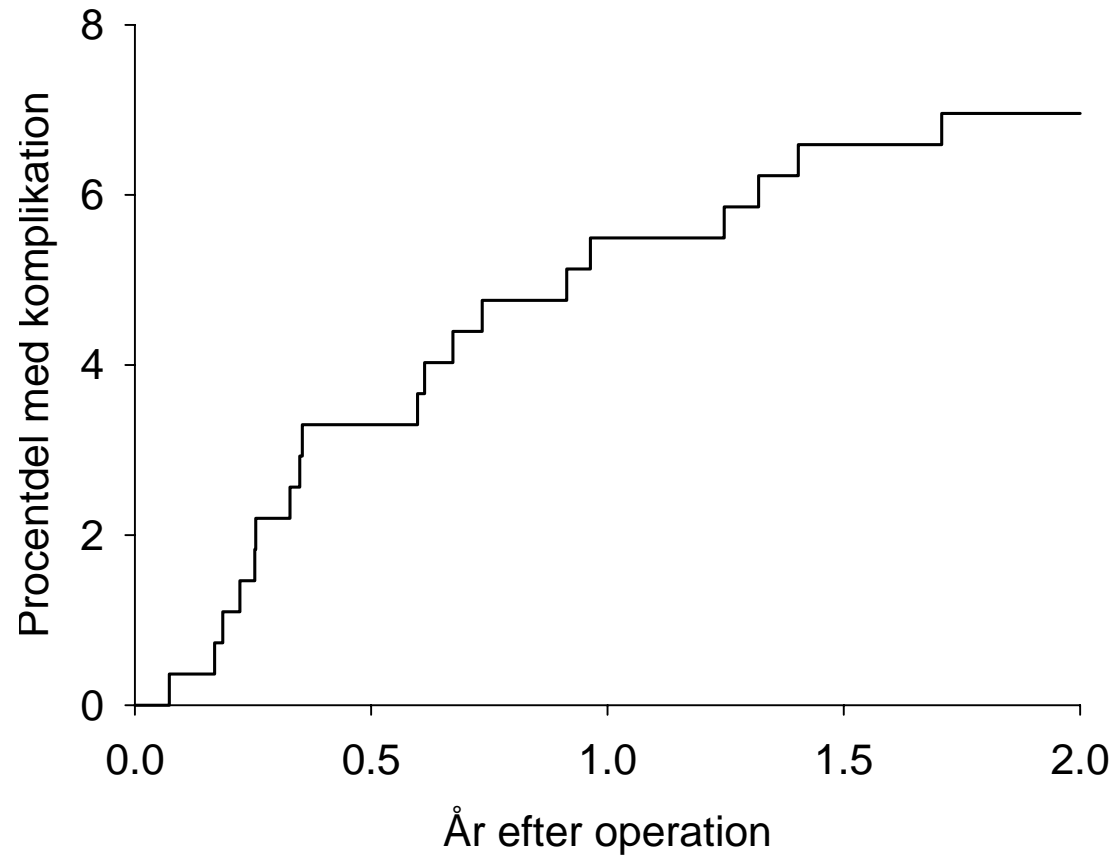
PP = 8,2% (95% CI: 6,4%; 10,0%)

Incidens: Hyppighed af begivenheder

Incidensproportion:

Andel af en *kohorte*, som udvikler sygdom gennem et givet tidsrum.

Udvikling i kumuleret incidensproportion



Incidence rate:

Den hastighed hvormed nye sygdomstilfælde indtræffer

$$IR = a/t = \frac{\text{Antal nye sygdomstilfælde}}{\text{Den tilsvarende risikotid}}$$

1978-80 diagnosticeredes 1.737 nye tilfælde af lungecancer blandt 55-64 årige mænd i Danmark. 1.7.1979 var der 263.787 mænd i aldersgruppen.

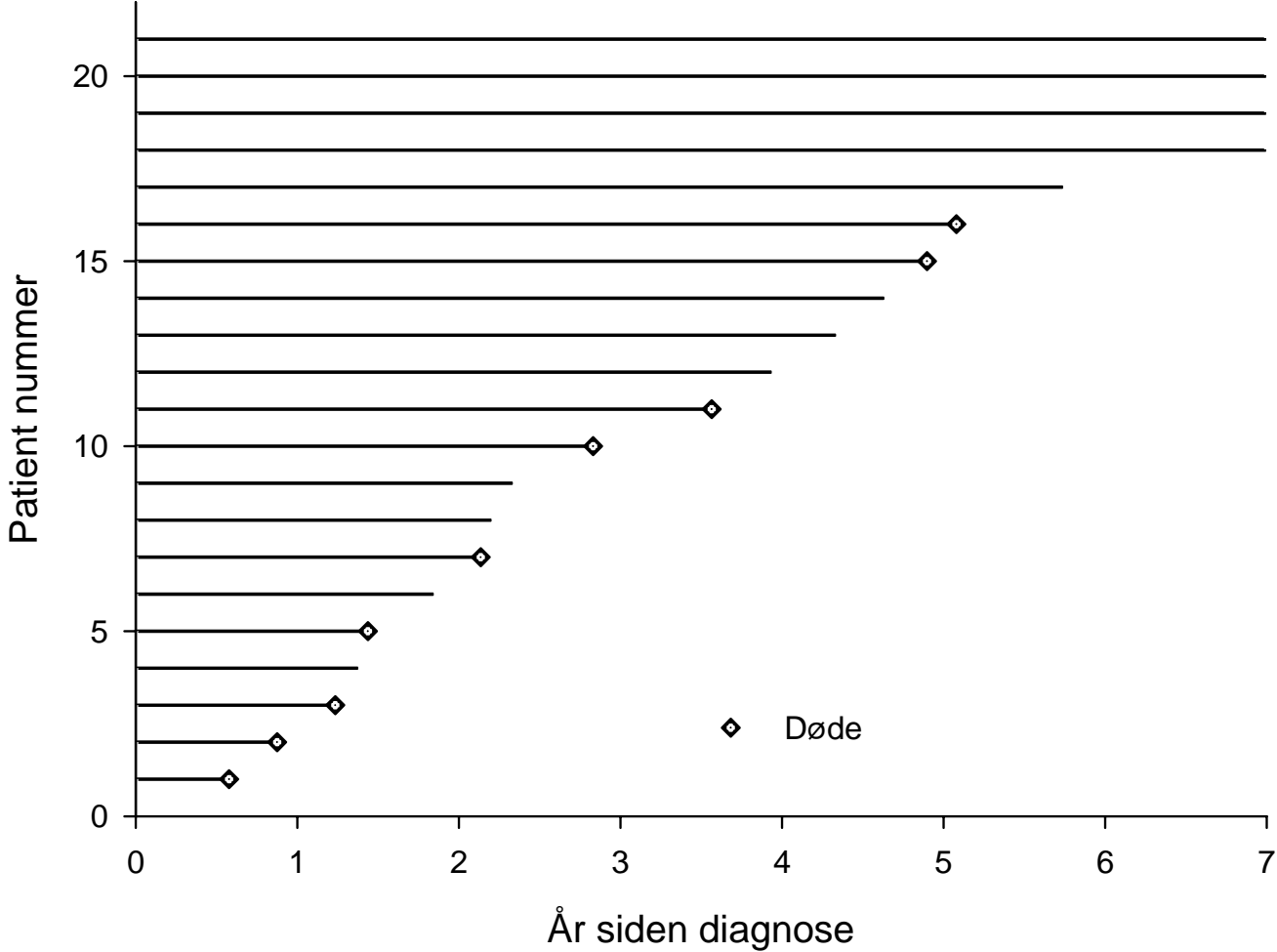
$$IR = \frac{1.737}{3 \text{ år} \times 263.783} = 0,0022/\text{år} = 2,2 \text{ pr. } 1000 \text{ år}$$

Danske mænd 55-64 år er en *åben population*.
Medlemmer tilkommer og forlader populationen.

En *kohorte* er en konkret persongruppe, som følges gennem et tidsrum, f.eks.

- danske mænd, som 1.1.1978 var 55-64 år,
- personer, som gennemgik en bestemt operation.

Risikotid i en kohorte



De første 3 år skete 6 dødsfald, mens for 4 personer observationen ophørte af anden grund (*censurering*). Den samlede risikotid for de første 3 år var 49,835 år, og incidensraten:

$$\text{IR} = \frac{6}{49,835 \text{ år}} = 0,120/\text{år} = 12,0 \text{ pr. } 100 \text{ år}$$

Sammenlignende undersøgelser

(Dwyer et al. Prospective cohort study of prone sleeping position and SIDS. Lancet 1991; 337: 1244-7)

For 2607 højrisiko-børn spurgtes ved 1 månedes alderen om sædvanlig sovestilling. Antal vuggedødsfald registreredes:

Sove-stilling	N	SIDS	Risiko pr. 1000	Risiko-differens	Relativ risiko
mave	9	821	10,8	7,4	3,2
andet	6	1.771	3,4		

RR = 3,2 (95% CI: 1,2; 9,0)

Design:	Follow-up
Studiepopulation:	Kohorte, højrisiko-børn
Tidshorisont:	11 mdr (fikseret)
Eksponering:	Sovestilling
Outcome:	SIDS-dødsfald
Associationsmål:	Relativ risiko

Ved kontrol for en række faktorer fandtes en endnu stærkere sammenhæng.

Rygning og osteoporose

30.000 mænd og kvinder oplyste ved tre befolkningsundersøgelser om deres rygevaner. Fra Landspatientregisteret indhentedes oplysning om hoftefrakturer i 1977-97.

Ryger	Antal tilfælde	Risikotid, år	IR pr. 1000 år	IRR	IRD pr. 1000 år
≥15/dag	139	99.742	1,39	1,68	0,56
Aldrig	30	36.181	0,83		

IRR = 1,68 (95% CI: 1,13; 2,49)

IRR justeret = 2,2 (95% CI: 1,5; 3,3)

Design:	Follow-up
Studiepopulation:	Kohorte
Tidshorisont:	5-20 år
Eksponering:	Rygning
Outcome:	Fraktur (i 1977-97)
Associationsmål:	Incidence-rate-ratio og -difference

Case-kontrol studier

Ved sjældne sygdomme skal man bruge en stor studiepopulation og vente længe, før der er så mange tilfælde (cases), at man kan beregne rimeligt sikre estimater. Genveje ønskes.

(Mitchell EA et al. Results from the first year of the New Zealand cot death study. NZ Med J 1991; 104: 71-6)

Blandt børn født 1987-88 registreredes i New Zealand 162 tilfælde af vuggedød; for 128 opnåedes interview. En kontrolgruppe blev trukket som en tilfældig stikprøve af børn født 1987-88.

Design:	Case-kontrol
Cases:	SIDS tilfælde blandt børn født 1987-88
Kontroller:	Stikprøve af børn født 1987-88
Eksponering:	Sovestilling
Associationsmål:	Odds ratio

Formål med kontrolgruppen: at opnå et estimat over eksponeringens hyppighed i den population, cases kommer fra (kildepopulationen).

Sovestilling	Cases	Kontroller
Maveleje	93 (73%)	216 (43%)
Andet	35 (27%)	287 (57%)
I alt	128 (100%)	503 (100%)

Maveleje var således væsentligt hyppigere i case-gruppen end blandt repræsentanterne for kildepopulationen (kontrollerne).

Den samlede studiepopulationen repræsenterer ingenting; den er skævt sammensat, og risikoen for SIDS kan ikke beregnes - ej heller relativ risiko.

