

Énprosentmedisinen – en kilde til overbehandling

Mange kontrollerte kliniske studier viser små effekter av behandling, men effekten fremstilles ofte som stor. Det er mange usikkerheter ved slike undersøkelser og vanskelig å omsette resultatene til klinisk praksis på en god måte. Da kan behandling skade mer enn den gagnar. Diagnostikk og screening trappes opp, risikofaktorer gjøres til behandlingsindikasjon og overbehandlingen øker. Sykdomsbegrepet utvides og friske gjøres syke. Det kan skade den enkelte og føre til feil ressursprioritering. Basal viten samt klinisk skjønn og erfaring må være et korrektiv til resultatene fra kontrollerte kliniske studier. Bred klinisk kompetanse bør supplere epidemiologisk og statistisk ekspertise ved utarbeiding av retningslinjer og prioriteringer.

Oppgitte interessekonflikter: Ingen | Se også side 279

Ivar Følling

ivar.folling@medisin.uio.no
Endokrinologisk seksjon
Akershus universitetssykehus
1478 Nordbyhagen

Kontrollerte kliniske studier er hovedgrunnlaget for kunnskapsbasert klinisk medisin. Vi styres i betydelig grad av dette kunnskapsgrunnlaget. Det fører imidlertid til potensielt skadelig diagnostikk og overbehandling. Flere forhold bidrar. Mange påviste behandlingseffekter er små, men fremstilles likevel som store, ved at resultatene presenteres som relative istedenfor absolutte effekter. Det kan gi et feilaktig inntrykk av stor effekt av medisinske tiltak. Videre er en rekke påviste behandlingseffekter usikre. Kontrollerte kliniske studier har feilkilder og begrensninger slik at effektene av behandling langt fra blir så store som undersøkelsene tilsynelatende viser (1). Det medfører at vi overbehandler, at vi skader mer enn vi gagnar og at vi prioriterer feil. Resultatet kan bli «énprosentmedisinen».

Relative effekter villeder

Et tenkt eksempel fra en kontrollert klinisk studie er vist i figur 1a. Mens den relative risikoreduksjonen er 33 %, er den absolutte effekt bare 1 %. Det er den lille absolutte effekten som som oftest er relevant i spørsmålet om valg av terapi for den enkelte pasient. Hvor liten den er, illustreres i figur 1b, der de samme resultatene er tegnet på en annen måte. Nyttig behandling er oppnådd hos 1 % av pasientene. Hos 2 % kom det

sykdom selv om man behandlet. Hos 97 % kom det ingen sykdom uansett. Hos 99 % av forsøkspersonene er det derfor ikke påvist noen nytte.

Når man skal vurdere om en person skal behandles kan man, på basis av undersøkelsen over, vurdere på to måter. Enten kan man legge hovedvekt på den påviste relative effekt – det er vist en betydelig virkning (33 %), noe som gjør at vi bør behandle. Eller vi kan vurdere det slik: Det er vist at 99 av 100 neppe vil ha noen nytte av behandlingen, og vi avstår derfor fra å behandle. Begrepet «numbers needed to treat» (NNT) er et mål for hvor mange vi gjennomsnittlig må behandle for å oppnå effekt hos én (100 i eksemplet over). Men begrepet er noe abstrakt og gir liten intuitiv mening for mange leger, og betyr i alle fall lite for mange pasienter (2). Som regel er det legens råd som avgjør.

Man kan tenke seg to temmelig forskjellige måter å gi råd på, begge med grunnlag i tallmaterialet over. Den første kunne være: «Det er vist klart at behandlingen hjelper ved denne tilstanden, og jeg råder deg derfor til at vi starter slik behandling.» Alternativt kunne vi si: «Hvis vi skulle starte slik behandling hos deg, måtte du komme hit til kontroll fire ganger per år, du kan risikere bivirkninger underveis, det vil koste deg og helsevesenet tid og penger i årevis, og likevel er det usannsynlig at behandlingen vil være nyttig for deg. Jeg foreslår ingen behandling.»

