

# Epidemiologi

## Begreber og metoder

Marianne Lindahl, fysioterapeut, MPH, Carsten Bogh Juhl, fysioterapeut, MPH  
Fysioterapeutuddannelsen CVU Syd, Næstved

Lindahl M, Juhl CB 2005, 25. februar)  
*Epidemiologi – begreber og metoder. Forskning i Fysioterapi (online), 2. årg.*,  
s. 1-9: URL: <http://www.ffy.dk/sw4732.asp>

Denne artikel er den første af to, som tilsammen beskriver grundlæggende begreber, metoder og associationsmål inden for epidemiologi. Den første artikel giver en kort indføring i de vigtigste begreber inden for epidemiologien, samt hvilke forskningsmetoder der traditionelt er knyttet til feltet. Den anden artikel beskriver associationsmålene relativ risiko (RR) og odds ratio (OR). Der tages udgangspunkt i eksempler om bevægeapparatsbesvær.

Epidemiologi stammer fra græsk (epi = hos, demos = folk, logos = lære) og er studiet af helbredsphenomener i befolkningsgrupper. Epidemiologi er læren om sygdommes hyppighed og fordeling i en befolkning eller befolkningsgruppe, samt hvilke årsagsfaktorer eller determinanter, der er bestemmende for sygdommens opståen og udvikling. Epidemiologisk forskning vedrører ikke kun sygdomme, der optræder i epidemier, men alle sygdomme og helbredstilstande. Det kan være sygdomsforekomsten i den danske befolkning mellem 16 og 66 år eller det kan være forekomsten af en bestemt sygdom, f.eks. nakke/skulderbesvær blandt en bestemt gruppe lønmodtagere. Epidemiologi handler således primært om grupper af mennesker i samfundet og tidsperspektivet for udvikling af sygdom.

Epidemiologi giver grundlaget for at vurdere, hvilket omfang en sygdom har, hvorfor mange sundhedsvidenskabelige undersøgelser i deres baggrund henviser til epidemiologiske undersøgelser. Omfanget af en sygdom skal i denne sammenhæng forstås som, hvor mange personer, der har sygdommen, hvor mange nye tilfælde der er om året, hvilke risikofaktorer der er for en sygdoms opståen, og eventuelt hvad der kan

forudsige forløbet. Epidemiologi belyser endvidere, hvor stor samfundsmæssig betydning en sygdom har, at belyse omfanget af sygdommens følger i form af f.eks. sygefravær, tab af arbejde, lægekontakter og hospitalsindlæggelser.

En grundlæggende antagelse inden for epidemiologi er, at sygdom ikke opstår tilfældigt, men skyldes nogle faktorer, som kan identificeres (8). Det afføder en anden grundlæggende antagelse, nemlig at sygdomme kan forebygges, derfor udgør epidemiologi en væsentlig del af folkesundhedsvidenskaben (public health) og en del af grundlaget for sundhedsfremme og forebyggelse. Epidemiologi beskæftiger sig både med ætiologi (årsager til sygdom) og patogenese (udvikling af sygdom) samt salutogenese (faktorer, som bevarer helbredet). Som en fælles funktion af disse begreber ligger der en stor udfordring i at udforske, hvorfor nogle bliver syge og andre ikke, selv om de udsættes for de samme risikofaktorer. I det følgende omtales de syge som cases.

Inden for epidemiologien ansues sygdommes opståen og udvikling som et samspil mellem levekår, livsstil, arbejdsforhold, sårbarhed, sociale forhold og miljø. Deri ligger også, at epidemiologien kan erkende sammenhænge, som ikke umiddelbart kan forklares med patologiske vævsforandringer, derfor indgår epidemiologi som en del af grundlaget for det socialmedicinske felt.

I tilrettelæggelsen af undersøgelser kan bedømmelse af årsag og virkning foretages på forskellig måde, men som hovedregel kan man sige, at undersøgelsen tilsigter at beskrive en association eller forklare den.

## Deskriptiv epidemiologi

Deskriptiv epidemiologi er, som navnet siger, beskrivende undersøgelser og beskæftiger sig med bestemmelse af en sygdoms hyppighed, i hvilke befolkningsgrupper den optræder, samt hvad den medfører i form af morbiditet (sygelighed) og mortalitet (dødelighed). Her går man efter en kvantificering af nogle mønstre. Det vil sige, at man tilstræber at kunne sige noget om, hvem i en population, der bliver ramt, f.eks. aldersgrupper, køn, geografisk område, socialgruppe, jobtype og livsstil. Man kan også være interesseret i hvornår, man bliver ramt, f.eks. om der er sæsonudsving eller sygdommen først optræder efter lang tids påvirkning. Herigennem får man et billede af et helbredsfaenomens omfang og mulighed for overvågning af udviklingen.

