



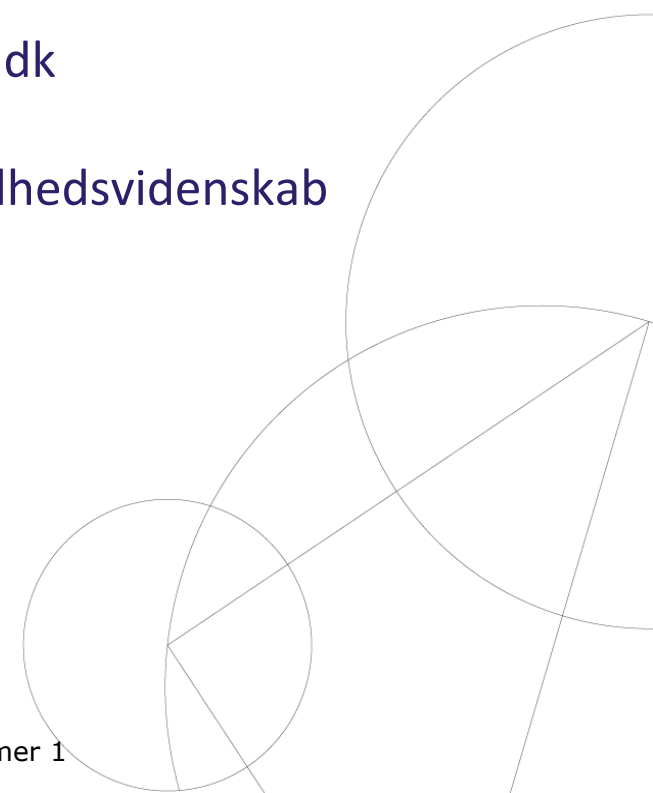
Faculty of Health Sciences



Informationsbias

Mads Kamper-Jørgensen, lektor, maka@sund.ku.dk

Afdeling for Social Medicin, Institut for Folkesundhedsvidenskab



Sidste gang

Vi snakkede om

- Tværsnitsundersøgelser: Oplysninger om eksponering og udfald er indsamlet på samme tidspunkt, individniveau. Ofte svært at afgøre temporalitet, så risiko for omvendt kausalitet. Bruges også til kun at vurdere forekomst
- Økologiske undersøgelser: Sammenligning over tid eller geografi, gruppeniveau. Pas på den økologiske fejlslutning
- Overvågningsundersøgelser: Kun oplysning om udfald, individ eller gruppe niveau. Ingen sammenligningsgruppe
- Nemt, hurtigt og billigt
- Kasuistikker, tvillinge-/søskendestudier, case-only studier, crossover studier, case-crossover studier



Typer af bias (skævhed)

Informationsbias

- Har at gøre med *informationen* om de individer der indgår i den udvalgte stikprøve

Selektionsbias

- Har at gøre med *selektionen* til den udvalgte stikprøve

Confounding

- Har at gøre med *forveksling af effekter* på grund af uombyttelighed

≥ 800 typer bias er identificeret – langt de fleste passer ind i kategorierne informationsbias eller selektionsbias

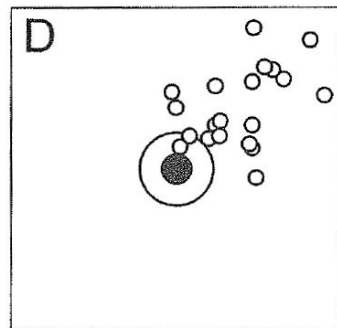
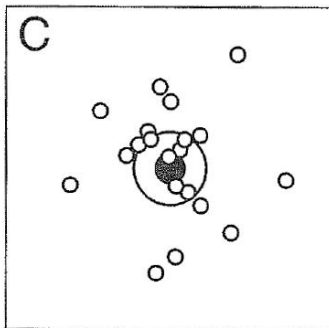
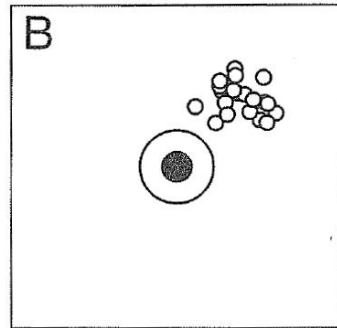
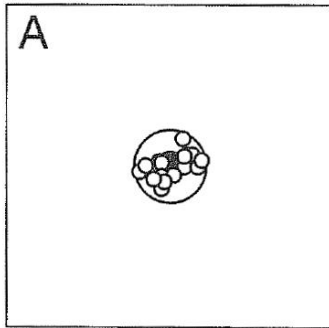


