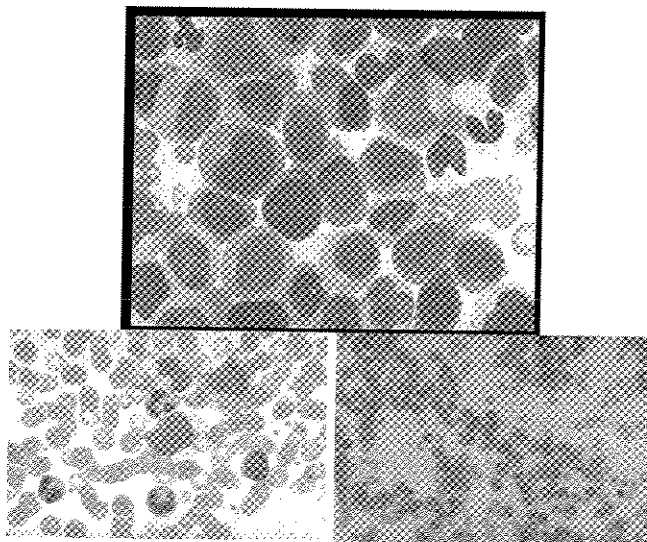


***Fordeler og ulemper ved bruk av kliniske studier med spesiell vekt på hematologiske studier.***

5. årsoppgave, stadium IV i medisinstudiet ved Universitetet i Tromsø

Av  
Geir Klepaker, Kull 2000.



***Veiledet av Inger Marie S. Dahl***

*Tromsø 15. September 2005*

## Innholdsfortegnelse

Resymé .....	2
Introduksjon.....	4
Metode.....	4
Resultat .....	6
<i>Omtale av de aktuelle maligne hematologiske lidelsene</i> .....	6
<i>The Nordic Myeloma Study Group (NMSG)</i> .....	7
<i>The Nordic Myelodysplasia (MDS )Group</i> .....	8
<i>Protokollene</i> .....	8
<i>Nøkkelartikler</i> .....	24
Diskusjon .....	38
<i>Høyeste bevisnivå</i> .....	38
<i>Bruk av metaanalyser</i> .....	39
<i>Hva med prosedyrer og operasjoner?</i> .....	40
<i>Muliggjøring av multicenter forsøk</i> .....	41
<i>Tolking av kliniske forsøk</i> .....	42
<i>Like riktig og likeverdig?</i> .....	43
<i>Kokebokmedisin</i> .....	44
<i>Administrasjon og kostnader</i> .....	44
<i>Fordel ved deltakelse for forsker og pasient</i> .....	46
<i>Interrimanalyser</i> .....	47
<i>Helhetsforskning</i> .....	47
<i>Noen etiske betraktninger</i> .....	49
<i>Til hjelp ved prioriteringer</i> .....	49
<i>Særtrekk ved hematologiske kliniske studier</i> .....	51
Konklusjon .....	54
Tabeller.....	56
Figurer .....	58
Referanseliste .....	61

## Resymé

Oppgaven er en litteraturstudie som søker å belyse fordeler og ulemper ved bruk av kliniske studier. Det er lagt spesiell vekt på studier som inkluderer hematologiske pasienter. Aktuelle artikler er funnet ved søk i Pubmed og Embase. Jeg har også fått noen nøkkelartikler av min veileder.

Forskningsprotokollene som pågår eller nylig er avsluttet ved hematologisk seksjon, Universitetssykehuset i Nord-Norge, er grundig gjennomgått. Protokollene illustrerer hvordan et klinisk forsøk gjennomføres. I tillegg har jeg lært mer om de maligne hematologiske lidelser studiene omhandler.

Artiklene konkluderer med at resultatet av kliniske studier representerer det høyeste bevisnivået innenfor medisinen, i tråd med tradisjonen for "evidence based medicin". Gode kliniske studier gir oss kunnskap om mennesket som en funksjonell enhet og et helhetsbilde av pasientene ved blant annet å inkludere livskvalitetsstudier. Helt nye perspektiver kan også åpne seg. Det viktigste med klinisk forskning er imidlertid å finne ut hvordan den metode eller det medikament som utprøves best kan gagne pasientene. Med bakgrunn i kliniske studier kan bedre behandlingsmetoder utvikles samt unyttige og endog skadelig behandling fjernes. I tillegg kan resultatene danne grunnlag for prioritering mellom ulike pasientgrupper. Det er dessuten vist at pasienter som deltar i kliniske forsøk har en positiv korttidseffekt, kalt forsøkeseffekten.