Odds ratio er normalt et godt estimat af den relative risiko:

Eksponering	Cases	Kontroller	OR
+	a_1	b_1	$(a_1 \times b_0) / (a_0 \times b_1)$
0	a_0	b_0	

Sovestilling	Cases	Kontroller	OR
Maveleje	93	216	$(93 \times 287) / (35 \times 216) = 3,53$
Andet	35	287	
I alt	128	503	

OR = 3,53 (95% CI: 2,30; 5,41)

Hvis kontrolgruppen udgjorde en 1% stikprøve af kildepopulationen, kan vi estimere, hvad resultatet af en follow-up undersøgelse ville have været:

Sovestilling	Cases	Antal børn	Risiko pr. 1000	RR
Maveleje	93	21.600	4,30	3,53
Andet	35	28.700	1,22	
I alt	128	50.300		

Kontrolgruppe-problemer

I eksemplet var kildepopulationen veldefineret; det er ikke altid tilfældet, og det bliver da et spørgsmål, hvordan en kontrolgruppe defineres.

Informations-problemer

Hvis man spørger til eksponering bagud i tid, kan informationen være usikker. Eller værre: informationen kan være præget af bevidstheden om begivenheden.

Alternative case-kontrol designs

Case-noncase undersøgelse.

Kontroller skal være ikke-syge. Kan give problemer med fortolkning af odds ratio.

Observerende undersøgelser

Beskrivende undersøgelser

Analytiske (sammenlignende) undersøgelser:

Follow-up

Case-kontrol

Eksperimentelle studier:

Styret eksponering; typisk for behandlingsforsøg.

Kan ikke gennemføres med potentielt skadelige eksponeringer.

Vurdering af epidemiologiske undersøgelser

Svend Juul, 9.2.2004

1

En hypotese:
Et udsagn om årsags-sammenhænge
'i virkeligheden'.

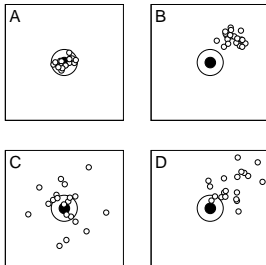
Effekt
Eksponering → Udfald

En empirisk undersøgelse:
En vurdering af hypotesen ud fra
indsamlede data.

Associationsmål
Eksponeringsmål → Mål for udfald

2

Måleproblemer

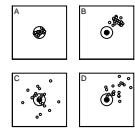


3

Usikkerhed og bias

• Præcision, sikkerhed, reproducibilitet, ryster ikke på hånden.
(modsat usikkerhed)

• Nøjagtighed, korrekthed
(modsat bias)



• **Usikkerhed** kan der kompenseres for ved gentagne målinger.

• **Bias** er en systematisk skævhed, som der ikke umiddelbart kan korrigeres for.

4

De vigtigste kilder til usikkerhed og bias

- Usikkerhed:
 - Stikprøveusikkerhed
 - Måleusikkerhed
 - Biologisk variation
- Bias:
 - Selektionsproblemer
 - Informationsproblemer (måleproblemer)
 - Fejltolkning (confounding)

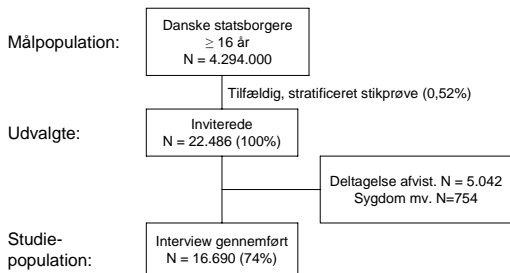
5

Stikprøveusikkerhed

- Hvis vi vil udtale os om danskernes vægtfordeling, er det billigere at undersøge en stikprøve end at undersøge alle danskerne.
- En *tilfældig stikprøve* er det ideelle. Det forudsætter en liste over alle danskere (CPR-registeret).
- Estimatet ud fra stikprøven af danskernes vægtfordeling er behæftet med usikkerhed.
- Denne usikkerhed er veldefineret og kan kvantificeres (SE, 95% CI).

6

Sundheds- og sygdomsundersøgelsen 2000



7

Bortfaldsproblemer

- Hvis ikke alle udvalgte medvirker, er der et *bortfald*.
- Bortfaldet kan medføre *bias* i vores estimat af danskernes vægtfordeling.
- Som læser: Er der gjort klart rede for stikprøvetagning, deltagelse og bortfald?
- Sensitivitetsanalyse: Hvad er den sande prævalens af fedme, hvis ikke-deltagerne:
 - har halvt så stor prævalens som deltagerne?
 - har dobbelt så stor prævalens som deltagerne?

8

Repræsentativitet

- Studiepopulationen skal være repræsentativ for målpopulationen *med hensyn til det, vi vil undersøge*.
- Repræsentativiteten kan trues af:
 - "forkert" stikprøve
 - skævt bortfald
- Et *selektionsbias* kan ikke korrigeres med statistiske metoder.
- Et selektionsbias' konsekvenser kan vurderes ved en sensitivitetsanalyse.

9

Selektionsproblemer i sammenlignende undersøgelser

- Selektionsproblemer kan have mindre omfattende konsekvenser i sammenlignende undersøgelser:
- Hvis et bortfald er "tilfældigt", dvs. uden sammenhæng med eksponering eller udfald, påvirkes associationsestimaterne ikke.
- Problemet er, at det kan vi sjældent vide; bortfaldet fører jo netop til manglende information.

10

Case-kontrol undersøgelse.
Bortfald associeret med eksponeringen, men ikke med udfaldet.

Udvalgte			Deltagere		
Eksponering	Cases	Kontroller	Eksponering	Cases	Kontroller
+	100	60	+	100	60
0	60	100	0	30	50
I alt	160	160	I alt	130	110
OR = 2,8			OR = 2,8		

11

Case-kontrol undersøgelse.
Bortfald associeret med både eksponeringen og udfaldet.

Udvalgte			Deltagere		
Eksponering	Case	Kontrol	Eksponering	Case	Kontrol
+	100	60	+	100	60
0	60	100	0	60	50
I alt	160	160	I alt	160	110
OR = 2,8			OR = 1,4		

12

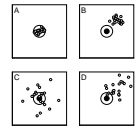
Bortfaldsproblemer i sammenlignende undersøgelser

- Et bortfald uden sammenhæng med eksponering eller udfald medfører ikke bias i associationsmålet.
- Et bortfald, som kun hænger sammen med eksponeringen, medfører ikke bias i associationsmålet.
- Et bortfald, som hænger sammen med både eksponering og udfald (dobbel-skævt), medfører bias.

13

Informationsproblemer

- Målefejl
 - Apparaturløst problem
 - Observatørløst problem
- Biologisk variation:



Ikke alle egenskaber er stabile; der er intra-individuel variation, f.eks. varierer den enkeltes blodtryk.

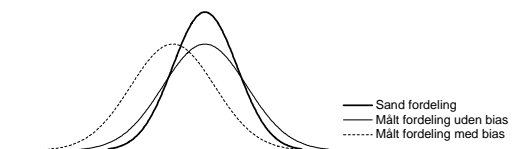
14

Konsekvenser af måleusikkerhed

- De observerede værdier får større spredning (SD) end de faktiske værdier.
- Det betyder, at middelværdien bliver bestemt med større usikkerhed (bredere sikkerhedsinterval).
- Ved sammenligninger overser man derfor lettere en reel forskel (type 2 fejl).
- Måleusikkerhed svarer til støj eller tåge: Tåge slører kontraster, det skaber dem ikke.

15

Målevariation og målebias



16

Målebias medfører ikke nødvendigvis bias i sammenlignende undersøgelser

- I en beskrivende undersøgelse: Et middelværdi-estimat bliver bias'ed
- I sammenlignende undersøgelser: Det kommer an på...

Børn af	Sandt	Målt
Ikke-rygere	3500 g	3400 g
Rygere	3350 g	3250 g
Differens	150 g	150 g

17

Misklassifikation (målefejl af kategoriske egenskaber)

- Graden af misklassifikation kan karakteriseres ved:
 - Sensitivitet: Evnen til at identificere de syge
 - Specificitet: Evnen til at identificere de ikke-syge

Klassificeret som	Sandheden	
	Syg	Ikke-syg
Syg	Sandt positiv	Falsk positiv
Ikke-syg	Falsk negativ	Sandt negativ

18

Estimation af sensitivitet og specificitet

Klassifikation	Sandheden	
	Syg	Ikke-syg
Syg	95	10
Ikke-syg	5	90
I alt	100	100

- Sensitivitet: 95% (95% CI: 89%; 98%)
- Specificitet: 90% (95% CI: 82%; 95%)

19

Konsekvenser af misklassifikation

- Beskrivende undersøgelser: Fejlagtigt hyppighedsestimat
- Sammenlignende undersøgelser: Det kommer an på ...
 - om misklassifikationen er *differentieret*

20

Ikke-differentieret misklassifikation

- Samme grad af misklassifikation af sygdomsstatus blandt eksponerede og ueksponerede
- Samme grad af misklassifikation af eksponering blandt syge og ikke-syge

21

Case-kontrol-undersøgelse: Ikke-differentieret misklassifikation af eksponeringen

Sand eksponeringsfordeling			Målt eksponeringsfordeling		
Eksponeret	Cases	Kontroller	Eksponeret	Cases	Kontroller
Ja	100	60	Ja	50	30
Nej	60	100	Nej	110	140
I alt	160	160	I alt	160	160
OR = 2,8			OR = 2,0		

- Både i case- og kontrolgruppen har 50% af de eksponerede "glemt" eksponeringen
- Der er altså samme grad af misklassifikation i de to grupper (ikke-differentieret misklassifikation)
- Det fører til en undervurdering af kontrasten (risiko for type 2 fejl).

22

Case-kontrol-undersøgelse: Differentieret misklassifikation af eksponeringen

Sand eksponeringsfordeling			Rapporteret eksponeringsfordeling		
Eksponeret	Cases	Kontroller	Eksponeret	Cases	Kontroller
Ja	20	40	Ja	20	20
Nej	100	200	Nej	100	220
I alt	120	240	I alt	120	240
OR = 1,0			OR = 2,2		

- 50% af kontrollerne havde "glemt" eksponeringen
- Det førte til en overvurdering af associationsestimatet
- Differentieret (dobbelt-skævt) misklassifikation fører til bias.

23

Problemer ved vurdering af konsekvenser af målefejl og misklassifikation

- Ikke differentieret målefejl eller misklassifikation svækker mulighederne for at se en kontrast. Det fører altså altid til bias i en bestemt retning.
- Differentieret målefejl eller misklassifikation kan føre til bias i begge retninger.
- Begynderfejl: "Der er måleproblemer, altså er der bias"

24

Eksempel: Kostoplysninger

- Det er givet, at kostoplysninger er usikre.
- Det betyder, at det er svært at studere effekten af kosten; vi kommer til at undervurdere effekten
- - med mindre der er tale om differentieret målefejl: At dem, der blev syge, rapporterer anderledes end dem, der forblev raske.
- Læseren må forsøge at vurdere, hvad der kan være sket.