Hvorfor informationsbias?

- Fordi vi kan komme til at over- eller undervurdere associationer, og drage forkerte konklusioner hvis vi ikke har korrekte informationer om deltagerne
- Indtil videre har vi forudsat at vi havde korrekte informationer om deltagerne: Det er (næsten) aldrig tilfældet
- Vi kan have mere eller mindre korrekte informationer om eksponering, udfald, og covariater
- Usikkerheden kan skyldes f.eks. biologisk variation, dårlig hukommelse, dårlig kodning, dårligt spørgsmål, upræcist måleudstyr, uvidenhed osv.
- Informationsbias er under- eller overvurdering af en association der skyldes *systematisk* forkert information om deltagerne



Præcision og bias



BT målt én gang for 20 personer

- A) Præcis, unbiased: BT, kviksølvapp.
- B) Præcis, biased: BT, fejlkalibreret kviksølvapp.
- C) Upræcis, unbiased: BT, iPhone
- D) Upræcis, biased: BT, fejlkalibreret iPhone

Præcision og bias

TILFÆLDIGE FEJL

- Øger usikkerheden af estimatet
- Reducerer præcisionen
- Har ingen retning
- Afhænger af undersøgelsens størrelse – mindskes med øget antal observationer

SYSTEMATISKE FEJL

- Fører til bias (skævvridning)
- Reducerer validiteten
- Medfører en over- eller undervurdering af associationen
- Kan ikke mindskes ved at øge undersøgelsens størrelse – så måler vi bare galt med større præcision



Variation

- Intra-individuel variation: variation indenfor individet. F.eks. blodsukker målt hver time i løbet af en dag
- Inter-individuel variation: variation mellem individer. F.eks. forskel i socio-økonomisk position mellem individer
- Intra-observatør variation: variation indenfor observatøren. F.eks. udseende for kronologisk alder
- Inter-observatør variation: variation mellem observatører. F.eks. meget lav/høj vægt



Usikkerhed og bias

Hvad kan man gøre ved det?

- Man kan øge præcisionen ved at inkludere flere personer i sit studie dvs. øge størrelsen på sin stikprøve
- Man kan ikke gøre noget ved informationsbias når først data er indsamlet
- Godt måleudstyr, træning og instruktion, standardiserede procedurer, blinding
- Hvis informationsbias ikke er til stede øges den interne validitet



Sensitivitet og specificitet

- Sensitivitet: en tests evne til at klassificere sandt positive (SP) som sandt positive. Beregnes som $SP/(SP+FN)$
- Specificitet: en tests evne til at klassificere sandt negative (SN) som sandt negative. Beregnes som $SN/(FP+SN)$

	Syg	Ikke-syg
Syg	SP	FP
Ikke-syg	FN	SN
I alt	SP+FN	FP+SN

- Bruges oftest ifm. kvaliteten af en test, men kan også bruges om hvor godt en variabel beskriver den sande værdi



Misklassifikation

- Forkert klassifikation af deltagerne
- Hvis misklassifikation er lige sandsynlig i de sammenlignede grupper snakker vi om non-differentiel misklassifikation
- Hvis misklassifikation *ikke* er lige sandsynlig i de sammenlignede grupper snakker vi om differentiel misklassifikation
- Både differentiel og non-differentiel misklassifikation betegnes informationsbias



Eksempler fra egen forskning

Uvidenhed

- Færre voksne Amerikanere end forventet har modtaget blod

Kultur

- Franskmændene drikker overraskende lidt alkohol

Dårligt spørgsmål

- Færre danske børn end forventet er motorisk alderssvarende



Hvorfor det?