Kliniske forsøk har også sine negative sider. Etisk og juridisk kan det være betenkelig å bruke mennesker som "forsøkskaniner". Hvem sine interesser skal telle og hvilke parametere skal vektlegges? Noen resultater kan ikke måles på en objektiv måte; - hva da? Det viser seg at kliniske forsøk heller ikke er spesielt egnet til å sammenligne ulike prosedyrer og operasjoner. Siden kliniske forsøk er et så kraftig vitenskapelig verktøy er det også sårbart m.h.p. misbruk. Det kan oppstå feil i tolkningen av statistiske beregninger ved at for eksempel P-verdien tillegges for stor vekt, fordi man ikke tar hensyn til forsøkseffekter, p.g.a ulike måter å designe et forsøk på etc. Dermed blir krav til dobbeltblinding og randomisering viktig.

Når det gjelder maligne hematologiske lidelser er kliniske forsøk spesielt viktig fordi så vel prevalens som insidens er lav og et større antall pasienter bør undersøkes før konklusjoner trekkes.

Selv om det finnes en del ulemper ved bruk av kliniske studier oppveies disse av fordelene. Jeg er ikke tvil i at denne typen forskning er til det beste for pasientene.

## Introduksjon

I de siste tiår er det blitt stadig mer vanlig å benytte seg av kliniske studier ved testing av nye behandlingsformer, inklusive nye medikamenter. I min oppgave har jeg forsøkt å se nærmere på de fordelene og ulempene som finnes ved kliniske studier. Jeg har lagt spesiell vekt på studier som inkluderer pasienter med maligne hematologiske lidelser.

Opgaven er en litteraturstudie hvor jeg har gjennomgått sentrale artikler som belyser fordeler og ulemper ved bruk av kliniske studier. I tillegg har jeg gjennomgått en rekke av de protokollerte studiene som pågår eller nylig er avsluttet ved hematologisk seksjon ved Universitetssykehuset i Nord-Norge. Disse har jeg brukt som eksempler på kliniske studier og de har gitt meg spesiell kunnskap om studier som omhandler maligne hematologiske lidelser.

Forventet utbytte av oppgaven har vært å skaffe meg mer inngående kunnskaper om maligne hematologiske sykdommer, spesielt utredning og behandling. Formålet har også vært å få kunnskap om kliniske protokoller, hvordan de planlegges og utføres. Jeg har videre studert fordeler og ulemper ved kliniske studier i forhold til pasientene, utprøvene, den vitenskapelige nytteverdi og samfunnet. På flere områder kan disse interessene ikke skilles og vil derfor bli omtalt sammen.

## Metode

Som metoder har jeg valgt en litteraturstudie. Jeg har da gjennomgått de protokollene som pågår eller som nylig er avsluttet ved hematologisk seksjon, Universitetssykehuset Nord-Norge (UNN). Disse har jeg fått kopi av fra seksjonsoverlege Inger Marie S. Dahl. De sykdommene som blir behandlet i henhold til disse protokollene blir kort gjort rede for. Som kilde har jeg vesentlig anvendt læreboka "Blodsykdommer" av Brinch et.al.(1).

Artiklene som er referert for å belyse pro og kontra i forhold til kliniske studier er funnet gjennom søk i Pubmed, Embase og store velrenommerte tidsskrifter som BMJ og NEJM. Jeg har også fått kopi av en del artikler fra min veileder. Søkeord har for eksempel vært : " Randomised controlled trial", "clinical trials" kombinert med "pro and contra" Et stort antall artikler er blitt

gjennomgått, hvorefter det ble gjort en kritisk utvelgelse av de artiklene som best belyser fordeler og ulemper ved bruk av kliniske studier.