25

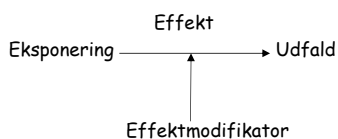
Effektmodifikation

- En eksponering eller behandling virker ikke ens på alle.
- Det må der være grunde til: andre faktorer, som modificerer effekten af eksponeringen:
 - køn
 - arvelige faktorer
 - livsstilsfaktorer
- Hvis vi kan identificere grupper, som er særligt modtagelige, er det nyttig viden:
 - planlægning af forebyggelse
 - differentieret behandling

26

Effektmodifikation

- Lægemiddel A påvirker effekten af lægemiddel B
- Vi taler om effektmodifikation eller *interaktion*



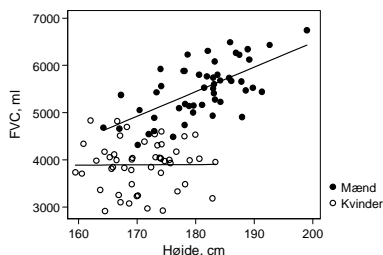
27

Effektmodifikation

- Eksempel 8-1: En ny mæslingevaccine var forbundet med øget generel dødelighed i forhold til den hidtidige:
 - MRR = 1,33 (95% CI: 1,02; 1,73)
- Imidlertid var der stor forskel på effekten hos drenge og piger:
 - Drenge: MRR = 0,91 (95% CI: 0,61; 1,35)
 - Piger: MRR = 1,86 (95% CI: 1,28; 2,70)
- Fundet er interessant, men mekanismen endnu ikke forstået.

28

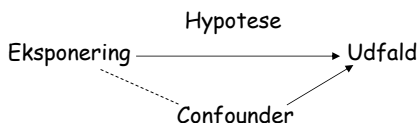
Effektmodifikation: Forskellig sammenhæng mellem højde og lungefunktion hos mænd og kvinder



29

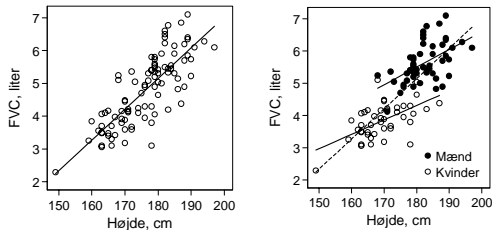
Confounding

- Sammenblanding, forveksling
- Confounding er en fortolkningsfejl
- Risiko for confounding: Hvis en anden årsag er associeret med den årsag, vi studerer.



30

Confounding: Lungefunktion og højde



31

Confounding: Alkohol og larynxkræft

Rygere		
Alkohol	Cases	Kontroller
Ja	116	230
Nej	124	240
I alt	240	470
OR = 0,98 (0,72; 1,33)		

Ikke-rygere		
Alkohol	Cases	Kontroller
Ja	15	110
Nej	55	450
I alt	70	560
OR = 1,12 (0,61; 2,05)		

32

Dette er ikke en confounder:

- En faktor, som er et led i årsagskæden fra eksponering til udfald:



- F.eks.:
Kost → S-kolesterol → Åreforkalkning

33

Forebyggelse af confounding: restriktion

- Hvis vi kun ser på sammenhængen mellem alkoholindtagelse og larynxcancer blandt ikke-rygere, kan der ikke opstå confounding på grund af rygning.
- Til gengæld ved vi ikke noget om sammenhængen blandt rygere.

34

Forebyggelse af confounding: matching

- I follow-up-undersøgelser:
 - For hver eksponeret, rygende, 50-årige fraskilte mand kunne man finde en ueksponeret, rygende, 50-årig fraskilt mand til den ueksponerede kontrolgruppe.
 - Det kræver en stor, velkarakteriseret gruppe at trække kontrollerne fra.
- I case-kontrol-undersøgelser:
 - Man kan ikke studere effekten af det, der er matchet på.
 - Man komplicerer den statistiske analyse (betinget logistisk regression).

35

Forebyggelse af confounding: randomisering

- I behandlingsforsøg kan man ved lodtrækning fordele patienterne tilfældigt til de behandlinger, hvis effekt skal sammenlignes.
- Herved sikres en ligelig fordeling af kendte og ukendte confundere mellem grupperne.
- Det kan ikke anvendes ved potentielt skadelige eksponeringer.
- Mere herom næste mandag.

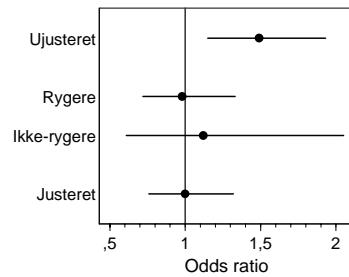
36

Stratificeret analyse

- Ved restriktion kan man undersøge associationen (alkohol/cancer), dels hos rygere, dels hos ikke-rygere.
- Er associationen (alkohol/cancer) væsentligt forskellig hos rygere og ikke-rygere, er der *effektmodifikation*. Der må laves to beskrivelser, for rygere og for ikke-rygere.
- Er associationerne ikke væsentligt forskellige, beregnes et fornuftigt vægdet gennemsnit.

37

Stratificeret analyse alkohol/larynxkræft Ujusteret og justeret estimat



38

Regressionsanalyse

- Ved regressionsanalyse kan man undersøge den samtidige effekt af flere faktorer, f.eks. alkohol, rygning, køn m.m.

Eksponering	OR	95% CI
Alkohol	1,00	0,76; 1,32
Rygning	4,08	3,01; 5,53
Køn		
osv.		

39

Effektmodifikation og confounding

- Effektmodifikation er noget virkeligt: Det skal erkendes og beskrives; det kan være værdifuld viden af betydning for differentieret forebyggelse og behandling.
- Confounding (forveksling af årsager) er en fejltolkning, som skal søges undgået.

40

Eksperimentelle undersøgelser

Poul Thorsen

Observationelle studier:

Studier af forekomst (incidens, prævalens)

Studier af sammenhænge eller kontraster "i naturen"

Eksperiment, forsøg:

Studiet af reaktionen på en styret eksponering

Det ideelle spørgsmål:

Hvordan var det gået den eksponerede person uden eksponering – eller med en anden eksponering?

(*counterfactual* tankegang)

Crossover forsøget:

Sammenligning af behandling A og B

Person	1	2	3	4	5	6	7	8	9	10	11	12
Tid 1	A	B	B	A	B	A	A	A	B	B	A	B
Tid 2	B	A	A	B	A	B	B	B	A	A	B	A

Forudsætninger:

at tilstanden uden intervention er stabil

at behandlingernes virkning er forbigående

at rækkefølgen (AB, BA) er uden betydning

Crossover forsøget er det, der nærmest svarer til en **counterfactual** tankegang. Personen er sin egen kontrol.

Person	1	2	3	4	5	6	7	8	9	10	11	12
Behandling	A	B	B	A	B	A	A	A	B	B	A	B

Hvis crossover forsøgets forudsætninger ikke er til stede, må man sammenligne **forskellige** personer:

A- og B-gruppen må ikke afvige systematisk fra hinanden mht. andet end behandlingen. Derfor **randomisering**.

Formål med randomisering:

Etablering af sammenlignelige grupper

Undgå confounding – kendte og ukendte faktorer

In the absence of differences in effect, the outcomes in the compared populations could be expected to be identical.

Randomisering:

Hvordan randomiseres?

Blok-randomisering

(MRC/BHF Heart protection study of antioxidant vitamin supplementation in 20536 high-risk individuals: a randomised placebo-controlled trial. Lancet 2002; 360: 23-33)

Hypotese: Antioxydante vitaminer (C, E, β -karoten) reducerer risikoen for vaskulær sygdom og cancer.

Studiepopulation: 20.536 40-80 år, UK, med coronarsygdom eller diabetes.

Design: Randomisering til vitaminer eller placebo.

Studieperiode: Inklusion juli 1994 - maj 1997. Follow-up slut maj-oktober 2001. Follow-up tid 4-7 år (gns. 5,3 år)

Udfald:

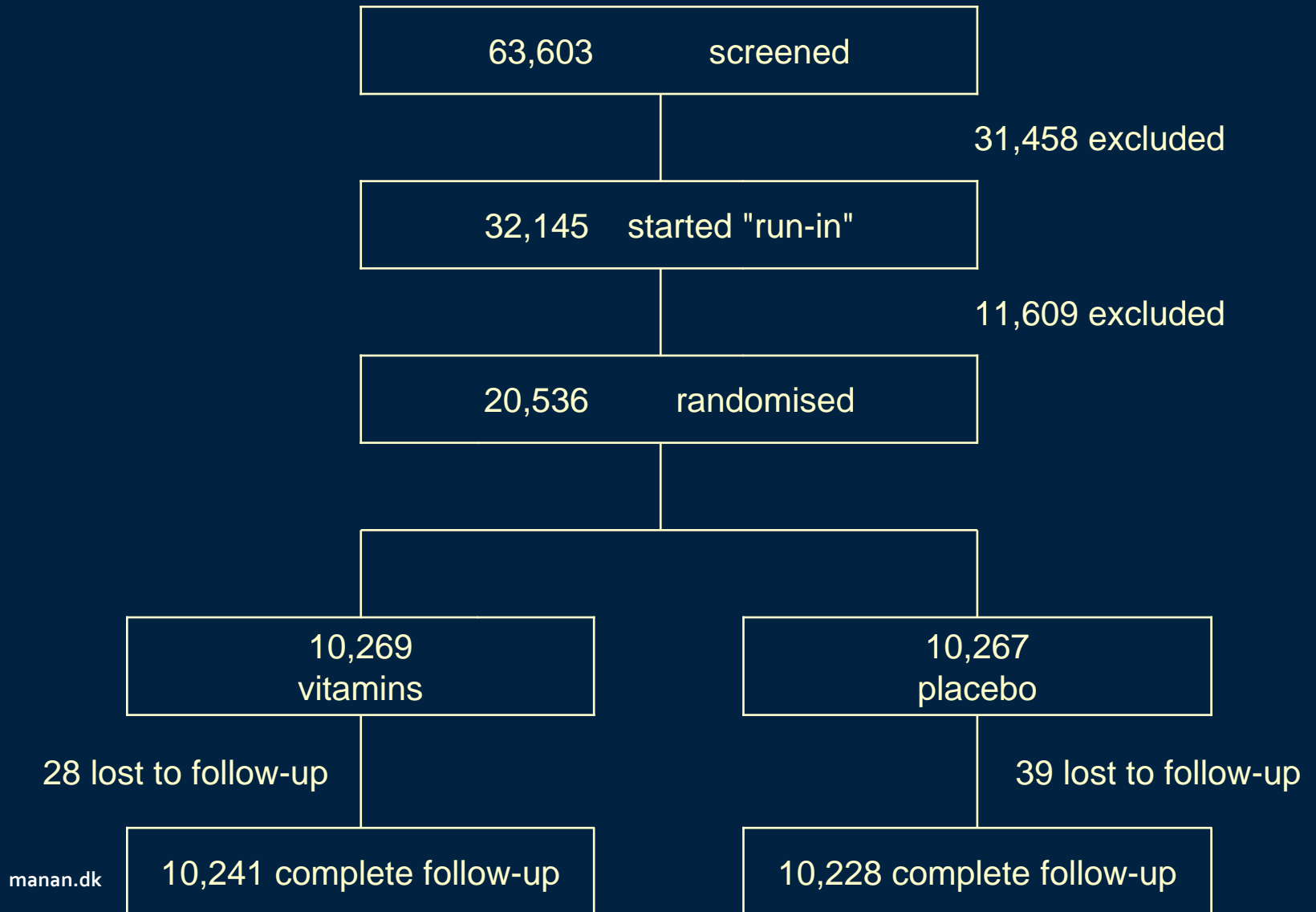
Total dødelighed

Død af kardiovaskulær sygdom, cancer

"Major vascular events"

Compliance:

Efter 1 år 89%, efter 5 år 77%.