Små effekter er særlig vanlige når en ny behandling sammenliknes med en som allerede er etablert. Mange nye medikamenter,

ofte langt dyrere enn de gamle, er lansert ut fra mulig marginalt bedre effekt enn tidligere opplegg. Vi er blitt så vant til at mange påviste effekter er små (énprosentmedisinen), slik at vi er blitt for beskjedne i våre krav. Selv med «tiprosentmedisinen», med en absolutt effekt på 10 %, vil vi trolig behandle ni av ti forgjeves. En pasient sa at han ville behandles bare hvis det var mest sannsynlig at behandlingen virket. Han krevde altså over 50 % absolutt effekt, eller $NNT \leq 2$. Kanskje ikke noe dumt syn ved mange tilstander. Det finnes andre mål for effekt i kontrollerte kliniske studier (3). Men det som er angitt over, er dekkende i svært mange situasjoner. Det er hevdet at små effekter påvist i kontrollerte kliniske undersøkelser kan bære bud om større effekt senere, utover den tid som er undersøkt (3). Men det er hypotetisk. Effekten er målt på det tidspunkt man planla for studiestart og er ikke tilfeldig.

Usikkerheter, ulemper og feil

Ramme 1 angir forhold som reduserer verdien av kontrollerte kliniske studier. Inklusions- og eksklusjonskriteriene begrenser deltakerne til bare en undergruppe av dem som har den sykdom man vil undersøke. En annen seleksjonsskjevhet ligger hos undersøkerne. De har ofte tilgang på bare en begrenset subpopulasjon av pasienter, som ikke trenger være representativ. I kon-

Ramme 1

Faktorer som gjør resultatene av kontrollerte kliniske studier usikre eller lite anvendelige i klinisk praksis

- Seleksjon
- Etterlevelse og oppfølging
- Diagnostikk
- Bivirkninger
- Sykeliggjøring
- Kostnader
- Markedsføring
- Studiedesign
- Publikasjonsskjevhet
- Metaanalyser
- Polyfarmasi

trollerte studier sikres riktig medikamentbruk med registrering av bivirkninger og adekvate dosejusteringer. I klinisk praksis er forholdene annerledes. Mange pasienter tar ikke medikamentene som foreskrevet, dosejusteringer avhengig av effekt og bivirkninger er mer tilfeldige, langtidsoppfølgingen kan svikte pga. skifte av lege og annet. Feilaktig eller unøyaktig diagnostikk med lavere presisjon enn i studier medfører dårligere behandlingsresultater.