## Resultat

### De maligne hematologiske sykdommene som protokollene omhandler

**Myelomatose** har en insidens på 6 pr 100 000 innbyggere pr år og rammer hovedsakelig eldre mennesker. Ved myelomatose foregår det en malign proliferasjon av plasmaceller i beinmargen, muligens ved at abberante B-lymfocytter fra kimsenter i lymfoid vev slår seg ned i beinmarg ved hjelp av adhesjonsmolekyler. Plasmacellene har ofte et umodent utseende, med sentral kjerne og tydelige nucleoli. Denne maligne klonen kan invadere hele beinmargsparenchymet og gi beinmargssvikt. Gjennom et nettverk av gjensidig cytokinpåvirkning trekkes osteoblaster, osteoclaster og stromale celler med i sykdomsprosessen. Dette fører bla til avkalkning og svekkelse av skjelettet. Symptomene på myelomatose er knyttet til dette ved at pasientene får benskjørhet, patologiske frakturer, skjelettsmerter og hypercalcemi. De maligne cellene lager monoklonalt IgG eller IgA, som gir et karakteristisk bånd ved serumelektroferese. Plasmacellene kan også produsere lette kjeder, som sammen med hypercalcemi er med på å ødelegge nyretubuli og gi nyresvikt. Er det svært høy konsentrasjon av immunglobuliner vil dette gi hyperviskositet i blodbanen som kjennetegnes ved hodepine og omtåketet. Akseptert behandlingen av Myelomatose per i dag er cytostatika, stråler og autolog stamcelletransplantasjon hos pasienter under ca. 65 år.(1)

**Myelodysplastisk syndrom (MDS)** er en heterogen gruppe sykdommer med en innsidens på 4 pr 100 000. MDS kan være en forløper til akutt leukemi. Som ved akutt leukemi regner vi med at tilstanden skyldes en abnorm klon av hematopoetisk stamceller i beinmarg som har fått en vekstfordel. Avkommet av denne klonen er ikke funksjonelt og har kort levetid. MDS diagnosen stilles etter at tilstanden har vært stabil i 2 måneder. Resultatet er en hypercellulær beinmarg og perifer cytopeni. Pasienter med MDS har ofte multiple kromosomale avvik. Ca. 25 % av pasienter med MDS utvikler akutt myelogen leukemi, det vil si  $> 30\%$  blaster i beinmargen.(1)

**Kronisk myelogen leukemi (KML)** er karakterisert av malign klonal vekst av hematopoetiske stamceller og opphopning av celler i den myelogene rekken. Det gir et bloduttstryk som viser

alle de ulike cellene i den myelogene rekken;- "blodutstryket ligner et normalt benmargsutstryk". Symptomer kan være lavgradig feber, vekttnap, økt abdominalomfang og hepato- og splenomegali. KML kan forekomme i flere ulike stadier. I den kroniske fasen, som kan vare fra uker til år, er det relativt få blaster (<12 %) i benmargen . Majoriteten av pasientene (ca. 2/3) vil gå over i en akselerert fase med økt antall blaster, økende splenomegali, myelofibrose og nye kromosomforandringer. Denne fasen kommer vanligvis etter noen få år. De øvrige 1/3 får en blasttransformasjon etter kort tid med >30 % blaster i benmargen. Over 90% av pasientene har en translokasjon mellom kromosom 9 og 22, det såkalte Philadelphiakromosom. Nye medikamenter av typen tyrosinkinasehemmere kan helbrede sykdommen, men resistens kan utvikles. Kurativ behandling for øvrig er allogen stamcelletransplantasjon til yngre pasienter, mens eldre KML kun får livsforlengende medikamentell behandling med hydroksyurea, interferon eller Myleran. For å få kontroll under en blasttransformasjon brukes cytostatikaregimer som ved AL.(1)

### *The Nordic Myeloma Study Group*

The Nordic Myeloma Study Group (NMSG) er en nordisk studiegruppe som jobber med myelomatose. Studiene har vært administrert slik at man har kontaktpersoner ved de fleste lokale sykehus i Norden, regionale koordinatorene ved Universitetsklinikkene og et sentralt sekretariat. Hvert nivå har sine bestemte ansvarsområder. Kontaktpersonen har ansvar for inkludering og registrering av pasienter og for å sende data til de regionale koordinatorene. Disse dataene blir skrevet inn på forhåndslagde blanketter. Den regionale koordinatoren er et bindeledd mellom klinikkene og det sentrale sekretariatet. Koordinatoren har ansvar for innsamling av blankettene og videresending av dem til sekretariatet. Dessuten fungerer vedkommende som en rådgiver og holder halvårlige samlinger innen sin region for å oppdatere kontaktpersonene og diskutere eventuelle problemer de har. Det sentrale sekretariatet har ansvaret for databehandlingen og for å holde koordinatorene løpende oppdatert om studienes framgang, problemer som oppstår og annen nødvendig informasjon. De arrangerer også koordinatormøter to ganger i året.( Dahl, pers. meddelelse)