Outcome	Begivenheder	IRR (95% CI)
Total dødelighed	1446; 1389	1,04 (0,97; 1,12)
Kardiovaskulær død	878; 840	1,05 (0,95; 1,15)
Anden død	568; 549	1,04 (0,92; 1,17)
Major vascular events	2306; 2312	1,00 (0,94; 1,06)
Cancertilfælde	800; 817	0,98 (0,89; 1,08)

Konklusion:

Evidence against major benefit or harm

Blinding

➤ af forsøgspersoner:

- Viden (tro) om behandling kunne have sundhedseffekter (placebo-virkning)
- Viden (tro) kunne påvirke rapportering (informations-bias)

➤ af bedømmere:

- Ønsketænkning (tro) kunne påvirke vurderinger

Blinding

Hvor blind kan man være?

Enkelt

Dobbelt

Tre- dobbelt

Placebo (jeg behager)

- Viden (tro) om en behandlings effekt kan have sundhedseffekter – og effekt på oplevelse og rapportering.
- Placebo bruges undertiden – bevidst eller ubevidst – i behandlingsøjemed
- Det bruges bevidst af mange alternative behandlere
- Det er et problem ved studiet af specifik behandlingseffekt (confounding-problem)
- Problemet er størst ved selvrapporeret helbred.

Regression mod gennemsnittet

Hvis:

sygdomsintensiteten varierer, og patienter inkluderes, når de har det dårligt,

så vil:

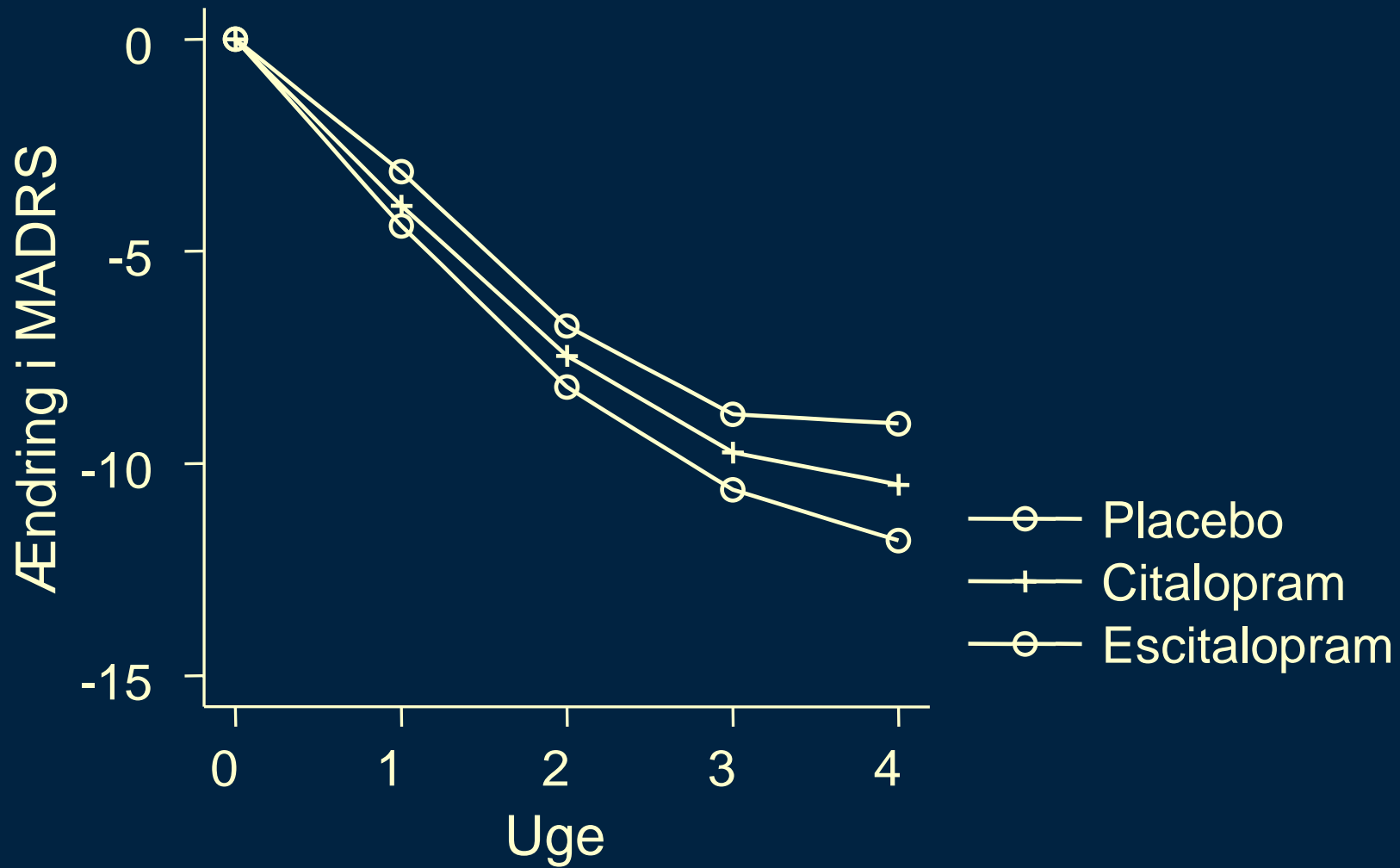
patienterne ved en efterundersøgelse have det bedre end da de blev inkluderet,

selv om:

behandlingen er virkningsløs

Derfor:

kontrolgruppe.



Hvad skal behandlingen sammenlignes med?

Hvis der allerede findes en anerkendt behandling:

- Er den nye behandling bedre end den gamle?

Hvis der ikke findes en anerkendt behandling:

- Er den nye behandling bedre end ingen behandling?

Formål med en 'placebo' kontrolgruppe:

at adskille specifik (farmakologisk) effekt fra uspecifik effekt (f.eks. omsorg)

at adskille biologisk effekt fra psykologisk effekt

Analyse: Intention-to-treat

Studiegrupperne som defineret ved randomiseringen fastholdes uanset compliance.

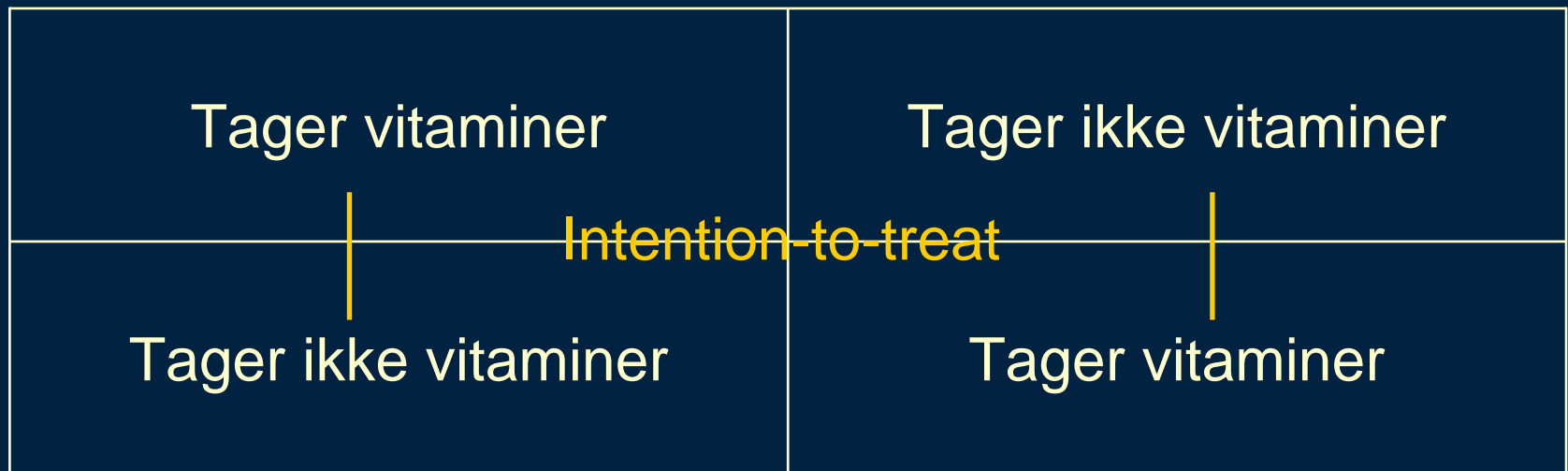
Ved analyse efter faktisk eksponering fås det samme confounding-problem som ved observerende undersøgelser.

Dårlig compliance ~ udifferentieret misklassifikation af eksponering ► undervurdering af effekt.

Randomiseret til:

Vitaminer

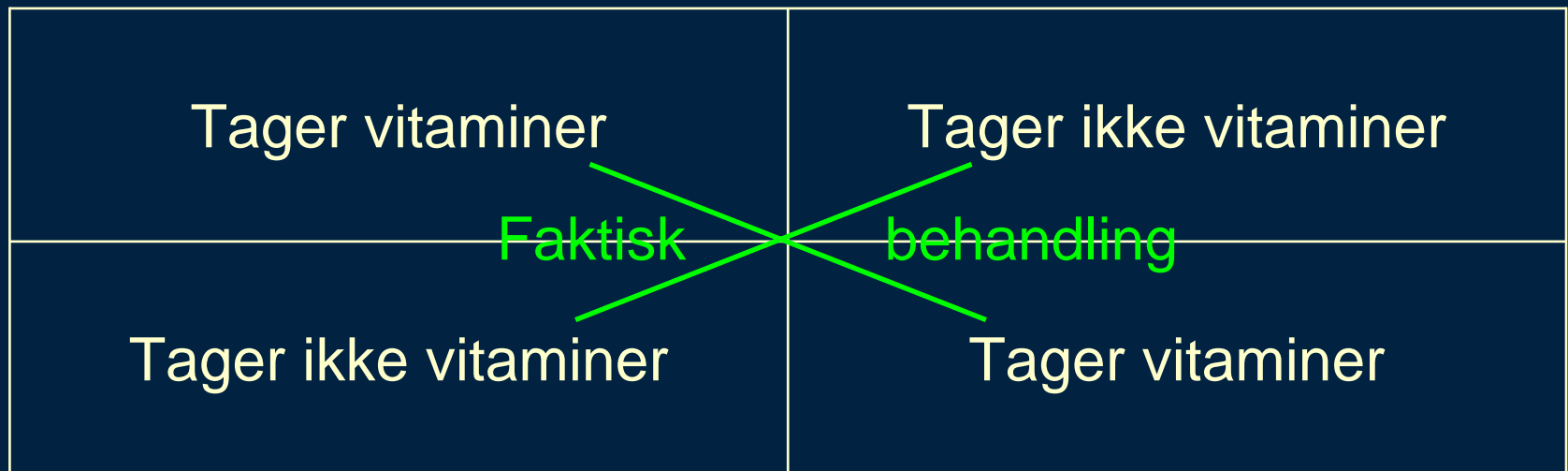
Placebo



Randomiseret til:

