

Peter C. Gøtzsche

# Det sundhedsvidenskabelige grundlag for medicinsk teknologivurdering

Det Ethiske Råd  
1996

Det sundhedsvidenskabelige grundlag for medicinsk teknologivurdering af Peter C. Gøtzsche, Direktionslæge, dr. med., Det Nordiske Cochrane Center, Rigshospitalet, København  
© Det Ethiske Råd og forfatteren, 1996  
Produceret af Forlaget Thorup  
Forside Anne-Li Engström

ISBN 87-89943-65-1

---

## Indhold

### Forord

### Det sundhedsvidenskabelige grundlag for medicinsk teknologivurdering

Teknologivurdering

Systematisk vurdering af litteraturen

Randomiserede forsøg

Surrogateffektmål

Kirurgi, obstetrik, omsorg og forebyggelsesprogrammer

Ethiske aspekter af randomiserede forsøg

### Diagnostiske undersøgelser

### Organisation af teknologivurdering i Danmark

#### Eksisterende organisationer

Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering

Sundhedsstyrelsens Lægemedjelafdeling

Dansk Sygehus Institut

Center for Helsetjenesteforskning og Socialpolitik

Det Nordiske Cochrane Center

## Forslag til organisation

### Litteratur

---

#### **Forord**

Det Etske Råd forbereder for indeværende en redegørelse om de etiske problemer, der er forbundet med prioritering i sundhedsvæsenet, til udgivelse i slutningen af 1996. I den forbindelse er rådet flere gange stødt på tilkendegivelser om, at det kan være vanskeligt at vurdere effekten af de ydelser, der leveres i sundhedsvæsenet. Det hævdes fra mange sider, at forholdsvis mange af de behandlinger, der tilbydes, ikke er dokumenteret effektive ved den pågældende indikation. Det betyder naturligvis ikke, at størstedelen af disse behandlinger er virkningsløse eller direkte skadelige. Men det betyder, at man for en række behandlinger faktisk ikke har nogen sikkerhed for, om de virker efter hensigten eller ej.

Når dette er relevant i en diskussion af de etiske aspekter af prioritering i sundhedsvæsenet, er det for det første fordi, det må betragtes som uetisk at give folk en behandling, som man ikke har tilstrækkeligt kendskab til virkningen af. Og for det andet fordi man ved at ophøre med virkningsløse eller ligefrem skadelige behandlinger muligvis kan frigøre ressourcer til at udføre virksomme behandlinger.

I den forbindelse bestilte Det Etske Råd i efteråret 1995 et responsum vedrørende det videnskabelige grundlag for vurderingen af effekten af forskellige behandlinger i sundhedsvæsenet hos Direktionslæge Peter C. Gøtzsche, Det Nordiske Cochrane Center, Rigshospitalet. Dette responsum foreligger nu.

Det Etske Råd har fundet, at dette responsum kunne have interesse for en bredere kreds end rådet selv, og har derfor besluttet at udgive det.

De vurderinger og kommentarer, der kommer til udtryk i teksten, er imidlertid ikke nødvendigvis udtryk for holdninger i Det Etske Råd, men er alene forfatterens.

*Lars Nordskov Nielsen, Formand for Det Etske Råd*

*Søren Holm, Formand for Det Etske Råds arbejdsgruppe vedr. prioritering i sundhedsvæsenet*

---

#### **Det sundhedsvidenskabelige grundlag for medicinsk teknologivurdering**

En vigtig forudsætning for, at sundhedsvæsenet kan fungere optimalt og effektivt, er, at der kun benyttes behandlinger med dokumenteret positiv effekt. Dette ideelle mål er svært at opnå. Både antallet af sygdomme og antallet af mulige interventioner er meget stort, hvorfor det vil være uoverkommeligt at undersøge alle interventioner ved alle sygdomme. På den anden side foreligger der for mange interventioners vedkommende en righoldig forskningslitteratur, hvilket i mange tilfælde skulle gøre det muligt at vurdere, hvilke behandlinger, det vil være mest rationelt at anvende.

I praksis indføres nye diagnostiske metoder og behandlinger ofte i sundhedsvæsenet, uden at der er foretaget en systematisk gennemgang og vurdering af forskningslitteraturen. I reglen er der heller ikke foretaget en afvejning af fordele, ulemper og pris, sammenlignet med mulige alternativer. Indførelsen er ofte en glidende proces: Først anvendes teknologien i forsøg, derefter som et tilbud til forsøgsparticipanterne og andre, og til sidst er den umærkeligt blevet rutine. På dette tidspunkt kan det være svært for samfundet at bakke ud igen, uanset hvor stort forholdet mellem prisen og gevinstens størrelse måtte være.

#### **Teknologivurdering**

Teknologivurdering er en hjælp til at overskue konsekvenserne af at anvende forskellige interventioner i sundhedsvæsenet. Teknologivurdering omfatter vurdering af effekt, bivirkninger, omkostninger og afledede virkninger af alle procedurer, der benyttes i sundhedsvæsenet for at forebygge eller behandle sygdomme. Ikke kun apparatur til undersøgelse og behandling, men også lægemidler, operationer, pleje og andre interventioner dækkes således af begrebet (1).

Oversigter over en lang række nye teknologier har vist, at de i gennemsnit ikke har været bedre eller kun marginalt bedre end den teknologi, man har sammenlignet med (2-7). Der er således langt mellem de markante fremskridt, hvorfor man let kommer til at fordyre sundhedsvæsenet uden at opnå noget til gengæld, hvis nye og allerede indførte teknologier ikke undersøges konsekvent og systematisk.

Erfaringen viser, at nogle af de teknologier, man kommer til at indføre, til og med vil være skadelige. Dette er selvsagt uholdbart, da sundhedsvæsenets fornemste opgave må være at sikre borgerne mod skadelige behandlinger. Selv om det også er vigtigt, at nye, gode behandlinger indføres uden unødigt forsinkelse, må undgåelse af skadevolden - *primum non nocere* - dog fortsat anses for at være det vigtigste princip, ikke kun i lægegerningen, men ved alle interventioner overfor borgerne. Derfor vil mange af eksemplerne i det følgende dreje sig om teknologier, der er formodet eller bevist skadelige. Hovedvægten i redegørelsen er lagt på det sundhedsvidenskabelige grundlag for teknologivurderingen. Selve teknologivurderingens teori, metoder og praksis vil ikke blive berørt.

Teknologivurdering er omkostningskrævende. Det er imidlertid langt dyrere at indføre utestede teknologier end at teste dem og filtrere de dårlige fra. Hvis arbejdet prioriteres fornuftigt, vil teknologivurdering kunne spare samfundet for store beløb (8, 9), hvilket Statens beredning for utvørdning af medicinsk metodik i Sverige er et godt eksempel på (10). Hvis man holder op med at bruge ineffektive behandlinger, vil der kunne frigøres ressourcer til at bruge flere både gamle og nye behandlinger med dokumenteret effekt, og til at etablere og vedligeholde systemer til tidlig varsling af nye teknologier (10). I erkendelse af dette har man i flere lande afsat betydelige forskningsmidler til kliniske forsøg, systematisk litteraturvurdering og teknologivurdering.

I Holland betales ny, dyr teknologi kun, hvis det anvendes som led i forsøg (11), og der er afsat ca. 100 mio. kr årligt til at vurdere både ny og eksisterende teknologi. I Storbritannien anbefales samme fremgangsmåde (8, 12), og der er afsat ca. 40 mio. kr årligt til National Health Service Health Technology Assessment Programme (13). Også Tyskland og Frankrig anvender betydelige midler til formålet.

## **Systematisk vurdering af litteraturen**

Kernen i teknologivurderingen er den videnskabelige dokumentation. Det er i reglen et meget stort arbejde at få et pålideligt overblik over forskningslitteraturen. For det første publiceres der over 2 mio. artikler om året. For det andet er elektroniske litteratursøgninger ufuldstændige; ved en god MEDLINE søgning finder man kun omkring halvdelen af de artikler, der er relevante for emnet (14). For det tredje kan søgning i de først indsamlede artiklers referencelister være vildledende, idet forfattere er tilbøjelige til at vælge referencer, der støtter deres eget synspunkt (15, 16). For det fjerde er størstedelen af litteraturen af dårlig kvalitet; det er anslået, at højst 6% af artiklerne - i de gode tidsskrifter vel at mærke - er videnskabeligt sunde, dvs. fører frem til pålidelige og holdbare konklusioner (17). Grunden til dette er først og fremmest, at forskningen ofte er for dårligt tilrettelagt eller er gennemført og analyseret på en måde, der behæfter den med skævheder (bias).