### The Nordic MDS Group

The Nordic MDS (myelodysplasi) Group omfatter alle helseregioner i Sverige og Norge to regioner i Danmark og en (Helsinki) i Finland. Gruppen er organisert på samme måte som NMSG, bortsett fra at det sentrale sekretariat er permanent lokalisert til Huddinge Sjukhus i Stockholm.( Dahl, pers. med)

### Protokollene

Som en del av oppgaven har jeg sett på de hematologiske forskningsprotokollene som i dag utføres eller har vært utført i nær fortid ved hematologisk seksjon på Universitetssykehuset Nord Norge (UNN)

1) Den eldste protokollen som jeg har valgt å ta med har betegnelsen ”NMSG #5/94: **Intensivbehandling med stamcellestøtte av yngre myelompasienter**”.(2) Den er meget illustrativ fordi den viser hvordan et klinisk forsøk har resultert i at behandlingen av yngre myelomatosepasienter er endret. Studien startet 1/1-94 og varte i 4 år, dvs den var ferdig i -98.

*Bakgrunn:* Melfalan-prednisolon har siden slutten av 60-tallet vært førstehåndsbehandling ved myelomatose, men dette er ikke en kurativ behandling. Langtidsoverlevelsen har ikke forandret seg nevneverdig, selv om stadig nye cytostatikaregimer har vært prøvd. Studier har imidlertid vist at høydose Melfalan har gitt høy responsfrekvens ved myelomatose, men mortaliteten har vært uakseptabel høy, vesentlig pga langvarig aplasi. På begynnelsen av 80-tallet begynte man derfor å gi allogene stamceller fra benmarg, men denne behandling begrenses av pasientens alder og av at egnet donor ikke er tilgjengelig for mer enn 25% av pasientene. Man startet derfor forsøk med høydose kjemoterapi etterfulgt av autolog stamcellestøtte, såkalt HMAS, til myelomatosepasienter. Tanken var at ved å kombinere høydose kjemoterapi med kjent effekt på myelomceller etterfulgt av tilførsel av egne stamceller ville apasifasen reduseres og dermed dødeligheten. (3,4). HMAS har senere vært anvendt i økende utstrekning. De senere år samles stamceller fra blod og det gis støttebehandling med hematopoietiske tilvekstfaktorer etter at disse cellene er tilbakeført.

*Design:* Studien er enarmet der kontrollene var historiske. Dette ble gjort fordi blinding og randomisering ikke lot seg gjennomføre da behandlingen allerede var kjent for pasientene og delvis tatt i bruk, noe som gjorde det vanskelig å motivere pasientene for inklusjon. I tillegg ville ha tatt 6-8 år før studien kunne blitt avsluttet, noe som ikke var hensiktsmessig på et forskningsfelt i rask utvikling. I studien besluttet NMSG å gi tilbud til myelomatosepasienter under 60 år. Aldersgrensen er satt slik fordi det er vist økende morbiditet og mortalitet med økende alder grunnet ved denne intensive behandlingen. Selve studien foregikk i fire faser, der det måtte påvises behandlingsrespons etter hver fase. Pasientmaterialet som fikk HMAS ble sammenlignet med et tilsvarende antall pasienter fra samme opptaksområde som hadde fått konvensjonell behandling.

Fase I var en induksjonsbehandling med tre kurer Vinkristin, Doxorubicin og Dexametason (VAD) for å undersøke om myelomcellene var kjemosensitive. Fase II bestod av cyclofosamid og tilvekstfaktor for å mobilisere stamceller fra beinmarg til perifert blod, fulgt av autolog stamcellehøsting. I fase III ble det så gitt høydose Melfalan, etterfulgt av autolog stamcelletilførsel. Fase IV bestod av vedlikeholdsbehandling med Interferon  $\alpha$ -2b. Ved relaps har legen fritt kunnet velge behandling. IFN ble gitt fordi det på tidspunktet for studiestart forelå primærdata på at medikamentet forlenget total overlevelse etter HMAS.