Et eksempel på deskriptive epidemiologiske undersøgelser er befolkningsundersøgelserne Sundhed og Sygelighed i Danmark, SUSY. Disse undersøgelser foretages af Statens Institut for Folkesundhed (SIF) tidligere Dansk Institut for Klinisk Epidemiologi (DIKE). På baggrund af omfattende spørgeskemaundersøgelser foretaget i 1987, 1994 og senest i år 2000 belyses befolkningens sundhedstilstand. Deskriptiv epidemiologi bidrager med værdifuld viden og lægger grunden til hypoteser (8), som mere dybdegående forskning må analysere nærmere. En måde at gå nærmere ind i problemstillingerne er gennem analytisk epidemiologi.

## Analytisk epidemiologi

Analytisk epidemiologi går mere i dybden og hører til den forklarende forskningstype. Her har man en hypotese om, hvorfor sygdom udvikler sig, og man sammenligner grupper i befolkningen med henblik på at forklare forskelle. Der søges efter årsager til sygdom og død, og forekomsten og fravær af sygdom analyseres ud fra, hvad den enkelte har været udsat for eller ikke været udsat for af risikofaktorer, det vil sige eksponering. Man søger gennem statistiske metoder at isolere en risikofaktors betydning og vurdere, hvor stor en del af sygdomsforekomsten, der kan tilskrives denne bestemte risikofaktor. Man afdækker mønstre og fællestræk ved case-ud fra undersøgerens valg af hvilke variable, man anser for at være relevante for problemstillingen. Selv om man i analytisk epidemiologi går

detaljeret ned i de enkelte cases, undersøger hvilke risikofaktorer den enkelte har været udsat for og relaterer det til forekomsten af sygdom, er undersøgelserne stadig populationsbaserede. Analytiske undersøgelser kan være baseret på data fra kohorteundersøgelser og case-kontrolundersøgelser, som beskrives senere, men kan også være interventionsstudier.

En undersøgelse kan godt benytte sig af både deskriptive og analytiske metoder. Eksempelvis i Harrebys undersøgelse af risikofaktorer for lændesmerter blandt skoleelever (7). Her har man dels beskrevet problemets omfang, (prævalens, se nedenfor), men også indhentet oplysninger om risikofaktorer. Det har gjort det muligt nærmere at analysere, hvad der har betydning for at skoleeleverne udvikler lændesmerter.

## Mål for hyppighed

Mål for, hvor hyppigt en sygdom forekommer, kan udtrykkes på flere måder, og i det følgende gennemgås de mest almindelige, nemlig rate, prævalens og incidens.

### Rate

Overordnet udtrykkes hyppigheden af en sygdom (også kaldet frekvens) i procent eller i rate. Når sygdomsfrekvensen udtrykkes i procent er enheden i tæller og nævner den samme, nemlig mennesker, og tællerens antal cases er indeholdt i nævnerens tal for hele populationen.

$$\frac{\text{antal cases}}{\text{hele populationen}}$$

Når man benytter rate for sygdomsfrekvensen, har tæller og nævner forskellige enheder, nemlig henholdsvis personer og personår. Rate benyttes, når man undersøger en befolkningsgruppe, hvor det ikke er alle, som hele tiden er i risiko for at få sygdommen. Man tager derfor ofte, hvor mange år hver observeret person bidrager med (persontid eller risikotid), og samlet får man således et udtryk for sygdommens forekomst pr. f.eks. 1.000 eller 100.000 personår. Derved får man raten, som udtrykker antal syge i forhold til en bestemt dimension for eksempel antal personår.

$$\frac{\text{antal cases}}{\text{antal personår}}$$

### Prævalens og incidens

Forekomsten eller hyppigheden af sygdommen på et givet tidspunkt udtrykkes som prævalens og angives i procent af en population. For eksempel angives prævalensen af reumatoid artrit i befolkningen til 0,47 % (2), dvs. at der er 0,47 % af den danske befolkning har fået konstateret reumatoid artrit. Prævalens er et tværnsnitsmål, men begrebet repræsenterer alligevel en vis dynamik, idet det indeholder en tilgang af nye cases lige så vel som en nedgang, fordi de syge (cases) bliver raske eller dør. Prævalensen afspejler således, sammenholdt med incidensen, om sygdommen er af lang varighed. Prævalensen kan måles for en bestemt periode og kaldes så periodeprævalens, se i øvrigt faktaboks nedenfor. Et andet meget anvendt udtryk i forbindelse med bl.a. lænderygbesvær er livstidspræ-

valens, det vil sige, hvor mange der i løbet af livet får ondt i ryggen.