Det blotte omfang af litteraturen bevirker, at man ikke kan følge med, hvis man kun læser originalartikler. Dette gælder også specialisten, som er nødt til at supplere sin viden ved at læse oversigtsartikler. Desværre benytter forskere ikke samme videnskabelige stringens, når de skriver oversigtsartikler, som når de forsker. Det er meget almindeligt, at oversigtsartikler ikke indeholder et metodeafsnit, der klart beskriver, hvordan litteraturindsamlingen blev foretaget, hvordan forfatteren vurderede og vægtede de indsamlede data, og efter hvilke kriterier, forfatteren udelukkede artikler fra oversigten. Kun én ud af 50 oversigtsartikler publiceret i nogle af de mest ansete, internationale tidsskrifter havde således gjort rede for disse afgørende forhold (18).

Usystematiske oversigtsartikler er ofte subjektive, og de kan være temmeligt upålidelige (18-20). I to

oversigtsartikler om behandling af blødende åreknuder i spiserøret, der blev publiceret i British Medical Journal og The Lancet i 1994, anbefalede forfatterne således forskellige behandlinger, fordi de kunne kontrollere blødningen i 90-95% af tilfældene (21, 22). Imidlertid standser blødningen ofte spontant, fx hos 83% af de patienter, der i et forsøg blev behandlet med placebo (virkningsløs medicin)(23). Det er derfor misvisende at beskrive disse behandlinger, som om de er lige så effektive som penicillin ved lungebetændelse, hvilket understreges af, at en systematisk oversigt over de forsøg, hvor et af de anbefalede stoffer var blevet sammenlignet med placebo, ikke kunne påvise nogen effekt af stoffet (24).

Når medicinsk viden indsamles og videregives på en så usystematisk måde, er det måske ikke så overraskende, at eksperter kan være mindre objektive end ikke-eksperter (2, 15, 25, 26), og at oversigtsartikler og eksperters anbefalinger kan være mere præget af, hvilket speciale forfatteren tilhører, end af hvad litteraturen viser. For eksempel anbefalede radioterapeuter strålebehandling efter radikal fjernelse af brystet, selv efter det var påvist, at det var skadeligt, i modsætning til andre lægers anbefalinger (27).

Også velanskrevne lærebøger kan anbefale skadelige eller formodet skadelige behandlinger i klar modstrid med resultater fra kontrollerede forsøg (28). For eksempel anbefaler mange lærebøger og oversigtsartikler rutinemæssig anvendelse af lidokain ved blodprop i hjertet, selv om de forsøg, der er udført med en ubehandlet kontrolgruppe, tyder på, at stoffet er skadeligt (28).

Vilkårligheden i sammenfatning og videregivelse af medicinsk viden har også konsekvenser for indførelse af nye, gavnlige behandlinger (28-30), selv når de er livsreddende (31-33). Indførelsen af streptokinase som rutinebehandling ved blodprop i hjertet blev forsinket i ca. 10 år, fordi ingen havde indsamlet, vurderet og syntetiseret litteraturen (32). En tilsvarende forsinkelse sås for behandling af truende for tidlig fødsel med binyrebarkhormon, der forebygger udvikling af svær sygdom hos barnet (33).

Systematiske oversigter over litteraturen kan undertiden stå alene, uden en efterfølgende teknologivurdering. Det kan fx være indlysende, at en teknologi er så dårligt dokumenteret, at den kun bør anvendes som led i kontrollerede forsøg. Eller vurderingen kan vise, at man kan opnå samme effekt med en billigere teknologi, eller den kan vise, at den nye teknologi er skadelig.

### **Randomiserede forsøg**

Det manglende overblik har også konsekvenser for forskningen. Man har vidst i 20 år, at risikoen for alvorlige infektioner efter kejsersnit nedsættes, hvis man forebygger med antibiotika (34). Alligevel er der i denne periode udført i størrelsesordenen 100 kontrollerede forsøg, hvor man har sammenlignet forskellige antibiotika med placebo, hvilket er uacceptabelt både fra et etisk og et ressourcemæssigt synspunkt. Det burde være et indlysende krav, at der forud for et klinisk forskningsprojekt foretages en systematisk litteraturindsamling og -syntese. Kravet om et grundigt kendskab til den videnskabelige litteratur forud for forsøg fremgår af Helsinki-deklarationen, og er for nylig gjort obligatorisk ved ansøgning om fondsmidler fra det engelske, medicinske forskningsråd.

Der er både teoretiske og empiriske holdepunkter for, at undersøgelse af ny teknologi så vidt muligt bør foretages i et randomiseret forsøg. I et sådant forsøg sammenlignes to eller flere teknologier, og patienterne fordeles til teknologierne efter lodtrækning. Lodtrækningens formål er at sikre, at grupperne ligner hinanden mest muligt, så eventuelle forskelle mellem dem kan tilskrives teknologierne og ikke uvedkommende faktorer. Randomiseringen skaber således en eller flere sammenlignelige kontrolgrupper, som effekten af den nye teknologi kan måles imod.

At randomiseringen er vigtig, illustreres af, at der opstår betydelig bias, hvis den ikke er udført godt nok (en dårlig randomisering er fx en metode, hvor behandleren har mulighed for at afsløre, hvilken behandling den næste patient skal have, før patienten endeligt har accepteret at indgå i forsøget). Studier udført med en suboptimal randomiseringsmetode overvurderer nye teknologiers gavnlige effekt med 30-40% (35). Dette betyder, at man risikerer at medvirke til at indføre skadelige behandlinger, hvis man udfører forsøg, hvor randomiseringen ikke er i orden. Fx kan en overdødelighed på 10-20% på en ny behandling i forhold til standardbehandlingen let maskeres i et dårligt forsøg, så det fejlagtigt

kommer til at vise, at den nye behandling er livsreddende.

Lodtrækningsprincippet strider mod patientens forventning om en individuel stillingtagen, og det er undertiden vanskeligt at gennemføre, især ved alvorlige sygdomme (36). Det har derfor været foreslået at anvende andre forskningsdesign eller logiske ræsonnementer. Den menneskelige organisme er imidlertid yderst kompliceret, og logisk tænkning i praktisk lægegering har vist sig ofte at føre til forkerte behandlinger. reladning har været anvendt i årtusinder i mange forskellige kulturer ved de forskellige sygdomme, også ved kolera hvor et af principperne i den moderne behandling er det modsatte af åreladning, nemlig at tilføre rigelig væske. Diæter og langvarigt sengeleje er nyere eksempler på årtiers anvendelse af sådanne futile behandlinger, endda på trods af dokumentation fra randomiserede forsøg om manglende effekt (37, 38).

For at undgå at konfrontere patienten med det ubehagelige faktum, at man ikke ved, hvad der er bedst og derfor ønsker at trække lod, foretages der undertiden sammenligninger mellem en serie patienter, der har fået en ny behandling, og en serie patienter, der tidligere har fået en anden behandling. Sådanne undersøgelser med en historisk kontrolgruppe kan undertiden give interessante oplysninger om bivirkninger, hvorimod de - med sjældne undtagelser - ikke bør benyttes til at vurdere effekten. Resultaterne er i reglen yderst upålidelige, også selv om man prøver at korrigere statistisk for de skævheder, der skyldes, at de to grupper af patienter ikke er sammenlignelige.

Historisk kontrollerede forsøg giver falske forhåbninger om, at nye behandlinger er bedre end gamle eller bedre end ingen behandling (3, 6, 17, 39-45). Det er ikke usædvanligt, at effekten overvurderes med 50% eller mere. Risikoen for at skadelige behandlinger ser ud, som om de er gavnlige, hvis de undersøges i historisk kontrollerede forsøg, er derfor betydelig.