På diagnosetidspunktet ble pasienter som hadde sagt seg villig til å delta i studien anmeldt av sin lege til den regionale koordinatoren. Regional koordinator fylte ut en sjekklister som ble sendt til det sentrale sekretariatet. Det ble tatt beinmargsprøve, funksjonsstatus, røntgen og en rekke blod- og urinprøver som er nærmere spesifisert i protokollen. Det ble også tatt serum, plasma og beinmarg som ble frosset ned for senere undersøkelser som ikke var nærmere definert. En helserelatert livskvalitetsundersøkelse av de pasientene som deltok i studien ble også gjennomført.

Pasienten ble undersøkt etter hver fase og hver 6 uke etter fase IV. Det ble igjen tatt en rekke blod- og urinprøver, bla til bestemmelse av M-komponent. Beinmargsprøver ble tatt etter fase I og II, ved mistanke om relaps, 3 og 6 måneder etter høydose Melfalan, seinere hver 6 måned. Røntgen ble tatt ved diagnosetidspunkt og 3 måneder etter fase III. Alle bivirkninger, komplikasjoner og død er blitt meldt til det sentrale sekretariatet på spesielle blanketter. Alvorlige bivirkninger og død skal meldes innen 24 timer.

*Inklusjons/eksklusjonskriterier:*

- Pasienter under 60 år
- sikker diagnose myelomatose
- det foreligger behandlingskrevende sykdom
- det finnes ikke kontraindikasjoner for fase I (alvorlig hjerte-karsykdom, alvorlig lungesykdom, annen malignitet, manglende samarbeidsevne, terminalt syk)

Alle disse momentene er veldefinerte i studiens protokoll. For at pasientene skulle kunne gå videre fra en fase til neste krevdes også at visse kriterier skulle være oppfylt. For eksempel krevdes det at før pasienten gikk videre til fase III skulle det ikke være tegn til progressiv myelomsykdom, nyresvikt, hepatopati eller alvorlig pågående infeksjon. Det skulle også foreligge adekvate stamcellehøstinger. Pasienter som ikke oppfylte disse kriteriene ble ekskludert fra studien.

*Administrasjon:* Studien var et multisenter forsøk hvor sykehus fra Norge, Sverige og Danmark deltok. Sykehusene har variert fra små sykehus til universitetssykehus, som har vært administrert etter den vanlige prosedyren for NMSG's ulike nivåer. Det sentrale sekretariatet befant seg ved universitetssykehuset i Lund i Sverige. Livskvalitetstudiene har blitt drevet og analysert ved Ullevål Universitetssykehus i Oslo.

*Resultat :* Resultatene ble publisert i BLOOD (5) og i Svenska Läkartidningen i 2000 (6) og i Tidsskrift for Norsk Lægeforening i 2001(7). 274 evaluerbare pasienter deltok i hver gruppe. Median overlevelse var 44 måneder i kontrollgruppen og 62 måneder i høydosegruppen ( $p=0,0001$ ). Ved hjelp av spørreskjema (EORTC QLQ) ble en moderat nedsatt livskvalitet registrert i den intense behandlingsfasen men ingen forskjeller senere. Nytt kostnadsanalyser gav en pris på 249 000 Nkr per livskvalitetsjustert leveår (Qaly), noe som ansees akseptabelt for samfunnet.

2) Den andre forskningsprotokollen i regi av NMSG jeg skal referere betegnes #7/ 1998, **"Intensivbehandling av yngre myelompasienter"**(8). Studien startet 1.1 1998 og inkluderte pasienter i 15 måneder. Pasientene ble deretter fulgt i 24 måneder, slik at hele studien hadde en beregnet tidsramme på 39 måneder fra behandlingsstart.

Som i den forrige studien sammenliknes den tradisjonelle behandlingen med Melfalan-Prednisolon med høydose kjemoterapi (Melfalan) etterfulgt av autolog stamcelle støtte (HMAS), men nå er aldersgrensen økt til 65 år . Også denne studien er avsluttet, men jeg har tatt den med som eksempel på en klinisk studie som har ført til at behandling av yngre myelomatosepasienter i Norden og verden forøvrig er endret. Det viste seg nemlig at også pasienter mellom 60-65 år som ble behandlet med HMAS hadde en 3 års forlenget overlevelse i forhold til de historiske kontrollene.