Vitaminer

Placebo

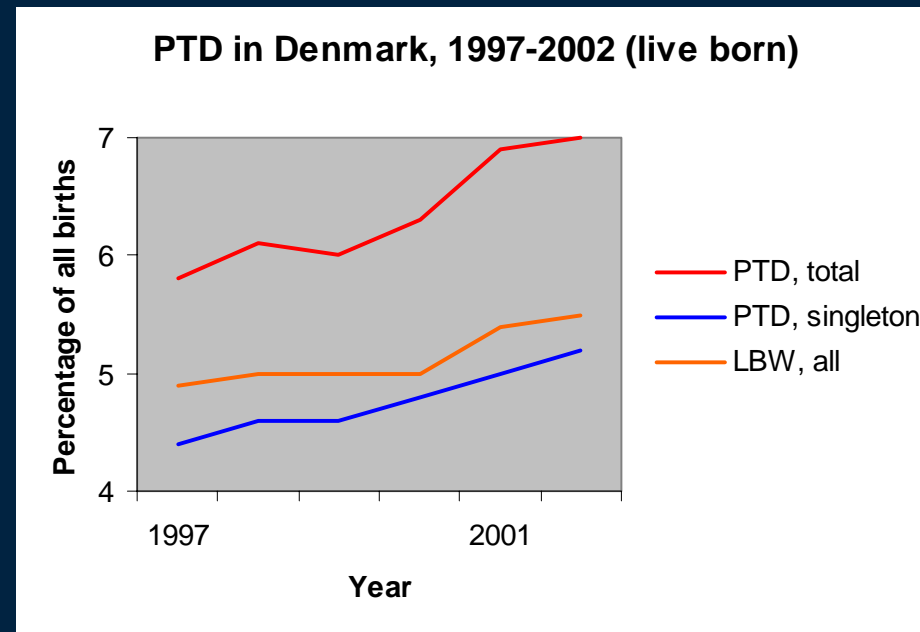


Designe eget studie

For tidlig fødsel før 37 graviditetsuger er steget med over 30% de sidste 10 år

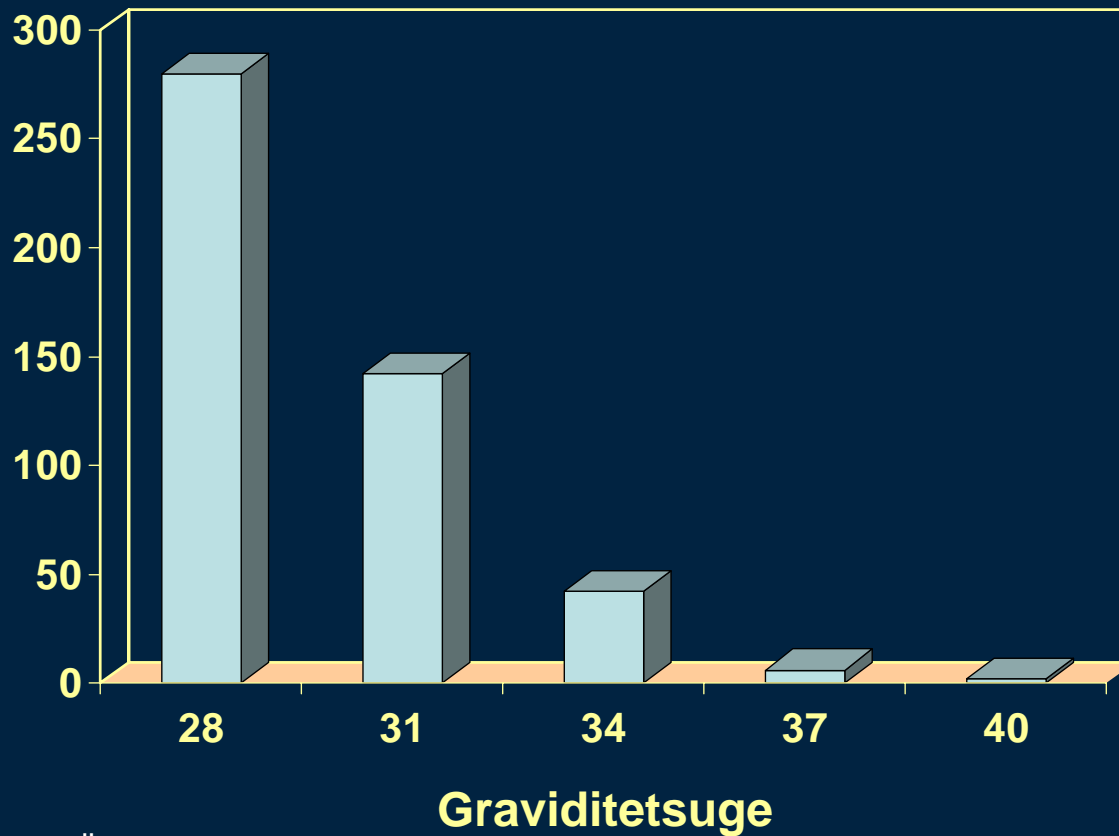
Det skyldes ikke blot anvendelsen af kunstig befrugtning og/eller flere med flerfoldsgraviditeter

Der er ikke nogen effektiv intervention



Designe eget studie

Dødelighed per 1000 fødte børn



8 gange højere dødelighed for for tidligt fødte børn

Designe eget studie

Hvad skal vi gøre?

Designe eget studie

Tailored intervention?

Lægemiddelforsøg

Fase 1 og 2: Forstudier inden en egentlig vurdering (toxicitet, dosering)

Fase 3: Randomiserede forsøg (efficacy)

Fase 4: Post-marketing surveillance. Bivirkningsmonitorering

For lægemidler er der strenge krav til dokumentation af sikkerhed og virkning. For andre behandlingsteknologier – f.eks. kirurgiske – er det meget mindre konsekvent.

Eksternt validitetsproblem

Fase 3 forsøg gennemføres ofte ved specialafdelinger på karakteristiske, velundersøgte patienter.

Efter lancering udvides anvendelse ofte til bredere patientgrupper – og industrien gør sit til det.

For hvilke patientgrupper er erfaringen fra fase 3 forsøget gyldige?

Etiske retningslinjer

Behandlingsforsøg skal godkendes af en videnskabsetisk komité, som vurderer:

- Forsøgets potentielle nytte**
- Risiko og belastning af forsøgspersoner**
- Kvaliteten af patientinformationen**
- Informeret samtykke**

Conflict of interest

- **Har forfatterne til en artikel interesse i et bestemt resultat?**
- **Er de tilknyttet et lægemiddelfirma?**
- **Har de aktier i et lægemiddelfirma?**
- **Hvorfra har de modtaget støtte til projektet?**
- **(Har de forelsket sig i et bestemt resultat?)**
- **Mange tidsskrifter forlanger en selvdeklaration fra forfatterne.**

Diagnostik

Svend Juul, 23.2.2004

1

Hvordan stiller man en diagnose?

- Symptomer
 - passive: patientens spontane rapport
 - aktive: svar på målrettede spørgsmål
- Fund

2

Fund

- Kliniske fund
 - passive: ernæringstilstand, cyanose
 - aktive: patellar-refleks
- Fysiske fund
 - passive: vægt, højde
 - aktive: lungefunktionsmål
- Biokemiske fund
 - passive: plasma-glukose
 - aktive: glukosebelastning
- Billeddiagnostiske fund
- Pato-anatomiske fund

3

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

- Den dygtige læge er effektiv ved:
 - udvælgelse af arbejdshypoteser
 - især mht. sygdomme, hvor behandling kan gøre en forskel
 - valg af test som kan bestyrke eller afsvække en hypotese
- Diagnoser er ofte behæftet med usikkerhed – også den erfarne læges
- Mange patienter fejler mere end én ting

4

Hvorfor er diagnoser nyttige?

- Diagnoser er skabt af mennesker
- Ikke to patienter er ens – men nogle har fælles træk. Diagnoser er et økonomisk middel i kommunikationen om patienter.
- Diagnoser strukturerer vores fælles viden.
- Studier af ætiologi mhp. forebyggelse
- Studier af behandlingseffekt
 - og diagnosen er særligt nyttig, når vi kender en behandling, der gør en forskel.

5

Diagnostiske kriterier

- Centrum:
 - Hvad er de centrale karakteristika for sygdommen?
- Periferi:
 - Hvordan afgrænses sygdommen over for andre sygdomme – og over for "rask"?

6

Diagnoseklassifikation: ICD-10

- I A00-B99 Infektiøse sygdomme
- II C00-D48 Svulster
- XI K00-K93 Sygd. i fordøjelsesorganer
 - K20-K31 Spiserør, mavesæk, duodenum
 - K25 Mavesår
 - K25.0 Akut mavesår med blødning
 - K25.1 Akut mavesår med perforation

7

Diagnostisk test: mammografi

- Forkert udtryk: Diagnoser stilles ikke af test, men af mennesker.
- Mammografi indførtes først som diagnostisk redskab: Vurdering af diagnosen brystkræft, 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

8

På en specialafdeling undersøgtes 1000 kvinder, som var henvist pga. mistanke om brystkræft. 40% havde det faktisk.

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: 40%

Sensitivitet: 90%

Specificitet: 95%

Prædiktiv værdi af positiv test: $360/390 = 92\%$

Prædiktiv værdi af negativ test: $570/610 = 93\%$

9

Der er tale om *estimer*:

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

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

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

10

Mere generelt:

Test	Syg		I alt	P(syg)
	+	0		
+	SP	FP	SP+FP	SP/(SP+FP)
0	FN	SN	FN+SN	FN/(FN+SN)
I alt	SP+FN	FP+SN		

Testen er ikke perfekt. Men den kan forbedres:

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

11

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.

12

Revision af sandsynligheder

- Et testresultat står ikke alene: Det skal kombineres med anden evidens.
- Før testresultatet foreligger, har vi en *a priori* sandsynlighedsvurdering:
 - $P(\text{brystkræft}) = 40\%$
- Testresultatet modificerer *a priori* vurderingen til en *a posteriori* sandsynlighedsvurdering (prædiktiv værdi):
 - Positiv test: $P(\text{brystkræft}) = 92\%$
 - Negativ test: $P(\text{brystkræft}) = 7\%$

13

Mammografi ved screening: *a priori* sandsynlighed = 1%

Test	Syg		I alt	P(syg)
	+	0		
+	900	4.950	5.850	15,4%
0	100	94.050	94.150	0,1%
I alt	1.000	99.000	100.000	1,0%
Sensitivitet	90%			
Specificitet	95%			

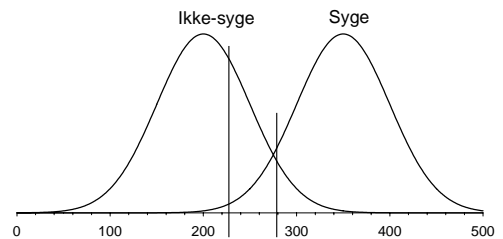
14

Test med kontinuert udfald, f.eks. biokemiske målinger

- Reference-interval: Karakteristik af testresultaternes fordeling blandt ikke-syge (et 95% prædiktionsinterval).
- Plasma-kreatinin:
 - Mænd: 62-133 $\mu\text{mol/l}$
 - Kvinder: 44-115 $\mu\text{mol/l}$
- Reference-intervallets grænser adskiller IKKE syge fra raske.
- Betegnelsen "normal-værdier" er uheldig.

15

Test med kontinuert udfald



16

- Ved fastlæggelse af cutpoint i en konkret diagnostisk situation må man tage hensyn til:
 - *a priori* sandsynligheden for sygdom
 - konsekvenserne af en falsk negativ diagnose
 - konsekvenserne af en falsk positiv diagnose
- Alvorlig konsekvens af en falsk negativ diagnose:
 - At en værdifuld behandling undlades.
- Alvorlig konsekvens af en falsk positiv diagnose:
 - At der iværksættes en belastende behandling uden sundhedsgevinst.