Nye tilfælde af sygdommen udtrykkes som incidens, det vil sige antal nye cases i en bestemt periode. Eksempelvis angives incidensen af reumatoid artrit til 0,04 % i MTV rapporten, hvilket betyder at 0,04 % af befolkningen hvert år diagnosticeres til at have reumatoid artrit, svarende til 4 ud af hver 10000 mennesker (2). Her er tidsperspektivet longitudinelt, idet man observerer i en vis periode. Prævalensen af en sygdom er således både afhængig af incidensen og af hvor mange, der holder op med at have sygdommen, enten fordi de bliver raske eller dør. Incidensen afspejler risikoen for at udvikle en sygdom, mens prævalensen afspejler risikoen for allerede at have den. Hvis der for eksem-

#### Faktaboks 1. Forskellige udtryk for mål for hyppighed

Prævalens (P)	Forekomsten på et bestemt tidspunkt	$P = \frac{\text{antal cases}}{\text{antal mennesker i en population}}$
Periodeprævalens	Forekomsten i løbet af en bestemt periode	$P = \frac{\text{antal cases}}{\text{antal mennesker i en population}}$ inden for et år
Punktprævalens	Forekomsten på et eller flere helt fastlagte kalendertidspunkter, eller et bestemt tidspunkt i den enkeltes liv.	
Incidens (I)	Antal nye cases i populationen i løbet af en given periode	
Cumuleret incidens (CI), også kaldet cumuleret incidens proportion	Opsummeret antal nye cases blandt personer i risiko	$P = \frac{\text{antal nye cases gennem en given periode}}{\text{totale population i risiko}}$
Incidens rate (IR) eller på engelsk incidence density (ID)	Antal nye cases pr. personår i risiko (risikotid)	$P = \frac{\text{antal nye cases gennem en given periode}}{\text{totale antal observerede personår}}$
Cumulativ incidens proportion	Antal cases opsummeret over flere delperioder	

pel er 20 % af en befolkningsgruppe, der har en sygdom er risikoen for at få sygdommen betydelig, under forudsætning af at alle i en gruppe har samme risiko for at få sygdommen. Ved incidens er det særlig relevant at forholde sig til begrebet risikotid, for hvis man tager en procentdel af befolkningen, er det ikke alle, som overhovedet er i risiko for at udvikle sygdommen. Det tager udtrykket incidensrate højde for, idet incidensen her kun er ud af personer, som er i risiko for at få sygdommen og altså ikke hele populationen. Man medregner blandt andet ikke dem, som allerede har sygdommen eller dem som af én eller anden grund ikke kan få den. Personerne indgår derfor med forskellig risikotid. Se i øvrigt faktaboks side 3.

### Undersøgelsestyper

Da det er en grundlæggende antagelse, at sygdomme ikke fordeles tilfældigt, må epidemiologiske undersøgelser tilrettelægges på en sådan måde, at tilfældigheder udelukkes eller kontrolleres. Inden for epidemiologi benytter man sig af mange forskellige måder at designe undersøgelser på, afhængig af hvilken type viden man ønsker. I det følgende omtales aspekterne ved tværsnitundersøgelser, kohorteundersøgelse, korrelationsundersøgelser samt case-kontrolundersøgelser.

Tidsperspektivet er af stor betydning inden for epidemiologien. Hvornår en sygdom kommer i udbrud i forhold til, hvornår man har været udsat for en risikofaktor, er et centralt spørgsmål, når man vil sige noget om en sammenhæng – også kaldet association. Jo længere tid, der går mellem eksponering (dvs. udsættelse for en risikofaktor) og sygdommens optræden, jo sværere er det med stor sikkerhed at bestemme, hvor meget og hvor længe, man skal være udsat for en risikofaktor for at udvikle sygdommen. Det gør det vanskeligt at sætte effektivt ind med forebyggelse. Samtidig kan der inden for et langt tidsperspektiv ske ændringer både i definitionen af sygdommen og den enkeltes eksponering, hvilket yderligere vanskeliggør sammenligning.

En undersøgelses tidsperspektiv kan være longitudinelt, det vil sige længdegående, hvor man følger en gruppe mennesker og bedømmer hver person mindst to gange med et vist tidsrum imellem, og således tidsmæssigt følger årsag

og virkning. Det kan gøres prospektivt (fremadskuende) eller retrospektivt (bagudskuende). Tidsperspektivet kan også være som tværsnit, idet man her tager et øjebliksbillede af situationen.