*Bakgrunn:* Som nevnt i den forrige studien har Melfalan- Prednisolon siden 1960- tallet vært standard behandling ved myelomatose. Langtidsoverlevelsen har ikke forandret seg nevneverdig, selv om nye cytostatika og kombinasjonsregimer har vært prøvd. Studier har imidlertid vist at høydose Melfalan har gitt høy responsfrekvens på myelomatose, men mortaliteten har vært høy pga langvarig aplasi. Mortaliteten har sunket ved å kombinere høydose Melfalan med tilførsel av autolog beinmarg. Senere har det vist seg at også perifere stamceller kan benyttes, etter at pasientene er forbehandlet med Cyclofosfamid og hematopoetiske vekstfaktorer. Liknende HMAS-studier har vist 70-90 % respons på denne behandlingen, mens 20-50 % av pasientene har oppnådd komplett respons. Median overlevelsestid har vært på 4-5 år. Det har tidligere bare vært gjennomført en randomisert studie som har sammenliknet konvensjonell behandling med HMAS, på et materiale på 200 pasienter. Denne viste en signifikant respons hos de pasienter som ble behandlet med HMAS i forhold til de som fikk den konvensjonelle behandlingen.

*Design:* NMSG hevder ikke at HMAS har et kurativt potensiale, men at målet med behandlingen er å gi en god og langvarig kontroll av sykdommen med lav behandlingsrelatert morbiditet og mortalitet. Tidligere studier har vist at yngre pasienter tolererer bivirkningene bedre og øvre aldersgrense for innmelding i protokollen har derfor vært 65 år. De som er eldre har fått den konvensjonelle behandlingen. Et problem er at HMAS er blitt godt kjent i befolkningen og at det ikke er praktisk mulig med en randomisering mellom HMAS og konvensjonell behandling. Det vil dermed blir umulig å få med nok pasienter som kontroller til studien innen en rimelig tid. Studiepopulasjonen sammenlignes derfor med en gruppe veldefinerte historiske kontroller. Ved diagnosetidspunkt ble pasienter som var villige til å delta i studien innmeldt av sin lege til den regionale koordinatoren. Regional kooordinator fylte ut en sjekklister som ble sendt til det sentrale sekretariatet. Det ble tatt beinmargsprøve, funksjonsstatus, røntgen og en rekke blod og urinprøver som er nærmere spesifisert i protokollen. Det ble også tatt serum, plasma og beinmarg som fryses ned for senere undersøkelser.

Behandlingen inndeles i fire faser, de samme som beskrevet i forrige protokoll.

Pasienten ble undersøkt etter hver fase og hver 6 uke etter fase 4. Da tok man en rekke blodprøver og urinprøver, bla bestemmelse av M-komponenten. Beinmargsprøver ble tatt etter fase 1 og 2, ved mistanke om relaps, etter 3 og 6 måneder etter høydose Melfalan, deretter hver 6 måned. Røntgen ble tatt ved diagnosetidspunkt og 3 måneder etter fase 3. Alle bivirkninger, komplikasjoner og død er umiddelbart (innen 24 timer). blitt meldt til det sentrale sekretariatet.

*Inklusjons/eksklusjonskriterier:* Inklusjonskriterier for studien har vært:

- Alder under 65 år
- sikker diagnose myelomatose
- det foreligger behandlingskrevende sykdom
- det ikke finnes kontraindikasjoner for fase I

Alle disse momentene var veldefinerte i studiens protokoll. For at pasientene skulle kunne gå videre fra en fase til neste kreves også at visse kriterier skal være oppfylt. For eksempel kreves at før de gikk til fase III skulle det ikke foreligge progressiv sykdom, det skulle foreligge et adekvat antall stamceller, ikke foreligge nyresvikt, hepatopati eller alvorlig pågående infeksjon. Pasienter som ikke oppfyller disse kriteriene ble ekskludert fra studien.

*Administrasjon:* Studien var et multicenter forsøk hvor 107 sykehus fra Norge, Sverige og Danmark har deltok. Sykehusene som har variert fra små sykehus til universitetssykehus, og administrert etter den vanlige prosedyren for NMSG. Det sentrale sekretariatet har vært ved universitetssykehuset i Lund i Sverige.