Der publiceres stadig en hel del historisk kontrollerede forsøg, hvor det havde været muligt at udføre randomiserede forsøg. Offentligheden har derfor krav på beskyttelse mod, at der indføres nye teknologier på dette grundlag. De falske resultater kan endvidere gøre det vanskeligt for andre at gennemføre randomiserede forsøg på grund af etiske betænkeligheder, hvilket er en af den dårlige forsknings alvorligste følgevirkninger.

Forskere kan undertiden være tilbøjelige til at mene, at netop deres behandling er så logisk velbegrunderet eller må forventes at være så effektiv, at det ikke er nødvendigt at udføre randomiserede forsøg. En sådan optimisme er desværre sjældent berettiget, hvilket lægemiddeludviklingen er et eksempel på, idet kun omkring 1 promille af nye substanser ender som et markedsført lægemiddel. Også for andre opdagelser er der et stort frafald undervejs fra laboratorieforsøg til de store, endegyldige randomiserede forsøg.

Da andelen af brugbare innovationer er så lille - og fremskridtene endnu færre - domineres den videnskabelige litteratur af alle de tilsyneladende så lovende meddelelser om ny teknologi (40). Man kommer derfor til at indføre megen ineffektiv og skadelig teknologi, hvis man baserer teknologivurderingen på de foreløbige, lovende meddelelser. Og hvad værre er: Der er en risiko for, at det aldrig bliver opdaget, fordi incitamentet til at udføre et randomiseret forsøg er lille, når en teknologi er accepteret og spredt på markedet.

Industrien er stort set enerådende, når det gælder planlægning og gennemførelse af større forsøg, og de opfylder derfor ikke altid samfundets behov. Fx er langt de fleste forsøg, der er udført med Zofran ved postoperativ kvalme, placebokontrollerede. Dette er ikke hensigtsmæssigt, idet man så ikke ved, om langt billigere præparater, som fx Primperan eller Droperidol, er lige så gode. Et andet eksempel er migrænepræparatet Imigran, der koster over hundrede kr. per tablet. En dansk forskergruppe sammenlignede i et meget stort forsøg Imigran med en kombinationsbehandling bestående af Primperan og et Aspirin-lignende stof og fandt samme effekt af kombinationen, som kun koster en tiendedel af Imigran at behandle med (46).

Som eksemplerne antyder, er det nødvendigt at etablere en stærk, klinisk forskning, der styres af samfundsinteresser og finansieres af samfundet, hvis man skal have en chance for at opnå en sund sundhedsøkonomi.

Det er selvfølgelig ikke alt, der kan eller bør undersøges i randomiserede forsøg. Positive og negative effekter kan være så umiddelbare eller så udtalte, at der ikke er tvivl om dem. Det gælder fx genoplivning ved hjertestop eller tobaksrygning som årsag til kræft, hvilket er overbevisende påvist med andre forskningsdesign. De fleste af de behandlinger, der anvendes i sundhedsvæsenet, har imidlertid langt mindre udtalte effekter, hvorfor randomiserede forsøg som hovedregel altid må anses for at være nødvendige.

Afslutningsvis bør det nævnes, at randomiserede forsøg ikke kan stå alene. Specielt ved behandlinger med lille margin mellem gavnlige og skadelige effekter kan fx observationelle studier af, hvad der sker, når teknologien spredes og anvendes i daglig praksis og på mindre specialiserede afdelinger, være meget værdifulde. Endvidere kan kvalitativ forskning med undersøgelse af patienters præferencer og oplevelser være vigtige, når der skal træffes beslutninger om anvendelse af teknologi.

## **Surrogateffektmål**

Da det ofte er kostbart og tidskrævende at udføre randomiserede forsøg med effektmål, der er relevante for patienterne, prøver man undertiden at nøjes med korttidsforsøg med surrogateffektmål. Et surrogat er associeret til det, man er interesseret i, hvorfor man antager, at en effekt på surrogatet ledsages af en nyttig klinisk effekt.

Denne antagelse holder imidlertid ikke. Fx er blindtarmsbetændelse associeret til feber og smerter, men en effekt af Aspirin på feberen og smerterne vil næppe ledsages af en gunstig effekt på blindtarmsbetændelsen. Snarere tværtimod, idet det kunne medføre, at operationen blev udsat, fordi patienten fik det bedre.

Der er mange eksempler på, at det kan føre til forkerte beslutninger at stole på surrogatmål (16, 42, 43, 47-50). Behandlinger, der har en god effekt på surrogatmålet, kan vise sig at være skadelige for patienterne (42, 47, 49-52). Knoglemineralindholdet er fx et surrogatmål for risikoen for at få fraktur, idet et lavt mineralindhold øger risikoen. Man skulle derfor tro, at præparater, der øger mineralindholdet, nedsætter antallet af frakturer. Fluorid øger imidlertid både mineralindholdet og hyppigheden af frakturer (49), hvilket man måske aldrig ville have opdaget, hvis man kun havde lavet surrogatforsøg.

Behandling af rytmeforstyrrelser i hjertekamrene efter blodprop i hjertet er et andet eksempel på, hvad tro på teori og surrogatmål kan føre til. Det anslås, at behandlingen kostede 50.000 patienter livet om året alene i USA, da den var mest udbredt (49). De antiarytmiske præparater blev godkendt i mange lande ud fra teoretiske overvejelser og forsøg, hvor man havde set en positiv effekt på farlige EKG-forandringer. Senere udførte National Institutes of Health et stort klinisk forsøg, hvor det viste sig, at der døde flere patienter, når de blev behandlet med præparaterne, end når de fik placebo. Forsøg med andre anti-arytmiske stoffer har også vist tendens til skadelig effekt (28, 54). Flere af disse præparater anvendes fortsat i Danmark. Lidokain og mexiletin er godkendt til forebyggelse af arytmier. Flekainid er forbeholdt livstruende rytmeforstyrrelser, men selv med dette forbehold er det betænkeligt, at præparatet fortsat er på markedet, da det er livstruende i sig selv. Det er blevet hævdet, men der er ikke overbevisende dokumentation for, at balancen mellem positive og skadelige effekter skulle være gunstig blandt de mest syge patienter (55).

Langvarigt forhøjet blodsukker disponerer til nerve-, nyre- og øjenskader. Det er derfor nærliggende at tro, at nedsættelse af blodsukkeret hos sukkersygepatienter altid er gavnlig. Det er imidlertid ikke sikkert. Tolbutamid benyttes ofte ved behandling af gammelmandssukkersyge i Danmark. Det eneste store forsøg, der nogensinde er lavet med dette stof, blev imidlertid stoppet i utide på grund af en tendens til overdødelighed blandt dem, der fik tolbutamid (27). Det gik oven i købet de patienter dårligst, der var flinkest til at tage medicinen (disse patienter plejer ellers at have den bedste prognose, selv når de indtager placebo (56). Dertil kommer, at effekten på blodsukkeret er beskeden og svinder efter få år. Stoffet burde derfor kun anvendes som led i forsøg.

Man kan diskutere, om den nuværende praksis for godkendelse af lægemidler er rimelig, idet forsøg med surrogatmål anses for at være tilstrækkelig dokumentation for effekten af et lægemiddel. Der er mere og mere, der tyder på, at behandling af forhøjet blodtryk med calciumantagonister er dårligere

end behandling med betablokkere eller med vanddrivende midler, og måske er behandlingen skadelig, i det mindste for nogle af præparaternes vedkommende (57).

## **Kirurgi, obstetrik, omsorg og forebyggelsesprogrammer**

Kirurgisk og obstetrisk teknologi indebærer større indgreb i den menneskelige organisme, og dermed større risiko for skadevirkninger, end mange lægemidler. Det er derfor paradoksalt, at samfundet ikke stiller krav om, at disse teknologier skal vurderes i randomiserede forsøg, før de frigives til almindelig anvendelse. Der kan være specielle problemer ved at udføre randomiserede forsøg inden for kirurgien, men de kan løses (58).

Kirurgiens og obstetrikens historie har mange eksempler på indførelse og senere opgivelse af skadelige behandlinger, fx rutinemæssig fjernelse af mandler, hvilket gik ud over en stor del af befolkningen. Kikkertkirurgien (laparoskopisk kirurgi) er et frisk eksempel på manglende vurdering af operationer inden generel udbredelse. Teknologien er efterfølgende blevet dokumenteret i randomiserede forsøg og har vist sig nyttig ved nogle sygdomme, men da der er tvivl om dens berettigelse ved andre, ønsker man i Storbritannien at undersøge den mere systematisk, fx ved gynækologiske operationer (13).