### Tværsnitundersøgelser

Det kan give et godt overblik over helbredstilstanden at undersøge en befolkningsgruppe ved en tværsnitundersøgelse. Metoderne kan være både kliniske undersøgelser og spørgeskema eller kombinationer heraf. Hvad man vælger afhænger af, hvad man vil vide. Tværsnitundersøgelser er velegnede til at sige noget om hyppigheden af både årsag og virkning og gentagne tværsnitundersøgelser som SUSY (15) kan, netop fordi de omfatter store populationer, give et fingerpeg om udviklingen af sygdom i en stor befolkningsgruppe i løbet af en årrække.

Ved tværsnitundersøgelser måles årsag (risikofaktor) og virkning (sygdom) samtidig, og deltagerne vælges enten efter risikofaktoren (årsagen) eller virkningen (sygdommen). Det gør det imidlertid vanskeligt at udlede noget om en association, dvs. om der er en sammenhæng, idet der kan være andre forhold, som har haft betydning. Eksempelvis har man inden for forskning i arbejdsmiljø begrebet healthy-worker effect, hvilket betyder, at dem, der arbejder på det tidspunkt tværsnitundersøgelsen gennemføres, er de raske arbejdere (5). De arbejdere, som blev syge af arbejdsmiljøet, har søgt væk fra det bestemte job, som undersøges og er enten gået til andre erhverv eller er udstødt fra arbejdsmarkedet netop på grund af sygdom.

### Kohorteundersøgelser

Kohorteundersøgelsen anses for at være et meget stærkt design inden for epidemiologien og betegner en undersøgelsestype, hvor man følger en stor gruppe mennesker over en længere periode. Man sammenligner sygdomsforekomsten alt efter udsættelsen for én eller flere risikofaktorer. Udtrykket kohorte stammer fra romertiden og betegner en hærenhed. Når soldaten indgik i kohorten, vedblev han at være der, og kohorten blev ikke suppleret med nye soldater, når nogen blev dræbt eller af anden årsag måtte forlade hæren. Der kan således godt ske et vist frafald fra kohorten, men der bliver ikke suppleret med nye deltagere.

Afgørende for kohorteundersøgelsen er, at deltagerne ved undersøgelsens start ikke har den pågældende sygdom, men godt kan have en eller flere risikofaktorer for sygdom. Deltagerne undersøges måske klinisk og udspørges desuden om alle relevante faktorer, herunder de risikofaktorer, som forskeren interesserer sig for. Undersøgelsestypen er longitudinel, også kaldet follow-up, og såvel den kliniske undersøgelse som spørgeskemaundersøgelsen (evt. kun det sidste) gentages flere gange i perioden, men under alle omstændigheder mindst to gange – nemlig ved baseline og ved undersøgelsens slutning. Personerne, som indgår i kohorten, skal være udvalgt ved en tilfældig stikprøve, som afspejler den befolkningsgruppe, de skal repræsentere eksempelvis befolkningsundersøgelserne i Glostrup, hvor befolkningen i Glostrup i vid udstrækning er repræsentativ for den danske befolkning.

Undersøgeren har besluttet, hvilke risikofaktorer, der skal medtages og skal også beslutte, hvilke faktorer, som ellers kan have indflydelse og påvirke resultatet. Ved undersøgelsens start (baseline) undersøges den udvalgte kohorte for eksempel ved en klinisk undersøgelse og/eller spørgeskema. Her afdækkes demografiske data (alder, køn, ægteskabelig status, antal børn, arbejde osv.) helbredsstatus, og desuden hvilke risikofaktorer, man er udsat for. Efter baselineundersøgelsen venter man en længere periode, hvorefter man undersøger, hvordan det er gået kohorten. Det kan igen være klinisk undersøgelse og/eller spørgeskema.

Nogle kohorteundersøgelser forløber over temmelig mange år, hvor kohortens deltagere med års mellemrum undersøges. En stor dansk kohorteundersøgelse er f.eks. Østerbroundersøgelsen, hvor man siden 1976 har fulgt ca. 14.000 mennesker i København (11,13). Ved at sammenholde forskellige variable har man for eksempel belyst, hvilken betydning fysisk aktivitet har for dødeligheden i populationen, og man fandt, at dødeligheden blev reduceret ved et vist fysisk aktivitetsniveau. Det vil sige, at man har anvendt data til en analytisk undersøgelse ved at sammenholde forskellige faktorer (11).