17

18

Screening

Ulrik Kesmodel

19

Definition

Systematisk undersøgelse af asymptomatiske befolkningsgrupper for en eller flere sygdomme mhp. at vurdere om (det er sandsynligt at) de har sygdommen

Aktiviteten finder oftest sted i et myndighedsregi (*Sundhedsstyrelsen*)
Sundhedsmyndighederne tager initiativet til undersøgelsen (*Etisk Råd*)

20

Eksempler

1. Celleskrab fra livmoderhalsen
2. Mammografi
3. U-glukose hos gravide
4. Hælbloodprøve hos nyfødte
5. Alder (fosterdiagnostik)

21

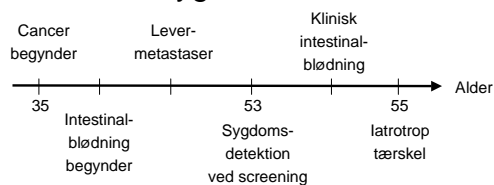
Formål med screening

1. At reducere dødelighed
2. At reducere lidelse og funktionstab

ved at finde sygdomstilfælde i tidligere stadier end ellers

22

Sygdommen



1. (Lang) præklinisk fase
2. Sygdommens naturhistorie skal være kendt

Rothman & Greenland.
Modern Epidemiology, 1998

23

Eksempler

1. Celleforandringer/livmoderhalskræft
2. Brystkræft
3. Gestationel diabetes mellitus?
4. Medfødte stofskiftesygdomme
5. Trisomi 21 (Downs syndrom)

24

Sygdommen

Hyppeghed

- Celleforandringer/livmoderhalskræft 5000/500
- Brystkræft 3000
- Gestationel diabetes 1200
- Trisomi 21 (Downs syndrom) 75
- Medfødte stofskiftesygdomme 45

25

Behandling

1. Bedre behandlingseffekt for tidligt end for sent diagnosticerede tilfælde
 - Celleforandringer/livmoderhalskræft
 - Brystkræft
 - Gestationel diabetes
 - Medfødte stofskiftesygdomme
 - Trisomi 21 (Downs syndrom)?
2. Behandlingsfaciliteter til stede

26

Testen(e)

Hvad er en "screeningstest"?

27

Testen(e)

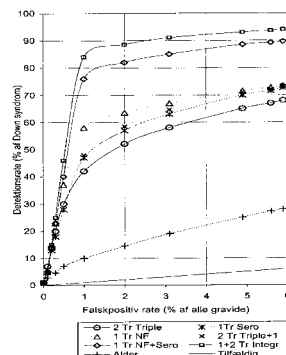
1. Skal med en rimelig sandsynlighed kunne identificere personer med høj sygdomsrisiko

28

Fosterdiagnostik – trisomi 21

1. Blodprøve 8-13 uger
2. Nakkefoldsscanning, 11-13 uger
3. Blodprøve 15-17 uger
4. Alder

29



30

Testen(e)

1. Skal med en rimelig sandsynlighed kunne identificere personer med høj sygdomsrisiko
2. Skal være minimalt belastende og (næsten) uden risiko

31

Fosterdiagnostik – trisomi 21

1. Blodprøve 8-13 uger
2. Nakkefoldsscanning, 11-13 uger
3. Blodprøve 15-17 uger
4. Alder
5. Fostervandsprøve/moderkageprøve

32

Sandsynlighedsmål

1. Sensitivitet, specificitet
2. Prædiktiv værdi af positivt/negativt resultat
3. Odds (1:250; 1:2500)

33

Økonomi og organisation

- Rimeligt forhold mellem nettosundhedsgevinst og omkostninger

34

	Alle 15-49 år	15-29 år	30-34 år	35-39 år	40-49 år
UD	188	152	27.3	7.9	1.5
IND	53.7	28.1	12.1	8.3	5.2
UD-IND	- 135	-124	-15.1	+0.4	+3.7

35

Fosterdiagnostik 2004

Økonomi

- Etableringsomkostninger 35 mio. kr.
- Driftsomkostninger/år 45 mio. kr.
- Besparelse/år (behandling, pleje, institutionspladser etc.) 48 mio. kr.

36

Fosterdiagnostik 2004

Andre gevinster/omkostninger

- Færre spontane aborter - 35
- Flere provokerede aborter 80-125
- Glæde/afsavn?

37

Økonomi og organisation

- Rimeligt forhold mellem nettosundhedsgevinst og omkostninger
- Kapacitet til undersøgelse og behandling

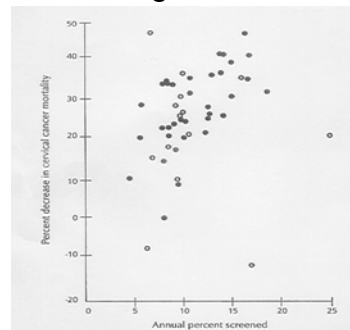
38

Evaluering af screeningsprogrammer

1. Makroundersøgelser

39

Screening for livmoderhalskræft



Cancer 1974;34:2018-27

40

Evaluering af screeningsprogrammer

1. Makroundersøgelser
2. Follow-up og case-kontrol studier

41

Observationelle studier

Fordele

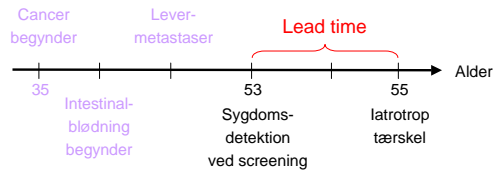
- Kan evaluere effekten af et screeningsprogram ved rutinemæssig brug

Ulemper

- Lead time bias

42

Sygdommen



Rothman & Greenland.
Modern Epidemiology, 1998
43

Observationelle studier

Fordele

- Kan evaluere effekten af et screeningsprogram ved rutinemæssig brug

Ulemper

- Lead time bias
- Manglende blinding af evaluering af dødsårsager
- Opportunistisk screening og forbedrede behandlingsmuligheder vanskeliggør identifikation af årsager til et evt. fald i mortaliteten

44

Evaluering af screeningsprogrammer

1. Makrounder søgelser
2. Follow-up og case-kontrol studier
3. Randomiserede studier

45

Fosterdiagnostik – trisomi 21

Er det overhovedet screening?

46

Definition

Systematisk undersøgelse af asymptomatiske befolkningsgrupper for en eller flere sygdomme mhp. at vurdere om det er sandsynligt at de har sygdommen

Aktiviteten finder oftest sted i et myndighedsregi (*Sundhedsstyrelsen*)
Sundhedsmyndighederne tager initiativet til undersøgelsen (*Etisk Råd*)

47

Fosterdiagnostik – trisomi 21

Er det overhovedet screening?

Ifølge Sundhedsstyrelsen: NEJ

Sundhedsstyrelsen
Notat vedrørende nye retningslinier
for fosterdiagnostik, 2003 48

Fosterdiagnostik – trisomi 21

Forebyggelsesparadigme

Formålet med fosterdiagnostik er at hindre fødsel af børn med alvorlige sygdomme eller handicaps



Screening

*Sundhedsstyrelsen
Fosterdiagnostik og risikovurdering,
2003* 49

Fosterdiagnostik – trisomi 21

1. Information hos egen læge om muligheder for rådgivning/undersøgelse
2. Henvvisning til specialafdeling
3. Nærmere udredning

Sundhedsvæsenet tager initiativ til undersøgelserne ved at forslå, anbefale, opfordre, invitere til eller ukaldet initiere dem

*Sundhedsstyrelsen
Notat vedrørende nye retningslinier
for fosterdiagnostik, 2003* 50

Fosterdiagnostik – trisomi 21

Selvbestemmelsesparadigme

Formålet med fosterdiagnostik er at bistå en gravid kvinde, som ønsker en sådan bistand, med at træffe sine egne valg. Neutral og fyldestgørende rådgivning er forudsætning herfor.



~~Screening~~
Informeret valg

51

Fosterdiagnostik – trisomi 21

1. Information hos egen læge om muligheden for nærmere information om muligheder for rådgivning/undersøgelse
2. Information hos egen læge om muligheder for rådgivning/undersøgelse
3. Henvisning til specialafdeling
4. Nærmere udredning

52

Fosterdiagnostik – trisomi 21

Informeret valg

Et tilbud om information om undersøgelser inden for svangreomsorgens regi er ikke at sidestille med en generel anbefaling eller opfordring til, at den gravide skal gennemgå sådanne undersøgelser.

*Sundhedsstyrelsen
Notat vedrørende nye retningslinier
for fosterdiagnostik, 2003* 53

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 undersøgte 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.

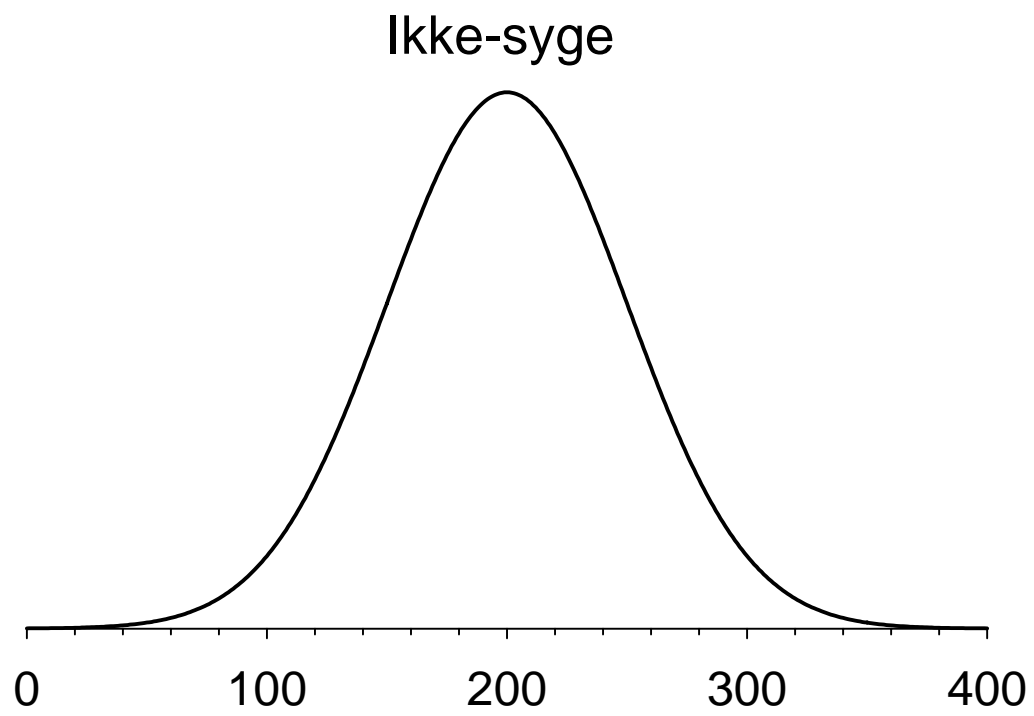
F.eks. plasma-kreatinin: Mænd: 62-133 $\mu\text{mol/l}$

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

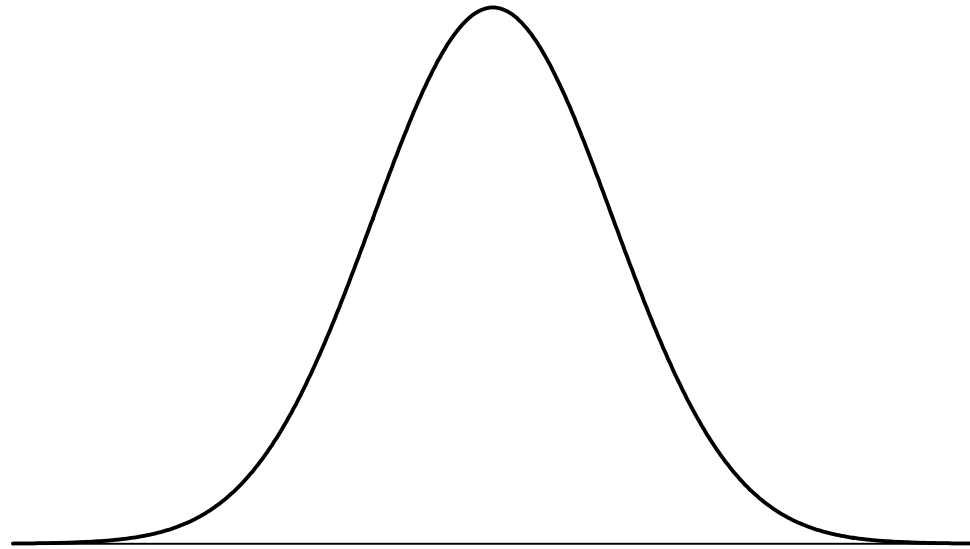
Reference-intervallet bestemmes typisk som et 95% prædiktions-interval i en rask population.