Randomiserede forsøg bør integreres i forskning og udvikling overalt, hvor det er muligt (59). Dette gælder også tilsyneladende banale ting som nye suturmateriale. Et nyt suturmateriale gav hyppigere smerter ved samlejet 3 år efter fødslen (60). Sådanne bivirkninger er oftest umulige at opdage uden randomiserede forsøg.

Der er mange uafklarede problemer inden for sygepleje, fysio- og ergoterapi om valg af den bedste behandling. Softlaser-behandling af smerter i bevægeapparatet er derimod et eksempel på en teknologi, hvis effekt er afklaret. Behandlingen er meget udbredt, selv om de teoretiske holdepunkter for, at den skulle virke, er dårligt funderede, og selv om hverken en statistisk sammenfatning (metaanalyse) af de placebokontrollerede forsøg (61) eller senere udførte forsøg har kunnet påvise nogen effekt af behandlingen. Da et laserapparat koster omkring 20.000 kr. i anskaffelse, og behandlingen indebærer gentagne besøg hos terapeuten, er det ikke helt ligegyldigt, hvor længe det varer, inden behandlingen afskaffes.

Omsorgsforskere angriber undertiden randomiserede forsøg med argumenterne, at de ikke tager hensyn til den samlede effekt af behandlingerne, og at de ikke kan bruges til at vurdere de kvalitative eller menneskelige aspekter, der fx kan være vigtige i sygeplejen. Kvalitativ forskning er selvfølgelig vigtig i sin egen ret, men kritikken er alligevel ikke berettiget. Fx har randomiserede forsøg påvist klar effekt af familiepleje ved skizofreni (62). De har også påvist dramatiske effekter af at have en støtteperson til den fødende kvinde, der er til stede i hele fødselsforløbet, sammenlignet med de sædvanlige forhold på en fødeafdeling. Anvendelsen af kejsersnit, bedøvelse, episiotomi og smertestillende medicin nedsættes, barnets tilstand ved fødslen bliver bedre, og der er færre tilfælde af svær fødselsdepression (63).

Føtal kardiokografi (elektronisk overvågning af fosteret) er en diagnostisk test, der er kontroversiel. En metaanalyse af de kontrollerede forsøg viste, at der døde flere nyfødte, når mødre blev overvåget, end når de ikke blev det (64). Det drejede sig imidlertid om små tal og i nogle tilfælde om misdannede børn, hvorfor resultatet skal tolkes med varsomhed. En medical audit af intrauterin og neonatal dødelighed i København giver derimod et fingerpeg om, hvad der kan gå galt, hvis man stoler for meget på apparaturet og for lidt på patienten og den kliniske vurdering. I to tilfælde, hvor kardiokografi blev benyttet, kunne akut kejsersnit være blevet udført tidligere, hvis obstetrikeren ikke havde brugt tid på yderligere tekniske undersøgelser med enten kardiokografi eller andet apparatur, efter der allerede forelå klare faresignaler (65). I en redegørelse til Folketinget om medicinsk teknologivurdering blev en række problemer ved kardiokografi påpeget allerede i 1980 (1), hvilket imidlertid ikke har forhindret den i at blive spredt over hele landet. Anvendelse af kardiokografi har ført til store ændringer i det psykologiske miljø ved fødslen, hvilket kan være u hensigtsmæssigt, idet psykologiske faktorer er meget vigtige for et optimalt fødselsforløb. Det er derfor spørgsmålet, om anvendelsen af kardiokografi med fordel kunne mindskes og erstattes af øget klinisk vurdering og øget omsorg.

Der anvendes meget store beløb på medicinsk behandling af svært syge kræftpatienter, også når de er kommet i en fase, hvor det ikke længere er realistisk at tro på, at behandlingen hjælper, eller at en eventuel livsforlængelse står mål med de lidelser, behandlingen påfører patienterne. Det er vist i en finsk undersøgelse, at den mest aggressive behandling af brystkræft fandt sted, når patienterne var i slutfasen (66). Det er tankevækkende, at en nylig, omfattende meta-analyse af 52 forsøg ved lungekræft udelukkende beskrev overlevelsen (der i den gunstigste gruppe kun blev forlænget med ca. 2 måneder) og ikke patienternes livskvalitet eller bivirkningerne ved behandlingerne (67). Også på hæmatologiske afdelinger benyttes store ressourcer efter patienterne har haft flere tilbagefald. Selv om patienterne ofte ønsker behandling også i håbløse situationer, var det måske alligevel rimeligt i større udstrækning end det finder sted i dag at prøve at frigøre nogle af ressourcerne til omsorg for patienterne selv og til andre formål i sundhedsvæsenet.

Forebyggelsesprogrammer bør også randomiseres (68). Der er for få eksperimenter og for mange beskrivende undersøgelser i Danmark inden for helsetjenesteforskning og epidemiologi (69). Dette gælder formentlig i endnu højere grad inden for socialvæsenet, hvor interventioners effekter ofte er ukendte, trods de meget store summer, der ofres på dem (fx afvænningsprogrammer for narkomaner). Lige så vel som det er påvist inden for sundhedsvæsenet, må man inden for socialvæsenet regne med, at velmente interventioner kan vise sig at være skadelige. Men det opdager man selvfølgelig ikke, hvis man ikke undersøger dem, hvilket man ofte ikke vil føle sig forpligtet til at gøre, da socialvæsenet ikke har samme tradition som sundhedsvæsenet for at undersøge sine interventioner.

### **Etiske aspekter af randomiserede forsøg**

Kravet om fuld information i forsøg kan af og til være barsk for både patient og læge, fx når der meddeles en cancerdiagnose. Det kan også være så vanskeligt at opfylde, at proceduren kan blive farcepræget, dvs. formelt opfyldt men uden reelt indhold, fx hvis patienten er meget bange, og der skal handles hurtigt.

Læger er frit stillet til at prøve nye behandlinger, der ikke er officielt godkendt, på enkelte patienter, hvis de skønner, at patienterne måske kunne have gavn af dem. Det er uundgåeligt, at resultaterne af sådanne "forsøg" kommer til at påvirke lægers opfattelse af behandlingerne, selv om det er veldokumenteret, at viden indhentet på denne måde er meget upålidelig.

Samfundet burde interessere sig mindst lige så meget for det etiske problem i denne lægelige vilkårlighed, som for forsøg hvor usikkerheden er sat i system med henblik på at opnå ny viden (70). En læge udtrykte paradokset således: "Jeg skal have tilladelse til at bruge et lægemiddel til halvdelen af mine patienter, men ikke hvis de får det allesammen" (71)! Den store, franske fysiolog Claude Bernard skrev allerede i 1866: Sammenlignende erfaring er en absolut betingelse for den eksperimentelle og videnskabelige medicin, ellers vandrer lægen på må og få og bliver offer for tusinde illusioner (72). Uvidenhed ledsages af store omkostninger, men uvidenheden er ofte usynlig. Derfor kan det være mere attraktivt for alle parter at opretholde uvidenheden end at stille den til skue, som man nødvendigvis må gøre i et randomiseret forsøg.

Patienter bør ikke kunne forlange, og læger bør ikke give, utilstrækkeligt gennemprøvede behandlinger uden for protokol. Sådanne usystematiske "forsøg" på enkeltpatienter kan medvirke til, at potentielt skadelige behandlinger indføres "ad bagvejen" (8).

Deltagelse i gode, velbegrundede forsøg bør være en naturlig del af en implicit kontrakt mellem borgerne og sundhedsvæsenet om parternes rettigheder og forpligtelser, idet borgerne får glæde af de resultater, der er opnået ved tidligere forsøg. Hvis patienterne sætter sig uden for fællesskabet og forlanger at få nye behandlinger uden for protokol, fordi de håber, det vil hjælpe dem, skaber de et dilemma for deres medpatienter. Hvis handlemåden blev generel, ville den nemlig føre til, at alle blev stillet ringere i al fremtid.