Ud over at finde ud af hvem der bliver syge i en kohorte, kan en kohorteundersøgelse også tage udgangspunkt i en gruppe syge. Her vil det så

være undersøgelsens formål at følge disse syge og se, hvordan de klarer sig, således at det bliver muligt at sige, hvem der klarer sig godt og hurtigt vender tilbage til normal funktion. Eksempelvis er der inden for rygområdet lavet en del undersøgelser af, hvilke faktorer der har betydning for, om lænderygbesvær udvikler sig til en kronisk tilstand (14). I Danmark har man dog enestående gode muligheder for gennem registerdata at finde ud af, hvordan det er gået folk for eksempel gennem cpr.-registeret, Dødsårsagsregisteret, Cancerregisteret og Landspatientregisteret.

En epidemiologisk undersøgelse kan også være en interventionsundersøgelse med kontrolgruppe (8). Det vil sige, at man randomiserer deltagerne i to eller flere grupper, som får forskellige former for intervention. Der laves baselinemåling inden randomiseringen, og der måles igen efter en passende periode i forhold til at interventionen skal kunne have haft effekt på de variable, der måles på. Når man intervenserer i epidemiologiske undersøgelser, er man særlig interesseret i outcomemål i form af sygelighed og dødelighed over en længere periode. Det er derfor en kvalitet, hvis der laves follow-up efter for eksempel 6-12 måneder eller 2-5 år. Et eksempel på intervention med 5 års follow-up er Indahls store undersøgelse af intervention overfor patienter med lænderygbesvær (9).

### *Korrelationsundersøgelser*

Ofte er det ikke muligt at tilrettelægge en undersøgelse med det stærkeste design, som helt sikkert giver svaret på, om en bestemt risikofaktor er årsag til en bestemt sygdom. Det kan være en langvarig proces at gennemføre en prospektiv kohorteundersøgelse, så derfor nøjes man ofte i første omgang med at undersøge, om der er en korrelation mellem risikofaktoren og sygdommen.

Det, man hermed er interesseret i, er i hvor høj grad, der er samvariation mellem to variable. Korrelation er et deskriptivt mål for association, hvor man ikke ser på de enkelte individer i populationen, men udelukkende på variationen inden for grupper. Det kan eksempelvis være, om der er korrelation mellem selvrapporterede smerter i bevægeapparatet og antal sygefraværsdage. Det vil sige, at jo større den ene variabel er (smerterne), jo større er også den anden variabel (sygefraværsdagene).



Selv om man finder en stærk korrelation, er det imidlertid ikke alene på den baggrund muligt at slutte kausalt, dvs. at smerter i bevægeapparatet er årsag til højt sygefravær. Korrelation beregnes statistisk, således at der fremkommer en korrelationskoefficient, som kan ligge mellem  $-1$  og  $+1$ . Hvis korrelationskoefficienten er  $-1$  viser beregningen, at der er en perfekt negativ korrelation, dvs. at når den ene variabel bliver større (f.eks. funktionsevnen) bliver den anden mindre (f.eks. smerterne). ) Hvis korrelationskoefficienten er  $0$ , betyder det, at der absolut ikke er nogen korrelation, men hvis den er  $+1$  indikerer det en perfekt positiv korrelation (figur 1 side 7). Det vil sige, at når den ene variabel bliver større (f.eks. graden af fysisk aktivitet, bliver den anden også større (f.eks. graden af funktionsevne).

Hvis vi vil være sikre på, at den fremkomne korrelation ikke er udtryk for tilfældigheder, beregnes også en  $p$ -værdi. Den beregnede  $p$ -værdi tester hypotesen  $y = 0$ , altså om den ene variabel giver en vis del af forklaringen på den anden. Hvis man f.eks. har en korrelationskoefficient på  $0,7$ , udtrykker tallet korrelationen mellem de to variable. Hvis man sætter  $0,7$  i  $2$ . potens får man  $0,49$ , som er udtryk for, hvor stor en del af variationen af den ene variabel, der kan forklares med variationen af den anden variabel. Dette kaldes også forklaringskoefficienten.

Ofte anvendes et Bland Altman plot til at vurdere sammenhængen mellem to variable. Ud af  $x$ -aksen i et koordinatsystem afsættes gennemsnittet af de to variable og ud af  $y$ -aksen afsættes forskellen mellem de to variable. Herved kan det ses, om der er en systematisk forskel mellem de to variable. For eksempel om forskellen mellem de to bliver større ved høje værdier af de to variable (data kommer til at ligne en trompet), eller det kan være, at den ene variabel systematisk scorer højere en den anden (alle data ligger på samme side linien  $y = 0$ ) og at den forskel bliver større ved høje værdier af de to variable (alle data ligger på samme side af aksens  $y = 0$ . og afstanden til aksens stiger med stigende værdier).