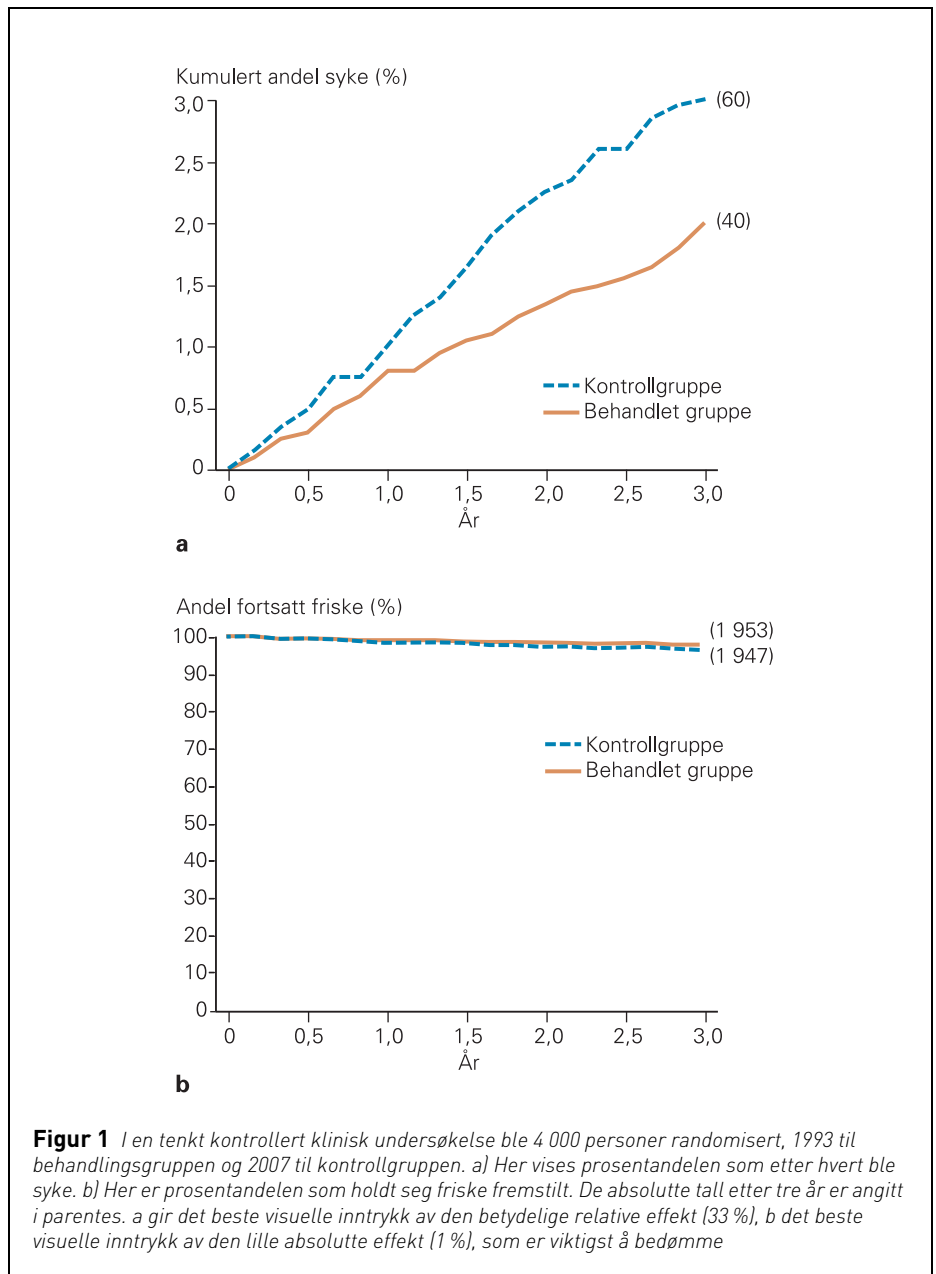
Bivirkninger registreres ikke alltid like nøye som virkningene av en behandling, og får mindre oppmerksomhet ved lansering av resultater. Viktigere er at det ved nye behandlingsmåter alltid er risiko for at bivirkningene kan erkjennes senere (antiarytmika, COX-hemmere, glitazoner, slankemidler, psykofarmaka). I mange undersøkelser påviser man effekter bare på risikofaktorer eller klinisk lite relevante endepunkter (HbA_{1c}, blodtrykk, kolesterolnivå, mild retinopati).

Det å starte behandling, eller bare det å påvise en risikofaktor, medfører ofte sykeliggjøring og kan ha uoversiktlige negative konsekvenser (4). Behandling medfører ofte utgifter for pasient og samfunn langt utover prisen på medikamentene. Den farmasøytiske industri markedsfører behandling tendensiøst (5). Relative effekter presenteres i stedet for absolutte. Små effekter lanseres som viktige. Begrensninger i studiers relevans overses, og resultatene beskrives som mer allmenngyldige enn det er grunnlag for. Markedsføringsstudier uten vitenskapelig betydning lanseres. Markedsføringen retter seg også i økende grad mot grupper som er lettere å overbevise enn ekspertene (leger med forskjellig grad av kompetanse, studenter, sykepleiere, pasientorganisasjoner, mediene og allmennheten).

Inklusjonskriterier kan være laget slik at de prioriterer undergrupper der man antar at behandlingseffektene vil bli størst. Resultater som viser god effekt av behandling, er lettere å få publisert enn de som ikke viser effekt (6). Når resultatene fra flere studier slås sammen i metaanalyser, får de større styrke til å vise små effekter. Men slike analyser er vanskelige og har feilkilder. Kvaliteten av metaanalyser initiert av legemiddelindustrien har vært dårligere enn de som er initiert av uavhengige (7). Mange pasienter har mer enn én sykdom. Polyfarmasi øker faren for interaksjoner og feilmedisinering.