Patientens ret til at nægte deltagelse i et forsøg skal selvfølgelig respekteres. Det er imidlertid lige så vigtigt, at samfundet forbeholder sig ret til at kun tillade anvendelse af forsøgsbehandlinger som led i randomiserede forsøg, idet samfundets motiv er at sikre en generel beskyttelse af borgerne mod, at der indføres skadelige behandlinger. Patientens ret til selvbestemmelse bør derfor ikke opfattes som en



egoistisk ret til at kræve uafprøvede behandlinger uden for protokol.

Som tidligere nævnt er det besluttet i Holland, at nye behandlinger kun må anvendes på visse sygehuse, og kun som led i forsøg. Først når afprøvningerne er afsluttede, tages der stilling til, om behandlingerne skal frigives til bredere anvendelse. Hvis dette princip blev fulgt i Danmark, ville man fritage lægerne for den nuværende, uholdbare situation, hvor de selv skal afgøre, om en patient, der ikke vil deltage i et forsøg og kræver at få forsøgsbehandlingen, også skal have lov til det. Da såvel patienter som læger ofte har urealistiske forventninger om nye behandlings gavnlige effekter, kan det selvsagt være meget svært at afslå sådan en anmodning.

---

## Diagnostiske undersøgelser

Det er ikke almindeligt anerkendt, at diagnostiske procedurer bør afprøves i randomiserede forsøg. Det ville imidlertid være meget relevant, idet mange procedurer både er meget ressourcekrævende og kan indebære en vis risiko for patienten, fx ved invasive undersøgelser.

Mange klinisk-fysiologiske undersøgelser er blevet indført, uden det er dokumenteret, at de hjælper patienten. Det gælder fx scintigrafisk undersøgelse af patienter med blodprop i hjertet, hvor man måler, hvor stor en del af hjertets volumen, der pumpes ud ved hvert hjerteslag. Hvis hjertets pumpeevne er nedsat, har patienten en dårligere prognose. Nedsat pumpeevne kan behandles med medicin, hvilket har effekt på overlevelsen, men det er ikke sikkert, at scintigrafien er nødvendig for at starte medicinen. I et randomiseret forsøg gjorde det således ingen forskel, om der forelå en scintigrafi eller ej (73). Det er heller ikke sikkert, det er nødvendigt at udføre en ekkokardiografi, hvilket er en anden hyppigt benyttet undersøgelse. Man kan undersøge pumpeevnen på andre måder, fx ved at udspørge om symptomer på hjertesvigt og undersøge patienten klinisk på sædvanlig måde, eventuelt suppleret med røntgen og EKG, hvilket er langt billigere end scintigrafi og ekkokardiografi.

Behovet for teknologivurdering og for effektmål, der er relevante for patienten, sættes især i relief af de nye billeddannende teknikker, fx MR- og PET-scanning, der er særdeles kostbare at anvende.

---

## Organisation af teknologivurdering i Danmark

Sundhedsstyrelsens forslag om en National strategi for medicinsk teknologivurdering (74) og rapporten Fremtidens medicinske metoder - tidlig opmærksomhed og vurdering (10), der begge publiceres i 1996, indeholder forslag om organisation af teknologivurdering i Danmark og etablering af et system til tidlig varsling. Rapporterne beskriver også organisationer i andre lande, herunder især National Health Service i Storbritannien. Betragtningerne i det følgende er udarbejdet uafhængigt af disse rapporter og er i høj grad inspireret af de strategier og metoder, der benyttes af National Health Service.

I Danmark eksisterer der ikke nogen enkelt organisation, der har til opgave at afveje nytte og omkostninger ved anvendelse af gamle såvel som nye teknologier. Interferon til behandling af dissemineret sklerose er et aktuelt eksempel på de akutte problemer, dette kan medføre. Stoffet er ved at blive godkendt af Sundhedsstyrelsen på baggrund af et enkelt randomiseret forsøg, hvor der var effekt på sygdomsaktiviteten. Der er imidlertid ikke dokumentation for, at behandlingen er helbredende, og da det vil koste 100 mio. kr om året at anvende interferon til en kernegruppe bestående af blot de 1.000 mest relevante patienter, er en teknologivurdering påkrævet, før stoffet eventuelt tages i bredere anvendelse. Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering har derfor nedsat et hurtigtarbejdende udvalg vedrørende interferonbehandling af sklerose.

Der mangler også et tværfagligt organ til overordnet prioritering, herunder vurdering af, om man vil betale en given pris for en given nytte.

Formålet med at indføre teknologivurdering i større målestok end de spredte vurderinger, der foretages i øjeblikket, ville være at opnå størst mulig gavn for patienterne af de interventioner, der benyttes i sundhedsvæsenet. Teknologivurderingen bør derfor føre til bedre diagnostik og behandling og/eller til

mere rationel anvendelse af ressourcerne.

## **Eksisterende organisationer**

### **Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering**

Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering har til opgave at medvirke til en forstærkelse af indsatsen for anvendelse af medicinsk teknologivurdering i sundhedsvæsenet (74). Udvalget skal udarbejde og bistå med implementeringen af en national strategi for medicinsk teknologivurdering i sundhedsvæsenet. Udvalget er således tildelt en overordnet, strategisk rolle, men det ville være naturligt, hvis udvalget også fik en central rolle ved den mere konkrete organisation og koordination af teknologivurderingen i Danmark. I overensstemmelse hermed indgår det i udvalgets kommissorium at tage initiativ til, finansiere, vurdere og give bidrag til udarbejdelse og implementering af konkrete projekter. Det er imidlertid ikke meningen, at udvalget selv skal gennemføre teknologivurderinger.

### **Sundhedsstyrelsens Lægemiddelafdeling**

Sundhedsstyrelsens Lægemiddelafdeling har ansvar for godkendelse af nye lægemidler. En vigtig opgave er derfor at bedømme den videnskabelige litteratur i registreringsansøgninger. Lægemidlet skal være formålstjenligt i sundhedsmæssig henseende, men behøver ikke være bedre end eksisterende alternativer. Afdelingen har ikke ansvar for at vurdere, om lægemidlets nytte er rimelig i forhold til de omkostninger for samfundet, godkendelsen vil medføre.

I 1998 træder en bekendtgørelse om medicinsk udstyr i kraft, men Sundhedsstyrelsen har allerede indskærpet, at der ved ibrugtagning af nyt medicinsk udstyr ikke må indføres et nyt behandlingsprincip, før der er gennemført en klinisk undersøgelse.

### **Dansk Sygehus Institut**

Dansk Sygehus Institut har siden 1980 gennemført konkrete teknologivurderinger og har siden 1983 i samarbejde med Statens Sundhedsvidenskabelige Forskningsråd gennemført konsensus-konferencer. Instituttet har lægefaglig, økonomisk, sociologisk og medikoteknisk ekspertise inden for området vurdering og implementering af medicinske teknologier. Det vil øge sin indsats inden for området i de kommende år samt yde rådgivning på konsulentbasis. Instituttet har et udbygget internationalt netværk til andre forskere og koordinerer og deltager i internationale forskningsprojekter. Instituttet er medlem af International Network of Agencies for Health Technology Assessment (INATHA). Det samarbejder med Det Nordiske Cochrane Center og deltager i Cochrane Review Group on Effective Professional Practice. Instituttet råder over et offentligt bibliotek og udgiver litteraturoversigter over teknologivurderinger.

### **Center for Helsetjenesteforskning og Socialpolitik**

Center for Helsetjenesteforskning og Socialpolitik ved Odense Universitet er et tværvidenskabeligt center for sundheds- og samfundsvidenskab. Centeret har en særlig interesse i forskning i sundhedsøkonomi, hvilket er en vigtig del af den medicinske teknologivurdering. Endvidere forskes der i metoder til tidlig varsling af nye medicinske teknologier, herunder hvordan teknologierne opstår og udvikles. Udover forskning i metodeudvikling har centeret gennemført flere konkrete sundhedsøkonomiske analyser, fx omkostningsanalyse af interferonbehandling af hepatitis C, samfundsøkonomisk analyse af anæstesi ved total hoftealloplastik og økonomisk analyse af en ny test for colorectal cancer.