### *Case-kontrolundersøgelser*

Det er ikke altid, at man kan eller har mulighed for at vente på resultaterne fra en longitudinel kohorteundersøgelse. Hvis der for eksempel er tale om en sjælden lidelse, kan man komme til

at vente meget længe og skal inddrage en meget stor befolkningsgruppe for at kunne sige noget om en association. Her er case-kontrolundersøgelsen et godt alternativ..

Case-kontrolundersøgelsen kan være både retrospektiv og prospektiv. Udgangspunktet for en retrospektiv case-kontrolundersøgelse er, at undersøgeren ser tilbage på, hvad folk blev syge af. Man udvælger tilfældigt en række cases, det vil sige mennesker som har den pågældende sygdom. Herefter finder man nogle mennesker, der ligner de syge på alle væsentlige områder (f.eks. alder, køn, arbejde), men ikke har sygdommen. Undersøgelsen afdækker dernæst hvem i de to grupper, der har været udsat for pågældende risikofaktor. Kontrolgruppen kan være væsentligt større end casegruppen, op til fire gange så stor. Alligevel kan analysen ikke sige noget om sygdommens hyppighed i en population, man har jo selv på forhånd bestemt det procentvise forhold mellem syge og raske.

Fordelen ved case-kontrol designet er, at det er forholdsvis nemt og hurtigt at indsamle oplysningerne, idet især casene som regel er nemme at finde. Ulempen er, at det kan være vanskeligt at sige noget om tidsperspektivet mellem udsættelse for risikofaktoren og udvikling af sygdommen. Derudover er det vanskeligt helt præcist at definere kontrolgruppen.

En væsentlig problemstilling ved de retrospektive case-kontrolundersøgelser er, at når mennesker bliver syge, vil de som regel spekulere på årsagen, og hvis de indgår i en undersøgelse, der skal afdække risikofaktorer, vil deres hukommelse, for hvad de har været udsat for igennem nogle år, sandsynligvis være betydelig bedre end hos de raske kontroller. Det kaldes recall-bias eller informationsbias.

### **Variable**

Variable er de faktorer, der antages at have betydning for sygdommen og de kendetegn, man beslutter at undersøge ved sygdommen. Man deler op i afhængige og uafhængige variable, og forskellige benævnelser er beskrevet i faktaboks 2 på side 7. Inden for epidemiologi er det særligt udtrykkene determinanter eller prædiktorer, som har interesse. Udfaldet (sygdommen) kaldes outcome og måles i én eller flere såkaldt

**Faktaboks 2 viser betegnelser for variable**

Uafhængige variable	Afhængige variable
Determinant	Outcome
Prædiktorer	Udfald
Indikator	Effekt
Eksposition	Sygdom
Risikofaktor	
Årsag	
Forklarende variable	

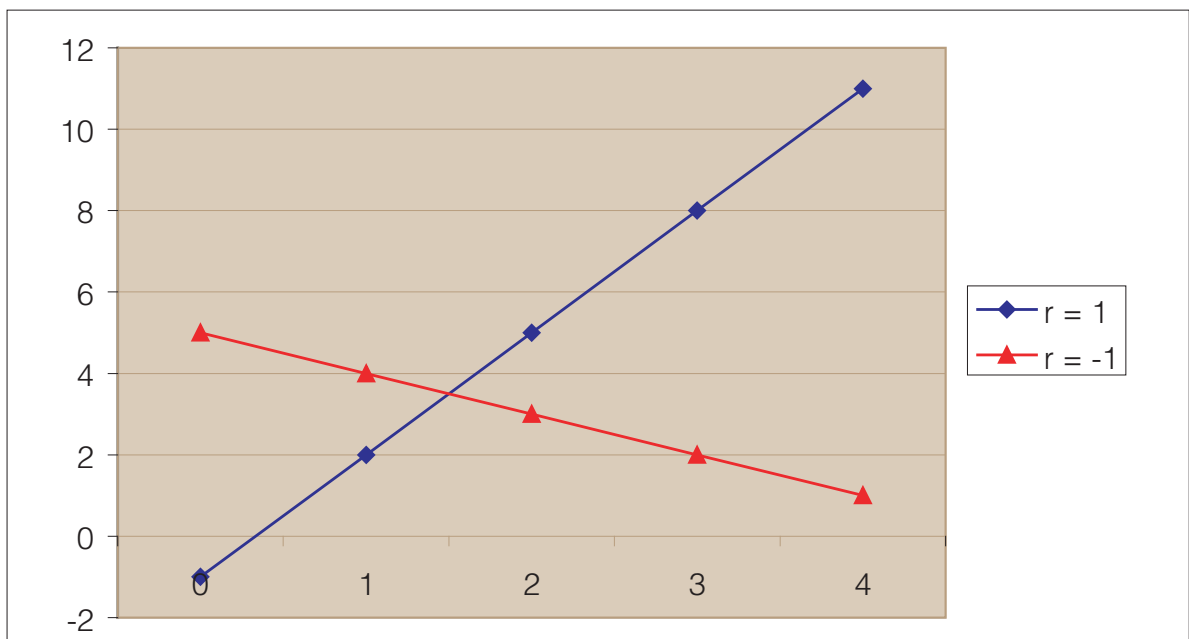
afhængige variable; det vil sige, at det at få sygdommen er afhængigt af noget. Det er de uafhængige variable, som inden for epidemiologi er årsagsfaktorer og andre faktorer, som kan have betydning for udfaldet. I interventionsundersøgelser er interventionen en uafhængig variabel på linie med andre forhold, som kan have betydning, f.eks. alder, køn. Sidstnævnte kaldes ofte baggrundvariable og er demografiske oplysninger.