Dagens afstemning på

- www.madskamper.dk/afstemning



Misklassifikation

Fiktivt kohortestudie af sammenhæng mellem alkoholforbrug og selv vurderet helbred med dårlig registrering af alkoholforbrug

	Godt	Dårligt	I alt
Afholdende	236	59	295
Drikkende	846	419	1265
I alt	1082	478	1560

- Sand information om alkoholforbrug: RR=1,66



Non-differentiel misklassifikation

	Godt	Dårligt	I alt
Afholdende	321	101	422
Drikkende	761	377	1139
I alt	1082	478	1560

- Sand information om alkoholforbrug: $RR=1,66$
- 10% af drikkende bliver misklassificeret: $RR=1,38$



Non-differentiel misklassifikation

	Godt	Dårligt	I alt
Afholdende	405	143	548
Drikkende	677	335	1012
I alt	1082	478	1560

- Sand information om alkoholforbrug: $RR=1,66$
- 10% af drikkende bliver misklassificeret: $RR=1,38$
- 20% af drikkende bliver misklassificeret: $RR=1,27$
- Association går mod ingen forskel (dvs. 0 hvis absolut og 1 hvis relativ)



Differentiel misklassifikation

	Godt	Dårligt	I alt
Afholdende	236	101	337
Drikkende	846	377	1223
I alt	1082	478	1560

- Sand information om alkoholforbrug: $RR=1,66$
- 10% af drikkende bliver misklassificeret , kun blandt personer med dårligt selv vurderet helbred: $RR=1,03$



Differentiel misklassifikation

	Godt	Dårligt	I alt
Afholdende	236	143	379
Drikkende	846	335	1181
I alt	1082	478	1560

- Sand information om alkoholforbrug: $RR=1,66$
- 10% af drikkende bliver misklassificeret , kun blandt personer med dårligt selv vurderet helbred: $RR=1,03$
- 20% af drikkende bliver misklassificeret , kun blandt personer med dårligt selv vurderet helbred: $RR=0,75$
- **Kan vende op og ned på konklusionen**



Differentiel misklassifikation

Case-controlundersøgelser, typiske eksempler

- Recall-bias: cases husker eksponeringer anderledes (typisk bedre) end kontroller. Ikke det samme som dårlig hukommelse
- Interviewerbias: Interviewer spørger anderledes (typisk mere uddybende) om eksponeringer blandt cases end blandt kontroller

Kohorteundersøgelser, typiske eksempler

- Detektionsbias: eksponerede har anderledes (typisk større) risiko for udfald end ikke-eksponerede
- Interviewerbias: Interviewer spørger anderledes (typisk mere uddybende) om udfald blandt eksponerede end blandt ikke-eksponerede



Konsekvenser

Non-differentiel misklassifikation

- Associationen går (næsten altid) mod ingen forskel
- **Differentiel misklassifikation**
- Alt efter retningen af misklassifikationen over- eller undervurderes associationen



Næste gang

Selektionsbias

- Dvs. de skævheder der kan opstå på grund af den måde hvorpå man har udvalgt sin stikprøve



Afleveringsopgave

- Aflevering er frivillig, men stærkt tilrådet
- Sæt jer (så vidt mulig) hver for sig
- Opgavetekst i kompendiet og på hjemmesiden
- Ingen hjælp fra mig eller medstuderende
- Vi slutter kl. 12
- I må gerne arbejde videre derhjemme
- Send jeres besvarelse til mig inden midnat på maka@sund.ku.dk
- Gitte giver jer individuel og fælles feedback d. 30. maj