### **Det Nordiske Cochrane Center**

Det Nordiske Cochrane Center indgår i The Cochrane Collaboration, som er en forskningsorganisation, hvis formål er at udarbejde systematiske oversigter over interventioners effekter, opdatere dem når nye studier publiceres, og gøre oversigterne let tilgængelige for sundhedsvæsenet. Organisationen udgiver to databaser, The Cochrane Pregnancy and Childbirth Database (75) og The Cochrane Database of

Systematic Reviews (76). Sidstnævnte inkluderer en database med over tusinde referencer til metaanalyser og rapporter over teknologivurderinger. Cochrane-samarbejdet udvikles meget hurtigt og har allerede tiltrukket en stor del af den internationale ekspertise, der beskæftiger sig med kritisk vurdering af forskningslitteraturen efter videnskabelige principper. Organisationen har et tæt samarbejde med organisationer, der arbejder med teknologivurdering. Dette har bl.a. til formål at sikre, at der ikke flere steder i verden udføres systematiske oversigter eller teknologivurderinger af samme problemstilling.

## **Dansk Institut for Klinisk Epidemiologi**

Instituttet spiller en vigtig rolle for den medicinske teknologivurdering, idet det bl.a. beskæftiger sig med at kortlægge sygdommes epidemiologi og vurdere effekten af forebyggelsesprogrammer.

---

## **Forslag til organisation**

Der bør i Danmark skabes en stærk organisation omkring medicinsk teknologivurdering, med solid, forskningsmæssig forankring. Som anført i Sundhedsstyrelsens publikation om teknologivurdering bør der anvendes samme systematik og statistiske metoder som i originalforskning (77). De eksisterende forskningsmiljøer bør endvidere styrkes, da der er mange uafklarede spørgsmål om de mest valide metoder til litteratur- og teknologivurdering.

Arbejdet er meget ressourcekrævende, og der bør derfor afsættes rigelige midler til det, hvis det skal have nogen gennemslagskraft som bolværk mod de mange nye, meget kostbare teknologier og mod indførelse af teknologier, der er udokumenterede, og derfor potentielt ineffektive eller skadelige.

Ideelt set burde alle nye teknologier evalueres systematisk, før de eventuelt tages i bred anvendelse, uanset deres natur. Dette ambitiøse mål vil dog nok næppe nogensinde blive opfyldt. Det er derfor mest rationelt at satse på interventioner, hvor der er stor chance for væsentlige forbedringer i patientbehandlingen og/eller besparelser.

Ved udvælgelse af områder for teknologivurdering kunne der lægges vægt på:

- teknologier der er meget anvendt og berører en stor del af befolkningen;
- teknologier der kunne indebære betydelige fremskridt i diagnostik eller behandling;
- teknologier der er meget kostbare;
- teknologier hvor der er mistanke om skadelig eller manglende effekt;
- teknologier hvor det forventes, at der kan opnås betydelige besparelser med uændret effekt.

Arbejdsopgaverne kunne være:

- identificering og prioritering af teknologier, der trænger til afklaring;
- identificering af teknologier hvis udbredelse bør kontrolleres, indtil der foreligger yderligere information;
- tidlig identificering af teknologier der forventes at få store konsekvenser for sundhedsvæsenet;
- identificering og prioritering af behov for metodeudvikling inden for teknologivurdering;
- koordination, også med internationale organisationer, så unødigt dobbeltarbejde undgås;
- kritisk, systematisk vurdering af forskningslitteraturen;
- konkrete teknologivurderinger; den systematiske litteraturvurdering udgør det videnskabelige

grundlag herfor;

- vurdering af ansøgninger om midler til kliniske forsøg, systematiske litteraturvurderinger og teknologivurderinger;
- implementering af forskningsresultater og opfølgning af at teknologien anvendes efter hensigten.

I lighed med National Health Service i Storbritannien, der fra 1991 er blevet en aktiv forskningspartner via dets Forsknings- og udviklingsprogram (commissioned research) (12, 78), burde der også i Danmark etableres en central fond til teknologivurdering.

Princippet for udvælgelse af prioriterede områder samt indhentning og prioritering af forslag kunne ligge tæt op ad retningslinjerne i Storbritannien, hvor midlerne gives til tre hovedformål (13):

- studier der involverer indsamling og analyse af nye data;
- systematiske oversigter over forskningslitteraturen;
- teknologivurderingsprojekter.

Forslag til teknologier, der trænger til vurdering, kunne indhentes ved bred annoncering i fag- og dagspressen, suppleret med henvendelse til de videnskabelige og faglige selskaber, patientforeningerne og sygehusejerne.

Udvælgelse af prioriterede områder kunne foretages af Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering, eventuelt efter supplerende med flere videnskabelige medlemmer, eller af en særlig komité oprettet til formålet (13). Det ville være hensigtsmæssigt at oprette et antal underudvalg med overvejende videnskabelig repræsentation. Et af udvalgene bør fokusere på videnskabelig metodologi (13); en mulighed vil være, ligesom ved universitetet i York i England, at lade dette underudvalg få hjemsted i et tværfagligt forskningsmiljø med styrke i metodeudvikling, fx Center for Helsetjenesteforskning og Socialpolitik.

Forslag til konkrete projekter på baggrund af prioriterede områder kunne indhentes i en tottrinsproces: Først skitseagtige forslag efter et standardskema, dernæst fuldstændige protokoller blandt de først udvalgte forslag. Forskerne bør angive i protokollen, hvordan resultaterne skal komme i anvendelse, fx gennem samarbejde med de videnskabelige selskaber (13). Der bør være fri konkurrence for at opnå et så højt niveau som muligt for projekterne.

Forslag bør indeholde en detaljeret tidsplan for arbejdet med konkrete, evaluerbare milepæle, og der bør afleveres en halvårsrapport til udvalget. Der bør endvidere være regelmæssige kontakter mellem udvalget og forskerne for at diskutere problemer og vurdere, om projektet skrider frem i overensstemmelse med planen.

Forslag bør også indeholde en plan for implementering og opfølgning, så det sikres, at det, der er evidens for, føres ud i livet og anvendes i overensstemmelse med dokumentationen. Dansk Sygehus Institut og Dansk Institut for Klinisk Epidemiologi kunne få en ledende rolle ved den opfølgende vurdering af implementeringen.

Det kunne foreslås at lade Statens Sundhedsvidenskabelige Forskningsråd foretage den videnskabelige vurdering af ansøgningerne, idet en vurdering af, om forbedringer eller besparelser vil være realistiske, ofte kræver kendskab til fagområdet. Ved vurderingerne bør anvendes mindst to videnskabelige konsulenter (peer reviewers). Der bør lægges vægt på videnskabelig værdi, herunder design og metodologi, og relevans for sundhedsvæsenet; endvidere bør omkostninger, varighed og muligheder for gennemførelse vurderes (13). Den endelige prioritering af ansøgningerne kunne eventuelt forbeholdes et råd, der både har videnskabelige og politiske medlemmer, inklusive repræsentanter for sygehusejerne, samt lægpersoner fx fra patientforeningerne, for at sikre, at forskningen og teknologivurderingen bliver så patient- og samfundsorienteret som muligt. Det bør sikres, at sygehusejerne løbende informeres og rådgives om ikke at indføre de teknologier, der er under evaluering (13).

I Storbritannien opererer man med et provider-purchaser forhold, som betyder, at de, der skal betale for interventionerne, vælger mellem de mulige tilbud. I Danmark kunne de enkelte sygehusmyndigheder organisere sig i et fælles forum, hvor man på baggrund af vurderingerne og i samråd med relevante specialister prøvede at opnå enighed om, hvilke teknologier man vil købe.

For at undgå dobbeltforskning bør forslag til prioriterede projekter konfereres med Det Sundhedsvidenskabelige Forskningsråd, Det Nordiske Cochrane Center, Dansk Sygehus Institut og National Health Service i Storbritannien, før de vedtages (13). For systematiske litteraturvurderinger eksisterer der en database, der indeholder oplysninger om planlagte og påbegyndte projekter (76). Det Nordiske Cochrane Center kunne få til opgave at undersøge mulighederne for at få udarbejdet prioriterede systematiske litteraturoversigter inden for rammerne af The Cochrane Collaboration.