3) Neste protokoll fra NMSG jeg har valgt ut har betegnelsen #8/00: "Effekten av i.v. bisfosfonat på helse-relatert livskvalitet og skjelettmorbiditet ved nydiagnostisert behandlingskrevende myelomatose. En prospektivt randomisert dobbelblindet dose-efficitive studie med cost-utility analyse "(9). Den startet i 2001 og var beregnet å ha en inklusjonstid på to år, til 750 pasienter var inkludert. Inklusjonen har imidlertid tatt fire år og avsluttes ikke før høsten 2005. Pasientene følges etter inklusjonen i minst 36 måneder.

*Bakgrunn:* Det er kjent at bisfosfonater hemmer osteoklastaktiviteten, og har dokumentert effekt på hypercalcemi ved maligne sykdommer (10), herunder myelomatose.

Den økte osteoklastaktiviteten skyldes at de maligne plasmacellene lager ulike cytokiner bl.a Interleukin 6 som aktiverer osteoklastene. Den økte osteoklastaktiviteten gir en nedbrytning av skjelettet med økt fraktur-risiko og skjelettsmerter som resultat. Tre større randomiserte studier med profylaktisk bisfosfonat (11,12,13) synes å vise symptomatisk effekt ved myelomatose. Problemet er at det er utviklet flere bisfosfonater og doseringen er uklar da det er brukt ulike typer med ulik potens. Det som imidlertid er uavklart er vikning på pasientenes livskvalitet og hvilken bisfosfonatdose som er optimal. Problemstillingene i denne studien er:

1 : Hvilken bisfosfonatdose er den optimale og kan effekt oppnås med 1/3 dose, eller bør dosen økes utover 90 mg? Påvises ikke dose-respons effekt kan man stille spørsmålet om myelomatosepasienter skal ha bisfosfonater rutinemessig.

2: Bedrer bisfosfonatbehandling pasientenes livskvalitet? Viktigste formål med bisfosfonater til myelomatosepasienter er å redusere og bremse utviklingen av skjelettsykdommen. Pasientenes livskvalitet er det primære endepunktet i studien.

*Design:* Studien er en randomisert dobbelblindet dose-respons studie. Pasientene randomiseres fortløpende til en av to grupper. Den ene gruppa får 30mg pamidronat (Aredia®) i.v., den andre 90 mg pamidronat (Aredia®) i.v. Pamidronat( Aredia®) gis som i.v. med 500 ml glukose hver måned i 36 måneder. Randomiseringen skjer også på grunnlag av stratifisering, dvs ar det innenfor hver gruppe finnes subgrupper: HMAS vs. konvensjonell terapi, WHO performance status, +/- thaliodomid og  $\beta$ -2-mikroglobulin. Legemiddelfirmaet har ikke kunnet levere forsøksmedisinen blindet, slik at blindingen foregår via apotek eller av en ansatt ved avdelingen. Behandlende lege, sykepleier eller pasient får ikke vite hvem av gruppene pasienten tilhører slik at dette er et dobbelblindet forsøk. Ved diagnositidspunktet tas røntgen og 10 ml serum og plasma som fryses ned. I tillegg tas de prøver som er nødvendige for å stille diagnosen myelomatose, inkl. beinmargsprøve. Røntgenundersøkelsen gjentas etter 9 og 24 måneder. Ellers tas en del blodprøver som for eksempel INR og kreatinin angitt i protokollen hver 3. måned. Hver 3.måned fylles også et livskvalitetsskjema, EORTC QLQ-C30, ut, det standardskjema NMSG anvender ved livskvalitetsstudier av myelomatosepasienter

*Inklusjons/eksklusjonskriterier:* Alle med behandlingskrevende nyoppstått myelomatose kan inkluderes i studien. Det er intet hinder at de er med i andre studier eller får annen behandling enn den etablerte (Melfalan-Prednisolon, HMAS).