Livsstilsintervensjon

Ved livsstilsintervensjon gjelder mange av de samme betenkelighetene som ved legemidler. Effektene er ofte små, selv ved svær omlegging, som er urealistisk i vanlig klinisk praksis. Kostholdsomlegging kan redusere livskvaliteten for mange, langt fra alle trives med trim, og idrettsskadene må tas inn i regnskapet.



Kort vei ned til null

Usikkerhetene i kontrollerte kliniske studier gjør at mange tilsynelatende påviste behandlingseffekter er falske eller ikke på noen god måte kan omsettes i klinisk praksis. Verst er dette når de påviste absolutte effekter er små, med kort vei ned til null. Behandlingen kan være uten tilsiktet effekt. I statistiske termer kan man si: Konklusjonene om terapi effekt blir usikre, konfidensintervallene blir bredere og omfatter null, signifikansen forsvinner. Terapien kan fort gjøre mer skade enn gagn.

Risikovurdering, utvidet sykdomsbegrep, diagnostikk og screening

Mange små terapi effekter er funnet i grupper av friske personer som har en eller flere risikofaktorer (blodtrykk, lipider, glukose, beintetthet, kroppsvekt, etnisitet,

slektsanamnese). Flere og flere risikofaktorer lanseres og kombineres, friske personer omfattes av utvidede sykdomsbegrep. Avisoppslagene følger: «Mange kan være rammet uten å vite om det.» Risikoangivelser kan være overdrevne. For eksempel viste en studie at 86 % av 40-årige menn i Nord-Trøndelags-undersøkelsen kunne plasseres i høyrisikogruppen for koronarsykdom dersom man fulgte europeiske retningslinjer fra 2003 bokstavelig. Det var ingen i lavrisikogruppen (8). Når dette er resultatet i et land med usedvanlig god folkehelse og høy levealder, kommer risikobegrepet i et underlig lys.

Hvis vi tror at det har effekt å behandle friske folk der det er små avvik i risikofaktorer, fører det til utbredt jakt på slike små avvik. Vi finner frem til personer som vil ha relativt god prognose ubehandlet, og vi kan fort gjøre mer skade enn gagn. Svært mange

må testes for mulig positiv effekt hos få, og falskt positiv screening kan gi angst (9). Det diagnostiseres og behandles omtrent 200 tilfeller av thyreoideakreft per år i Norge. Ut fra obduksjonsmaterialer er det holdpunkter for å anta at 3 000–4 000 hvert år dør med udiagnostisert thyreoideakreft. Disse dør med, men ikke av kreften. Den tiltakende bruk av ultralydscreening av thyreoidea vil avdekke flere av disse uskyldige krefttilfellene og forårsake en overbehandling som bare er skadelig.

De siste tiår har mange sykdommer fått bredere definisjoner. Flere inkluderes, særlig de med mild sykdom og god prognose. Det skyldes dels endring av definisjonene, dels mer avansert diagnostikk og utbredt screening. Referanseområder brukes ofte som definisjon på patologiske tilstander. Men minst 2,5 % av den friske befolkningen ligger per definisjon over referanseområdene slik de vanligvis er definert, og hyppigheten av en sykdom blir derved automatisk minst 2,5 % hvis man velger en slik tilnærming. I noen sammenhenger går man lenger og definerer avvik som noe som øker sykdomsrisikoen utover det ideelle (kolesterolverdier, blodtrykk). Da kan man sette grensen der man vil, og de fleste av oss kan inkluderes. Hos de aller fleste er det risikofaktorer.

Kunnskapsbasert

Det er blitt et utbredt krav at vår medisinske virksomhet skal være kunnskapsbasert. Med det menes ofte kunnskap om resultater fra kontrollerte kliniske studier og metaanalyser. Men begrepet «kunnskapsbasert» burde også inkludere kunnskap om alle begrensningene. Da kan medisinsk praksis bli ganske annerledes. La det være nevnt at

observasjonsstudier, som også brukes som grunnlag, har enda flere feilkilder.