Der er i Storbritannien taget initiativ til at oprette et katalog over eksisterende interventioner i sundhedsvæsenet og over igangværende teknologivurderinger; også Statens beredning för utvärdering av medicinsk metodik har oprettet en database over løbende projekter. Det er vigtigt, at der - ligesom flere steder i udlandet - tages initiativ til at etablere en database over påbegyndte kliniske forsøg for at undgå utilsigtet dobbeltforskning. Her kunne det videnskabsetiske komitéssystem spille en førende rolle, da overflødig forskning er dårlig forskning, og da det er de videnskabsetiske komitéers aksiom, at al dårlig forskning er uetisk (79).

---

## Litteratur

1. Andreasen, P.B. 1980. Medicinsk teknologivurdering. Rapport til Folketingets udvalg angående videnskabelig forskning. København.
2. Gøtzsche, P.C. 1990. Bias in double-blind trials [disputats]. *Dan Med Bull* 37: 329-36.
3. Berlin, J.A., C.B. Begg and T.A. Louis. 1989. An assessment of publication bias using a sample of published clinical trials. *Journal of the American Statistical Association* 84: 381-392.
4. Gilbert J.P, B. McPeck, F. Mosteller. 1977. Progress in surgery and anesthesia: benefits and risks of innovative therapy. In *Costs, risks and benefits of surgery: 124-69*. Ed. J.P.Bunker. New York: Oxford University Press.
5. Colditz, G.A.J.N. Miller and F. Mosteller. 1989. How study design affects outcomes in comparisons of therapy. I: medical. *Statistics in Medicine*; 8: 441-54.
6. Miller, J.N, G.A. Colditz and F. Mosteller. 1989. How study design affects outcomes in comparisons of therapy. II: surgical. *Statistics in Medicine*; 8: 455-68.
7. Bailar III, J.C and E.M. Smith. 1986. Progress against cancer? *New England Journal of Medicine*; 314: 1226-32.
8. Advisory Group on Health Technology Assessment for the Director of Research and Development. 1992. *Assessing the effects of health technologies*. London: Department of Health.
9. Ray, W.A, M.R. Griffin and J. Avorn. 1993. Sounding board: evaluating drugs after their approval for clinical use. *New England Journal of Medicine*; 329: 2029-32.
10. Poulsen, P.B, M. Hørder og T. Jørgensen. 1996. *Fremtidens medicinske metoder - tidlig opmærksomhed og vurdering*. København: Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering (i trykken).
11. Borst-Eilers, E. 1993. Assessing hospital technology in the Netherlands: new treatments are paid for only if they are part of an evaluation [editorial]. *British Medical Journal*; 306: 226.
12. Department of Health. 1993. *Research for health*. London: Department of Health.
13. Report of the NHS Technology Assessment Programme. 1995. Leeds: Department of Health.
14. Dickersin, K., R. Scherer and C. Lefebvre. 1995. Identifying relevant studies for systematic reviews. In *Systematic reviews: 17-36*. Eds. I.Chalmers and D.G.Altman. London: British Medical Journal Publishing Group.
15. Gøtzsche, P.C. 1987. Reference bias in reports of drug trials. *British Medical Journal*: 295: 654-6.
16. Ravnskov, U. 1992. Cholesterol lowering trials in coronary heart disease: frequency of citation and outcome. *British Medical Journal* 305: 15-9.
17. Williamson, J.W, P.G. Goldschmidt and T. Colton. 1986. The quality of medical literature: an analysis of validation assessments. In *Medical uses of statistics*. Eds. J.C.Bailar III and F. Mosteller. Waltham: NEJM Books.

18. Mulrow, C.D. 1987. The medical review article: state of the science. *Annals of Internal Medicine* 104: 485-8.
19. Squires, B.P. 1989. Biomedical review articles: What editors want from authors and peer reviewers. *Canadian Medical Association Journal* 141: 195-7.
20. Oxman, A.D, G.H. Guyatt, J. Singer, C.H. Goldsmith, B.G. Hutchison BG, R.A. Milner et al. 1991. Agreement among reviewers of review articles. *Journal of Clinical Epidemiology* 44: 91-8.
21. Williams, S.G.J and D. Westaby. 1994. Management of variceal haemorrhage. *British Medical Journal* 308: 1213-7.
22. Bornman, P.C, J.E.J. Krige and J. Terblanche. 1994. Management of oesophageal varices. *Lancet* 343: 1079-84.
23. Valenzuela, J.E, T. Schubert, M.R. Fogel, R.M. Strong RM, J. Levine, P.R. Mills et al. 1989. A multicenter, randomized, double-blind trial of somatostatin in the management of acute hemorrhage from esophageal varices. *Hepatology* 10: 958-61.
24. Gøtzsche, P.C, I. Gjørup, H. Bonnén, N.E.B. Brahe, U. Becker and F. Burcharth. 1995. Somatostatin v placebo in bleeding oesophageal varices: randomised trial and meta-analysis. *British Medical Journal* 310: 1495-8.
25. Oxman, A.D. and G.H. Guyatt. 1993. The science of reviewing research. In *Doing more good than harm*. Eds. K. S Warren and F. Mosteller. *Annals of the New York Academy of Science* 703: 125-33.
26. Cooper, H. and R.G. Ribble. 1989. Influences on the outcome of literature searches for integrative research reviews. *Knowledge* 10: 179-201.
27. Chalmers, T.C, C.S. Frank CS and D. Reitman. 1990. Minimizing the three stages of publication bias. *Journal of the American Medical Association* 263: 1392-5.
28. Antman, E.M, J. Lau, B. Kupelnick, F. Mosteller and T.C. Chalmers. 1992. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *Journal of the American Medical Association* 268: 240-8.
29. Stross, J.K and W.R. Harlan. 1979. The dissemination of new medical information. *Journal of the American Medical Association* 241: 2622-4.
30. Lomas, J. and R.B. Haynes. 1988. A taxonomy and critical review of tested strategies for the application of clinical practice recommendations: from 'official' to 'individual' clinical policy. *American Journal of Preventive Medicine* 4(suppl 2): 77-94.
31. Baum, M.L, D.S. Anish, T.C. Chalmers, H.S. Sacks, H. Smith and R.M. Fagerstrom. 1981. A survey of clinical trials of antibiotic prophylaxis in colon surgery: evidence against further use of no-treatment controls. *New England Journal of Medicine* 305: 795-9.
32. Lau, J., E.M. Antman, J. Jimenez-Silva, B. Kupelnick, F. Mosteller and T.C. Chalmers. 1992. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *New England Journal of Medicine* 327: 248-54.
33. Crowley, P., I. Chalmers and M.J.N.C. Keirse. 1990. The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. *British Journal of Obstetrics and Gynaecology* 97: 11-25.
34. Smaill, F. [revised 03 August 1994]. Prophylactic antibiotics in Caesarean section (all trials). In: *Pregnancy and Childbirth Module*. Eds. Keirse, Renfrew, Neilson and Crowther. In: *The Cochrane Pregnancy and Childbirth Database* [database on disk and CDROM]. The Cochrane Collaboration; Issue 2, Oxford: Update Software; 1995. Available from British Medical Journal Publishing Group, London.
35. Schulz, K.F, I. Chalmers, R.J. Hayes and D.G. Altman. 1995. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *Journal of the American Medical Association* 273: 408-12.
36. Blichert-Toft, M. og J. Beckman. 1981. ben information: en begrænsende faktor i klinisk forskning. *Ugeskrift for Læger* 143: 1199-202.
37. Chalmers, I., M. Enkin and M.J.N.C. Keirse eds. 1989. *Effective care in pregnancy and childbirth*. Oxford: Oxford University Press.
38. Chalmers, T.C. 1974. The impact of controlled trials on the practice of medicine. *Mount Sinai Journal of Medicine* 41: 753-9.
39. Easterbrook, P.J, J.A. Berlin, R. Gopalan and D.R. Matthews. 1991. Publication bias in clinical research. *Lancet* 337: 867-72.
40. Haynes, B.R. 1990. Loose connections between peer-reviewed clinical journals and clinical practice. *Annals of Internal Medicine* 113: 724-8.