Betegnelserne bruges lidt i flæng, men man kan inden for epidemiologi sige, at en determinant er et svagere udtryk end prædiktor og indikator, som igen er svagere end risikofaktor og årsag. Outcome måles jo på forskellig måde alt afhæn-

gig af, hvad man vil undersøge, og der findes utallige helbredsmaal, som udtrykker forskellige variationer og grader af sundhed og sygelighed.

Det store problem i tilrettelæggelsen af epidemiologiske undersøgelser er, hvor mange og hvilke variable man skal inddrage. For nogle sygdomme er sammenhængen forholdsvis enkel, men inden for bevægeapparatsbesvær er det ofte meget vanskeligt entydigt at pege på nogle få variable. Et eksempel er inden for lænderygbesvær, hvor der er lavet utallige undersøgelser af, hvad der øger risikoen for lænderygbesvær, og hvilke prædiktorer der er for at udvikle en kronisk tilstand. Det afspejles i MTV-rapporten Ondt i Ryggen (14), hvor man kan pege på nogle risikofaktorer som påviste, men lige så mange som mulige. Det skyldes ikke mangel på undersøgelser inden for området, men derimod den kompleksitet der ligger bag problemet, og som er vanskeligt at indfange i én undersøgelse. Et andet eksempel er Bigos et al.'s store longitudinelle, prospektive undersøgelse af risikofaktorer for udvikling lænderygbesvær. Den er foretaget på over 3.000 ansatte på Boeing fabrikkerne, og der indgår mere end 60 variable inden for antropometri, klinisk undersøgelse, funktion, medicinsk historie, arbejdspladsfaktorer (fysiske og psykosociale) samt psykologiske faktorer (4). Hver variabel

**Figur 1 viser eksempler på korrelationskoefficienter.**



dækker over flere spørgsmål/oplysninger, som indgår i datamaterialet. I en dansk undersøgelse af risikofaktorer og udvikling af lænderygbesvær indgår 58 variable og der er ikke megen overlapning i de to undersøgelsers valg af variable (3).

## Standardisering

Når man sammenligner befolkningsgruppers sygelighed eller dødelighed, skal man sikre, at sammenligningen er fair. Det vil sige, at man skal f.eks. ikke sammenligne forekomsten af osteoporose blandt 60+-årige kvinder i X-købing amt med 60+-årige kvinder i Y-købing amt, hvis der er flere meget gamle blandt gruppen af kvinder i Y-købing amt. Man vil sandsynligvis få det resultat, at forekomsten af osteoporose er større i Y-købing amt, men sammenligningen er ikke fair på grund af forskellig aldersfordeling i populationerne. Her foretager man så det, det hedder standardisering dvs. en statistisk procedure, hvor man tager højde for aldersforskellen.

Standardisering kan være direkte eller indirekte. I direkte standardisering sammenholdes antallet og sygeligheden i de forskellige aldersgrupper i de to byer med antal og fordeling af sygeligheden i en standardpopulation, f.eks. den danske befolkning. Ved indirekte standardisering er udgangspunktet, at man vil undersøge forekomsten i X-købing, og sammenligner så med Y-købing som såkaldt referencepopulation. Så tælles, hvor meget hver aldersgruppe procentvis vægter i X-købing og sammenlignes med referencepopulationens fordeling. Man beregner derefter den forventede sygelighed ud fra referencepopulationen.

Kort sagt, når man i en epidemiologisk undersøgelse læser, at der er standardiseret for f.eks. alder, betyder det, at man har taget højde for aldersforskelle i de grupper, der sammenlignes for dermed at opnå en fair sammenligning. For yderligere gennemgang henvises til litteraturen (6,10).