Men vår virksomhet er heldigvis også basert på andre verdifulle grunnpilarer. Den ene er medisinsk basalvitenskap og resonnement, som ofte gir god veiledning, den andre er solid klinisk erfaring og alminnelig menneskelig innsikt og erfaring. Disse har også sine klare begrensninger, men de må ikke vrakes. Vi har fått solid medisinsk statistisk og epidemiologisk ekspertise i bl.a. Nasjonalt kunnskapssenter for helse-tjenesten. Det er nødvendig og bra for å veilede oss i et terreng med fallgruver og irriganger. Men vurderingene mangler korrektiv fra klinikere med bred erfaring. Kontrollerte kliniske undersøkelser likner lite på det virkelige liv.

Ingen skarp grense

Når bør man intervensere, diagnostisk og terapeutisk? Det finnes ingen skarp grense og neppe noen fasit. Skal vi akseptere å belaste mange forgieves for å oppnå effekt hos én? Hvor skeptiske skal vi være til resultatene av kliniske forsøk og til effekten av screening? Synet på dette vil klart og legitimt variere. Sykdommens alvorlighetsgrad, terapeutisk effekt, bivirkninger/kostnader/ulempen og personens alder er viktige faktorer.

Vi er i ferd med å gå altfor langt i å utsette for mange for uvirksom og potensielt skadelig diagnostikk og behandling. Handlingsprogrammer, konsensusrapporter og sjekklistene kan bli en tvangstrøye. Det er etisk vanskelig. Vi må akseptere at det er livsfarlig å leve. Det skal ikke nødvendigvis helsevesenet alltid gjøre noe med.

Man bør i den medisinske grunn- og etterutdanningen gi økt kompetanse i

å vurdere begrensningene i resultater fra kontrollerte kliniske forsøk. Bred klinisk kompetanse bør supplere epidemiologisk og statistisk ekspertise ved uavhengig utarbeiding av retningslinjer og prioriteringer (10). Varsomhet med bruk av screening for risikofaktorer og med inklusjon av friske personer i utredning og behandling er nødvendig.

Litteratur

1. Figueras A, Laporte J-R. Failures of the therapeutic chain as a cause of drug ineffectiveness. *BMJ* 2003; 326: 895–6.
2. Halvorsen PA, Selmer R, Kristiansen IS. Different ways to describe the benefits of risk-reducing treatments. *Ann Intern Med* 2007; 146: 848–56.
3. Wisløff T, Halvorsen PA, Kristiansen IS. Antall som må behandles (NNT) – misvisende, misforstått, misbrukt? *Tidsskr Nor Lægeforen* 2004; 124: 1926–9.
4. Haynes RB, Sackett DL, Taylor DW et al. Increased absenteeism from work after detection and labeling of hypertensive patients. *N Engl J Med* 1978; 299: 741–4.
5. Solhaug HR, Indermo H, Slørdal L et al. Skriftlig legemiddelreklame – til å stole på? *Tidsskr Nor Lægeforen* 2006; 126: 1314–7.
6. De Angelis C, Drazen JM, Frizelle FA et al. Registrering av kliniske forsøk. *Tidsskr Nor Lægeforen* 2004; 124: 2339.
7. Jørgensen AW, Hilden J, Gøtzsche PC. Cochrane reviews compared with industry supported meta-analyses and other meta-analyses of the same drugs: systematic review. *BMJ* 2006; 333: 782–5.
8. Getz L, Sigurdsson JA, Hettlevik I et al. Estimating the high risk group for cardiovascular disease in the Norwegian HUNT 2 population according to the 2003 European guidelines: modelling study. *BMJ* 2005; 331: 551–6.
9. Brewer NT, Salz T, Lillie SE. Systematic review: the long-term effects of false-positive mammograms. *Ann Intern Med* 2007; 146: 502–10.
10. Reikvam Å. Retningslinjer – av hvem og for hvem? *Tidsskr Nor Lægeforen* 2007; 127: 1332.

Manuskriptet ble mottatt 12.9. 2007 og godkjent 26.11. 2007. Medisinsk redaktør Michael Bretthauer.