41. Sacks, H., T.C. Chalmers and H. Smith. 1982. Randomized versus historical controls for clinical trials. *Annals of Internal Medicine* 72: 233-40.
42. Chalmers, T.C. 1988. The role of clinical trials in health care: a personal view. *Proceedings of the XIVth International Biometric Conference; 1988 July 18-23; Namur. Gembloux: Société Adolphe Quetelet: 47-56.*
43. Moertel, C.G. 1984. Improving the efficiency of clinical trials: a medical perspective. *Statistics of Medicine* 3: 455-65.
44. O'Brien, W.M. 1968. Indomethacin: a survey of clinical trials. *Clinical Pharmacology and Therapeutics* 9: 94-107.
45. The EC/IC Bypass Study Group. 1985. Failure of extracranial-intracranial arterial bypass to reduce the risk of ischemic stroke. Results of an international randomized trial. *New England Journal of Medicine* 313: 1191-1200.
46. Tfelt-Hansen, P., P. Henry, L.J. Mulder, R.G. Scheldewaert, J. Schoenen and G. Chazot. 1995. The effectiveness of combined oral lysine acetylsalicylate and metoclopramide compared with oral sumatriptan for migraine. *Lancet* 346: 923-6.
47. The Cardiac Arrhythmia Suppression Trial (CAST) Investigators. 1989. Preliminary report: effect of encainide and flecainide on mortality in a randomized trial of arrhythmia suppression after myocardial infarction. *New England Journal of Medicine* 321: 406-12.
48. Aboulker, J.P. and A.M. Swart. 1993. Concorde Coordinating Committee. Preliminary analysis of the Concorde trial. *Lancet* 341: 890-1.
49. Riggs, B.L., S.F. Hodgson, W.M. O'Fallon, E.Y.S. Chao, H.W. Wahner, J.M. Muhs et al. 1990. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. *New England Journal of Medicine* 322: 802-9.
50. Lewis, E.J., L.G. Hunsicker, S.-P. Lan, R.D. Rohde and J.M. Lachin for the Lupus Nephritis Collaborative Study Group. 1992. A controlled trial of plasmapheresis therapy in severe lupus nephritis. *New England Journal of Medicine* 326: 1373-9.
51. Packer M., J.R. Carver, R.J. Rodeheffer, R.J. Ivanhoe, R. DiBianco, S.M. Zeldis et al. 1991. Effect of oral milrinone on mortality in severe chronic heart failure. *New England Journal of Medicine* 325: 1468-75.
52. Gøtzsche, P.C., A. Liberati, P. Luca and V. Torri. 1996. Beware of surrogate outcome measures. *International Journal of Technology Assessment in Health Care* 12 (2): 238-246.
53. Moore, T.J. 1995. *Deadly medicine*. New York: Simon & Schuster.
54. Impact Research Group. 1984. International mexiletine and placebo antiarrhythmic coronary trial: I. Report on arrhythmia and other findings. *Journal of the American College of Cardiology* 4: 1148-63.
55. Boissel, J., J. Collet, M. Lievre and P. Girard. 1993. An effect model for the assessment of drug benefit: example of antiarrhythmic drugs in postmyocardial infarction patients. *Journal of Cardiovascular Pharmacology* 22: 356-63.
56. The Coronary Drug Project Research Group. 1980. Influence of adherence to treatment and response of cholesterol on mortality in the coronary drug project. *New England Journal of Medicine* 303: 1038-41.
57. Psaty, B.M., S.R. Heckbert, T.D. Koepsell, D.S. Siscovick, T.E. Raghunathan, N.S. Weiss et al. 1995. The risk of myocardial infarction associated with antihypertensive therapies. *Journal of the American Medical Association* 274: 620-5.
58. Russell, I. 1995. Evaluating new surgical procedures: needs collaboration between surgeons and trialists [editorial]. *British Medical Journal* 311: 1243-4.
59. Forsknings- og Teknologiministeriet. 1993. *Forsknings- og Teknologipolitik 93. Redegørelse til Folketinget*. København: Forsknings- og Teknologiministeriet.
60. Grant, A., J. Sleep, H. Ashurst and J.A.D. Spencer. Dyspareunia associated with the use of glycerol-impregnated catgut to repair perineal trauma. Report of a 3-year follow-up study. *British Journal of Obstetrics and Gynaecology* 96: 741-3.
61. Gam, A.N., H. Thorsen and F. Lønnberg. 1993. The effect of low-level laser therapy on musculoskeletal pain: a meta-analysis. *Pain* 52: 63-6.
62. Mari, J.J. and D. Streiner. [revised 21 July 1995]. The effects of family intervention for those with schizophrenia. In: *Schizophrenia Module*. Eds. C.J. Anderson and J. De Jeus Mari. In: *The Cochrane Database of Systematic Reviews* [database on disk and CDROM]. The Cochrane Collaboration; Issue 2, Oxford: Update Software; 1995. Available from British Medical Journal Publishing Group, London.
63. Hodnett, E.D. [revised 05 December 1994] Support from caregivers during childbirth. In:

- Pregnancy and Childbirth Module. Eds. Keirse, Renfrew, Neilson and Crowther. In: The Cochrane Database of Systematic Reviews [database on disk and CDROM]. The Cochrane Collaboration; Issue 2, Oxford: Update Software; 1995. Available from British Medical Journal Publishing Group, London.
64. Neilson JP. [revised 12 May 1994]. Cardiotocography for antepartum fetal assessment. In: Pregnancy and Childbirth Module. Eds. Keirse, Renfrew, Neilson and Crowther. In: The Cochrane Pregnancy and Childbirth Database [database on disk and CDROM]. The Cochrane Collaboration; Issue 2, Oxford: Update Software; 1995. Available from British Medical Journal Publishing Group, London.
65. Stadslægen. 1995. Medical audit: Intrauterin og neonatal dødelighed i forbindelse med københavnske kvinders fødsler 1993-94. København.
66. Holli, K. and M. Hakama. 1989. Treatment of the terminal stages of breast cancer. *British Medical Journal* 298: 13-4.
67. Non-small Cell Lung Cancer Collaborative Group. 1995. Chemotherapy in non-small cell lung cancer: a meta-analysis using updated data on individual patients from 52 randomised clinical trials. *British Medical Journal* 311: 899-909.
68. Statens Sundhedsvidenskabelige Forskningsråd. 1991. Sundhed og sygdom II - en strategiplan fra SSVF. København: Statens Sundhedsvidenskabelige Forskningsråd.
69. Danish Council for Research Policy. 1992. International evaluation of Danish health research: final report (SOFIE). Copenhagen: Danish Council for Research Policy.
70. Chalmers, I. and W.A. Silverman. 1987. Professional and public double standards on clinical experimentation. *Controlled Clinical Trials* 8: 388-91.
71. Smithells, R.W. 1975. Iatrogenic hazards and their effects. *Postgraduate Medicine* 15: 39-52.
72. Bernard, C.L. 1866, reprinted 1966. Introduction à l'Etude de la Médecine Expérimentale. London: Garnier-Flammarion.
73. Gjørup, T., H. Kelbæk, B. Vestergaard, O. Munck and J. Godtfredsen. 1986. Prospective, randomised, double-blind study of radionuclide determination of left-ventricular ejection fraction in acute myocardial infarction. *Lancet* i:583-5.
74. Sundhedsstyrelsens Udvalg for Medicinsk Teknologivurdering. 1996. National strategi for medicinsk teknologivurdering. København.
75. The Cochrane Pregnancy and Childbirth Database [database on disk and CDROM]. 1995. The Cochrane Collaboration; Issue 2, Oxford: Update Software. Available from British Medical Journal Publishing Group, London.
76. The Cochrane Database of Systematic Reviews [database on disk and CDROM]. 1995. The Cochrane Collaboration; Issue 2, Oxford: Update Software. Available from British Medical Journal Publishing Group, London.
77. Arbejdsgruppen vedrørende Medicinsk Teknologivurdering. 1994. Medicinsk teknologivurdering: hvad er det? København: Sundhedsstyrelsen.
78. Gillett, R. and J. Harrow. 1993. Prescriptions for medical research. I - Management within the Medical Research Council. *British Medical Journal* 306: 1668-72.
79. Riis, P. 1991. Ethiske synspunkter. In: Good Clinical Practice: kvalitetskrav til kliniske undersøgelser: 57-63. Ed. I. Schou. København: Academic Publishing.