Hyppigt anvendte udtryk inden for epidemiologien er standard mortality rate (SMR). Rate betyder som tidligere nævnt (side 2), at man tager højde for den tid personerne gennemsnitligt er fulgt, det vil sige antal personår. SMR udtrykker forholdet mellem det forventede antal dødsfald og det observerede antal, og i ordet standardiseret ligger som ovenfor nævnt, at man har justeret for noget. I forbindelse med SMR betyder det som regel, at man har justeret for alder, og altså taget det procentvise antal dødsfald i hver aldersgruppering og opsummeret dem. Dette tal (det forventede procentvise antal) sammenlignes med tallet for den gruppe, som man observerer (det observerede procentvise antal).

De observerede kan for eksempel være en gruppe lønmodtagere i et bestemt udsat erhverv, der sammenlignes med andre lønmodtagere. SMR udtrykker således forholdet mellem de to grupper. SMR på 140 % for den observerede gruppe betyder således en overdødelighed på 40 %.

Tæt knyttet til SMR er middellevetid, som har fået meget opmærksomhed i Danmark, fordi det ses som udtryk for befolkningens generelle helbredstilstand. Middellevetid sammenlignes ofte på tværs af landegrænser og særligt i lande, vi normalt sammenligner os med, er middellevetiden interessant. Middellevetiden forklares, som den tid en nyfødt kan forventes at leve med udgangspunkt i dødeligheden i de forskellige aldersgrupper (10). Et andet hyppigt anvendt mål er standard hospitality rate, som udtrykker antal hospitalsindlæggelser i populationen. Det anvendes for eksempel til sammenligning af lønmodtagergrupper, og giver et direkte, men uspecifikt, mål for den mere alvorlige sygelighed. Det giver imidlertid kun mening, hvis man justerer for alder og køn.



## Referencer

1. Andersen LB, Schnohr O, Schroll M, Hein HO. All-cause mortality associated with physical activity during leisure time, work, sports and cycling to work. *Arch Intern Med* 2000;160(11):1621-28.
2. Bech K, Andersen LS, Lange M, Sigmund H (red). *Leddegigt – medicinsk teknologivurdering af diagnostik og behandling*. København: Sundhedsstyrelsen; 2002
3. Biering-Sørensen F, Thomsen C. Medical, social and occupational history as risk indicators for low-back trouble in a general population. *Spine* 1986;11:720-25.
4. Bigos S, Battie M, Spengler D et al. A longitudinal, prospective study of industrial back injury reporting. *Clin Orthop Related Res* 1992;279:21-34.
5. Borg V, Burr H. *Danske lønmodtageres arbejdsmiljø og helbred 1990-95*. København: Arbejdsmiljøinstituttet; 1997.
6. Gerstmann B. *Epidemiology kept simple. An introduction to classic and modern epidemiology*. Toronto: Wiley-Liss; 1998.
7. Harreby MS, Nygaard B, Jessen T et al. Risikofaktorer for lændesmerter hos 1.389 skolelever i 8.- og 9 klasse. Et epidemiologisk studie. *Ugeskr Læger* 2001;163(3):282-86.
8. Hennekens C, Buring J. *Epidemiology in medicine*. Boston/Toronto: Little Brown and Company; 1987.
9. Indahl A, Haldorsen E, Holm S, Reikerås O, Ursin H. Five years follow-up study of a controlled clinical trial using light mobilization and an informative approach to low back pain. *Spine* 1998; 23: 2625-30.
10. Olsen J, Overvad K, Juul S. *Analytisk epidemiologi. En introduktion*. København: Munksgaard Danmark, 2002.
11. Osler M. *Befolkningsundersøgelsers bidrag til forskning i folkesygdomme med fokus på danske kohorter*. *Ugeskr Læger* 2004;166:1428-31.
12. Rothman K. *Epidemiology. An introduction*. New York: Oxford University Press; 2002.
13. Schnohr P. *Østerbroundersøgelsen snart 25 år. (Elektronisk version)*, *Ugeskr Læger* 2000;43: 5788. Retrieved juli 2004, from [http://www.dadlnet.dk/ufl/ufl2043/v\\_p/33347.htm](http://www.dadlnet.dk/ufl/ufl2043/v_p/33347.htm).
14. *Statens Institut for Medicinsk Teknologivurdering: Ondt i Ryggen: Forekomst, behandling og forebyggelse i et MTV-perspektiv*. Medicinsk Teknologivurdering Serie B 1999 (1).
15. Kjølner M, Rasmussen N. *Sundhed og sygelighed i Danmark 2000 og udviklingen siden 1987*. Statens Institut for Folkesundhed; 2002.