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

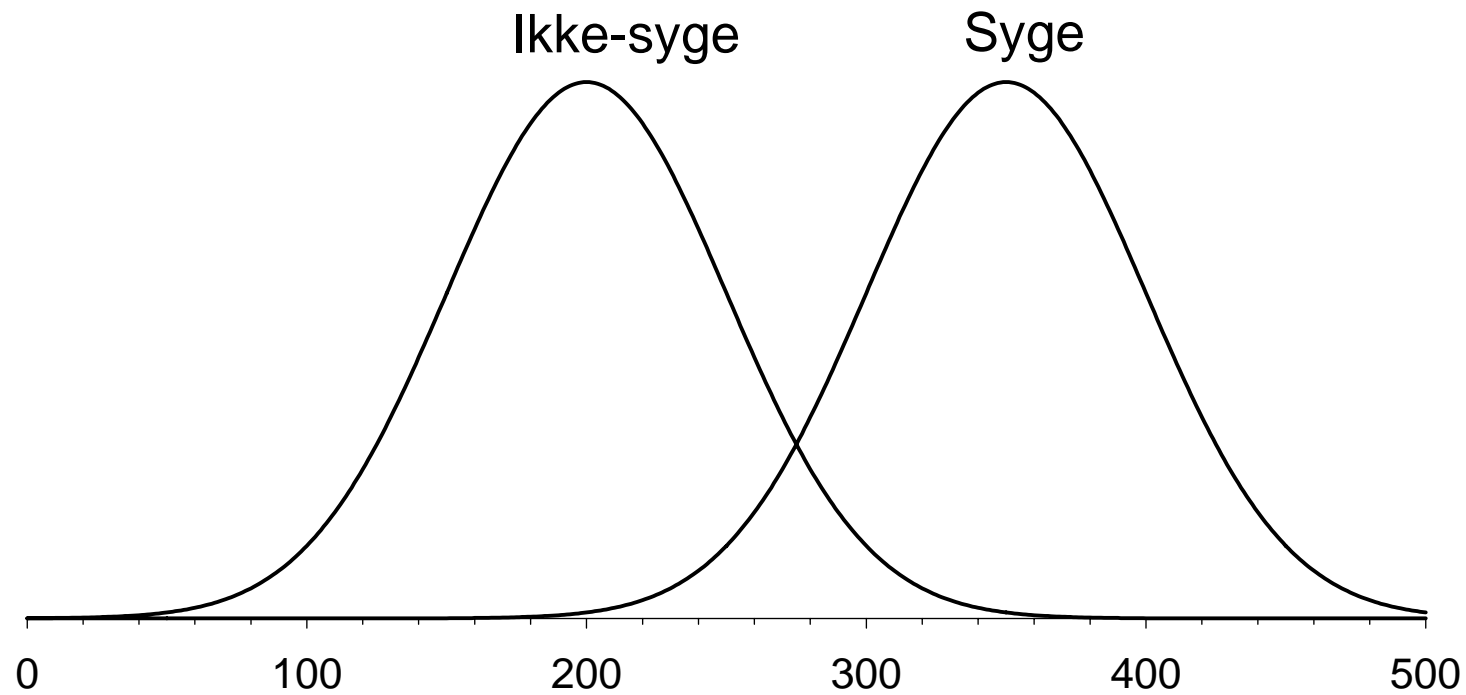
Betegnelsen "normal-værdier" er uheldig.



Syge



Tests med kontinuerligt 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:

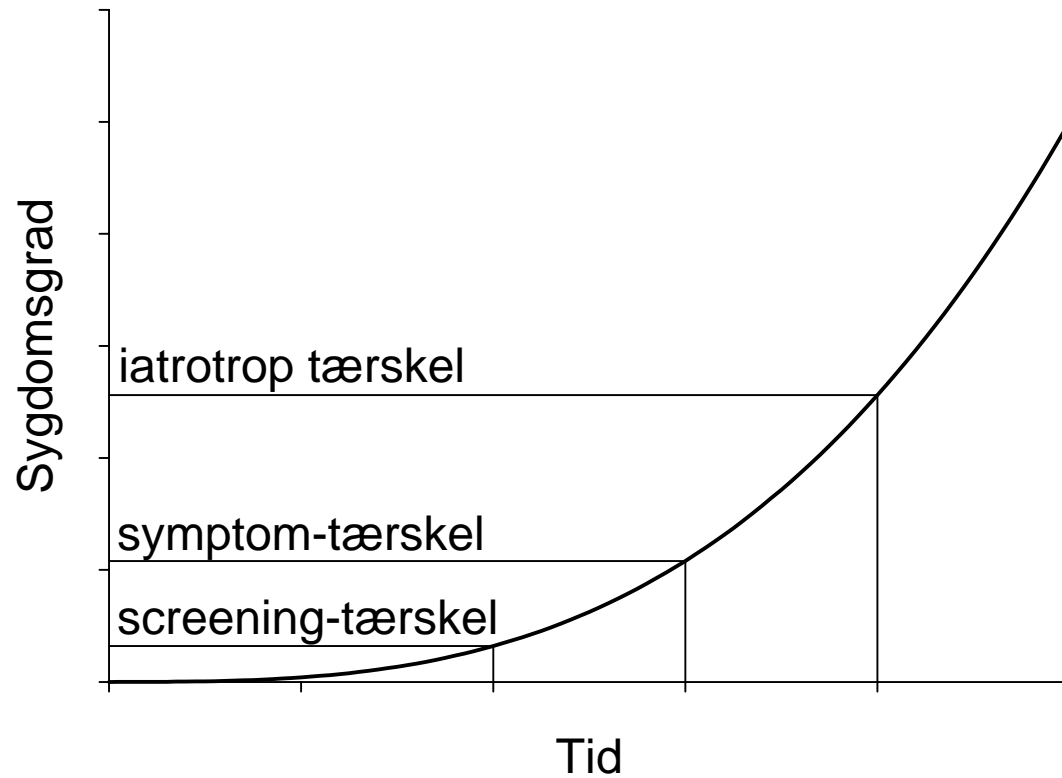
- Hvis en behandlelig sygdom med alvorlig prognose overses

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

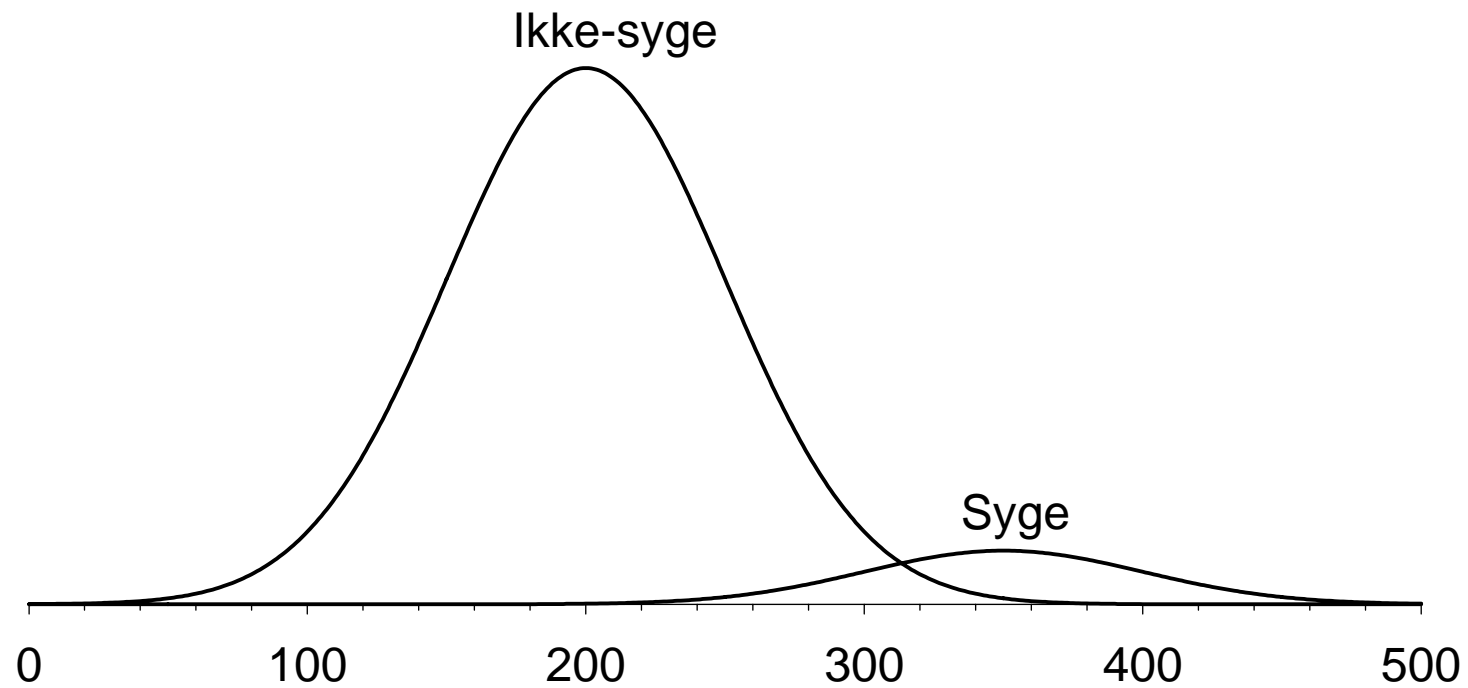
Sensitivitet = 0,90 (90% af de syge diagnosticeres korrekt)

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.

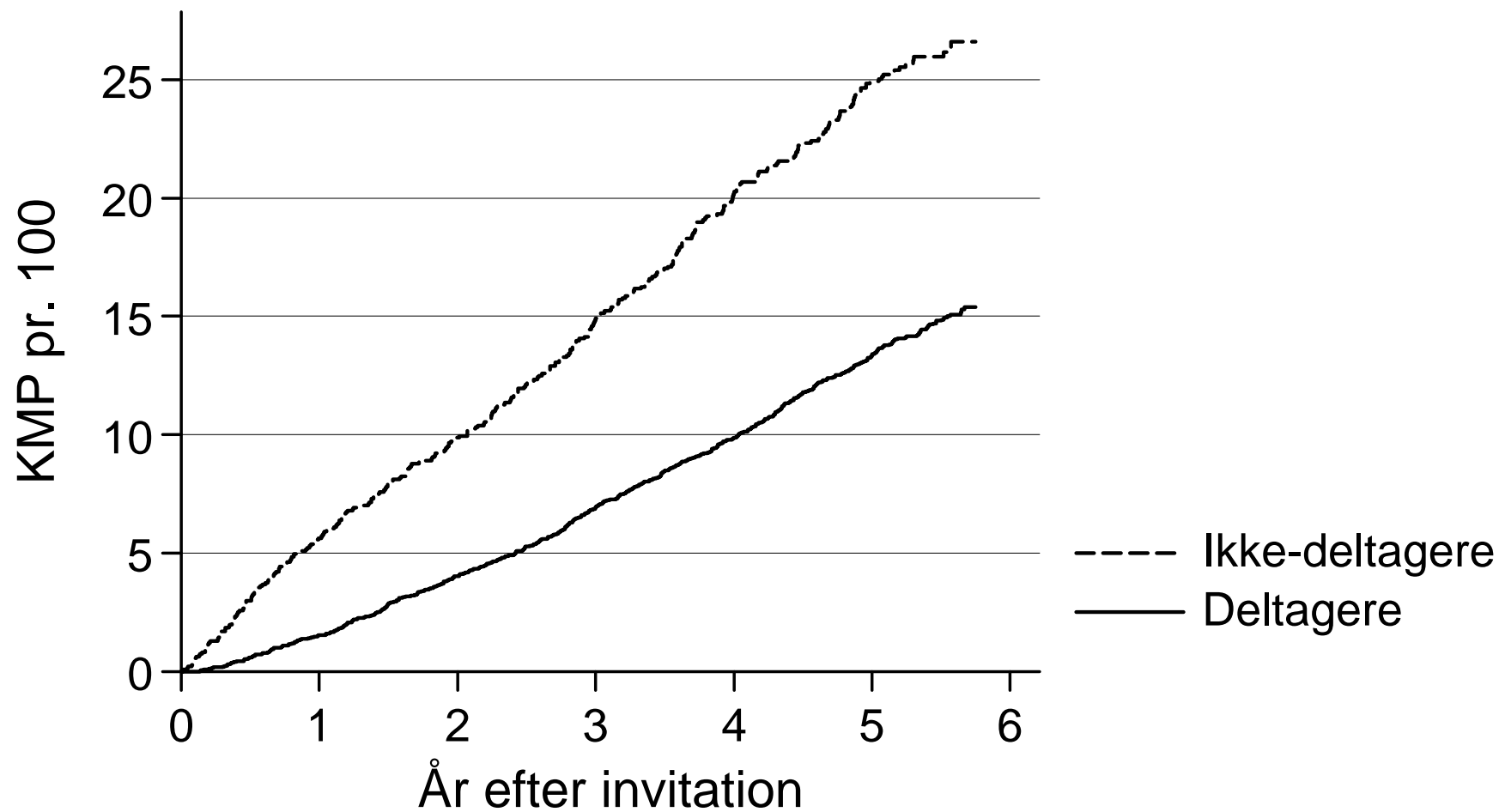
Der forudsættes typisk mange personer og lang tidshorisont.

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?

Vurdering af epidemiologiske undersøgelser

Mogens Vestergaard
Institut for Epidemiologi og Socialmedicin
Aarhus Universitet

Målsætning

At belyse en videnskabelig problemstilling ved at indsamle, analysere og fortolke informationer fra en stikprøve af befolkningen

Fejkilder

- Selektionsproblemer
- Informationsproblemer
- Fejltolkning (confounding)

Stikprøve



Fejl

- Stikprøveusikkerhed
- Systematisk fejl (Bias)

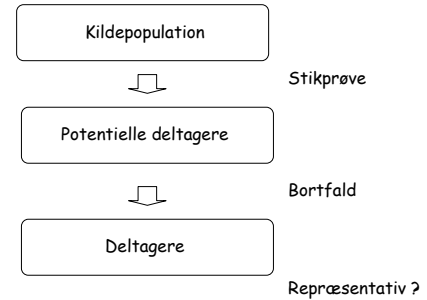
Stikprøveusikkerhed

Stikprøve størrelse	Antal positive	PP	95% CI
20	7	0,35	(0,11-0,66)

Stikprøveusikkerhed

Stikprøve størrelse	Antal positive	PP	95% CI
20	7	0,35	(0,11-0,66)
100	35	0,35	(0,23-0,48)
1.000	350	0,35	(0,31-0,39)
10.000	3.500	0,35	(0,34-0,36)

Stikprøve, bortfald og repræsentativitet

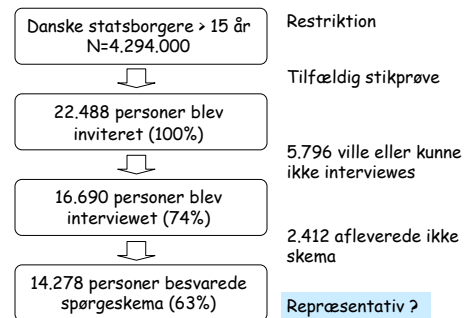


Bortfald



Storm P. 1910. Tilbage til naturen. Storm P Museet

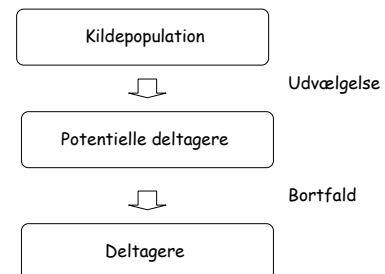
Sundheds- og sygelighedsundersøgelsen i DK 2000



Selektionsproblemer i sammenlignende undersøgelser

Selektionsbias opstår i sammenlignende undersøgelser, hvis udvælgelse af studiepopulationen eller bortfald medfører systematiske fejl.

Selektionsproblemer i sammenlignende undersøgelser



Bortfald i sammenlignende undersøgelser

Bortfaldet afhænger **hverken** af eksponering **eller** outcome

Stikprøve			Deltagere		
Eksponering	Cases	Kontroller	Eksponering	Cases	Kontroller
+	100	60	+	50	30
0	60	100	0	30	50
I alt	160	160	I alt	80	80

OR=2,8 (1,8 - 4,4)

OR=2,8 (1,5 - 5,3)

Ingen bias

Bortfald i sammenlignende undersøgelser

Bortfaldet afhænger af **enten** eksponering **eller** outcome

Stikprøve			Deltagere		
Eksponering	Cases	Kontroller	Eksponering	Cases	Kontroller
+	100	60	+	50	60
0	60	100	0	30	100
I alt	160	160	I alt	80	160

OR=2,8 (1,8 - 4,4)

OR=2,8 (1,6 - 4,8)

Ingen bias

Bortfald i sammenlignende undersøgelser

Bortfaldet afhænger af **både** eksponering **og** outcome

Stikprøve			Deltagere		
Eksponering	Cases	Kontroller	Eksponering	Cases	Kontroller
+	100	60	+	50	60
0	60	100	0	60	100
I alt	160	160	I alt	80	160

OR=2,8 (1,8 - 4,4)

OR=1,4 (0,9 - 2,3)

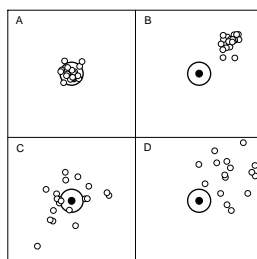
Biased estimat

Fejlkilder

- Selektionsproblemer
- Informationsproblemer
- Confounding (årsagsforveksling)

Usikkerhed og bias

Valid og præcis



Bias men præcis

Valid men usikker

Bias og usikker

Misklassifikation

Klassifikation	Sandheden	
	Syg	Ikke syg
Syg	Sandt positiv	Falsk positiv
Ikke syg	Falsk negativ	Sandt negativ

Sensitivitet: $SP/(SP+FN)$

Specificitet: $SN/(SN+FP)$

Misklassifikation

Ikke-differentieret misklassifikation

Fødselsvægt og sekretorisk otit.

Differentieret misklassifikation

Medicin under graviditeten og malformationer (recall bias)
Ryning og emfysem

Fejlkilder

- Selektionsproblemer
- Informationsproblemer
- Fejltolkning (confounding)

Confounding

Alkohol	Lungekræft	
	Ja	Nej
Ja	90	60
Nej	60	90
I alt	150	150

OR=2,3 (1,4 - 3,6)

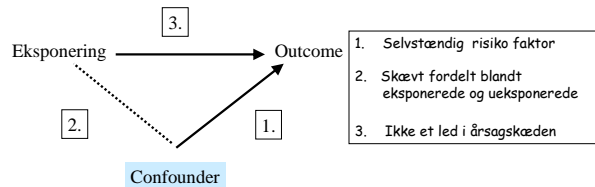
Confounding

Ryger			Ikke ryger		
Alkohol	Lungekræft		Alkohol	Lungekræft	
	Ja	Nej		Ja	Nej
Ja	80	40	Ja	10	20
Nej	20	10	Nej	40	80
I alt	100	50	I alt	50	100

OR=1,0 (0,4 - 2,3) OR=1,0 (0,4 - 2,3)

Confounding

Confoundere: at blande sammen (årsagsforveksling)



Et eksempel på confounding

Regelmæssig motion → Hjertekar sygdom

Er rygning en confounder?

Overvej de tre forhold én for én.

1. Rygere har højere risiko for hjertekarsygdom
2. Rygere motionere generelt mindre end ikke rygere
3. Rygning er næppe et led i årsagskæden

Kontrol for confounding

- Experimentelle undersøgelser
 - Randomisering
- Observationelle studies
 - Designfase
 - Restriktion
 - Matching
 - Analysefase
 - Stratifikation
 - Regressionsanalyse

Restriktion

Ikke ryger

Alkohol	Lungekræft	
	Ja	Nej
Ja	10	20
Nej	40	80
I alt	50	100

OR=1,0 (0,4 - 2,3)

Restriktion

• Fordele

- Let at forstå og formidle

• Ulemper

- Begrænser den eksterne validitet
- Færre deltagere
- Man kan ikke studere betydningen af faktorer som man har ekskluderet

Matching

1. Kan fjerne effekten af kendte confoundere i follow-up studier
2. Kan øge præcisionen ved at balancere antallet af cases og kontroller i de enkelte strata

1. Matching skal gøre i designfasen og er irreversibelt.
2. Man kan ikke studere effekten af variable man har matchet for
3. Kan være en tidsrøvende og dyr løsning
4. Kan føre til overmatching

Stratifikation

P-pillar og AMI

P-pillar	Cases	Kontroller	
Ja	39	24	63
Nej	114	154	268
	153	178	331

OR = 2.20

Stratifikation

Alder	P-pillar	Cases	Kontroller	OR
< 39	Ja	21	17	38
	Nej	26	59	85
		47	76	123
≥ 40	Ja	18	7	25
	Nej	88	95	183
		106	102	208

OR vægtet = 2.79

Rygning, kaffe og præterm fødsel

	Præterm	RR (95% CI)
Ikkeryger	3,8%	1,0
Ryger	5,3%	1,4 (1,2-1,9)

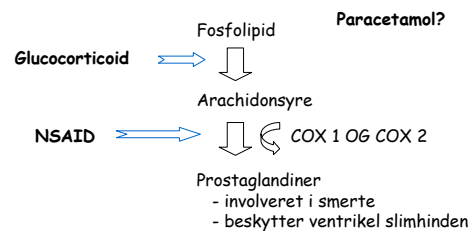
Rygning, kaffe og præterm fødsel

	Præterm	RR (95% CI)
<400 mg kaffein/dag		
Ikkeryger	4,1%	1,0
Ryger	3,9%	1,0 (0,6-1,5)
400+ mg kaffein/dag		
Ikkeryger	2,5%	1,0
Ryger	6,9%	2,8 (1,6-4,9)

NSAID, paracetamol og ventrikel blødning

Medicin	Ventrikel blødning		OR (95% CI)
	Ja	Nej	
Nej	852	6768	1,0
NSAID	193	489	3,6 (2,9 - 4,3)
Paracetamol (> 2g/dag)	52	124	2,4 (1,7 - 3,5)
NSAID + Paracetamol (> 2 g)	80	41	16,6 (11,0 - 4,9)

Smertestillende medicin og ventrikel blødning



Vurdering

- Undersøgelsen formål
- Sammenhæng mellem mellem formål og design
- Intern validitet
- Ekstern validitet