*Inklusjonskriterier:*

- Oppfylder diagnosekriteriene
- Har fylt ut livskvalitetsskjemaer
- Omattes ikke av noen eksklusjonskriterier
- Har gitt skriftlig informert samtykke
- Planlegges HMAS eller konvensjonell kjemoterapi?
- WHO performance status
- Deltagelse i thalidomidprotokoll?
- Konsentrasjon av beta-2-mikroglobulin

*Eksklusjonskriterier:*

- forventet livslengde på mindre enn 3 måneder
- samtidig annen aktiv malign sykdom
- vært behandlet med bisfosfonater i mer enn 3 måneder innenfor siste 6 måneder
- s- kreatinin > 400 $\mu$ mol/l
- månedlige injeksjoner ikke er praktisk mulig
- vil ikke delta

*Administrasjon:* Studien er et multicenter forsøk som involverer sykehus i Norge, Sverige og Danmark. Vanlige ansvarsfordelingen innenfor NMSG med lokale klinikker, regional koordinator og sentralt sekretariat i København foreligger. Sekretariat for livskvalitetsundersøkelsen befinner seg på Ullevål sykehus. Det er også opprettet en Komité som skal gjøre en interimanalyse som er uavhengig av studien når de 200 første pasientene har vært fulgt i 6 måneder. Viser det seg at behandlingen er overlegen signifikant bedre ( $P < 0,001$ ), eller at bivirkningene er uakseptabelt store, har komiteen mandat til å stoppe studien. Det gjøres flere interimanalyser når komiteen bestemmer det.

**4) Protokoll #11/00 fra NMSG kalles VAD mot Cy+Dex (14).** Denne startet i januar 2001 og ble avsluttet desember 2004.

*Bakgrunn:* Dette var en studie for pasienter der HMAS-behandling var planlagt. Tidligere studier, som protokoll # 7/98 (2), har vist at induksjonsbehandlingen, som består av Vinkristin, Dexametason og Doxorubicin, har en toksisk dødelighet på 2,1 %. Dessuten døde også 2,1 % av

progredierende myelomatose under VAD-behandlingen. Totalt døde 4,2% under den innledende VAD-behandling. I tillegg hadde ytterligere 4,0% fått progresjon av myelomatosen slik at de ikke kunne gjennomgå HMAS. Totalt utgjør dette et frafall under VAD på 8,2 %. NMSG har vist at det er HMAS som er den meste effektive tumorreduserende behandlingen, og at pasienter som responderer dårlig på VAD kan ha god effekt av HMAS. VAD brukes fordi kuren ikke skader stamcellene slik som Melfalan+Prednisolon gjør. Problemet med VAD er toksisitet og at den må gis som infusjon over 96 timer. Dette vil i praksis si en sykehusinnleggelse. Med denne nye studien vil NMSG undersøke om man kan bruke Cyclofosfamid+Dexametason (Cy+Dex) i stedet for VAD som induksjonsbehandling. Cy+Dex har også den fordel at den kan tas p.o. Målet med induksjonsbehandlingen er en rask og effektiv stabilisering av sykdommen slik at pasienten blir mottakelig for HMAS.

*Design:* Målet for studien er å studere pasienter som gjennomgår full HMAS og se om antallet blir høyere når Cy+Dex velges som induksjonsbehandling sammenlignet med VAD i en åpen randomisert studie. Det er også aktuelt å se på respons, relaps og mortalitet. Samtlige pasienter er under 65 år og det er planlagt behandling med HMAS. Det stratifiseres mellom de nordiske landene og for alder. Alderen stratifiseres i <60 år og 60-64 år. I den konvensjonelle armen gis 3, maks 4 sykler med VAD med syklusintervall på 28 dager, i den eksperimentelle 2, maks 3 sykler med Cy+Dex med syklusintervall på 21 dager. Etter induksjonsbehandlingen er de videre fasene like. Ved primær progresjon og relaps er behandlende lege fri til å velge fortsatt terapi.

Ved registrering av pasienter tas røntgen, M-komponent i serum og urin, beinmargsprøve og en rekke blodprøver som er angitt i protokollen. Det gjøres også en WHO performance status. De videre undersøkelsen gjøres før hver syklus av induksjonsbehandlingen, for hver fase i HMAS og 1,2 og 3 måneder etter fase 3. Da tas som regel beinmargs- og blodprøver. Pasienten er ferdig i protokollen 3 måneder etter fase 3.

*Inklusjons/eksklusjonskriterier:* Alle pasienter med nydiagnostisert myelomatose som planlegges behandlet med HMAS kan inkluderes i studien, dvs de må ha behandlingskrevende sykdom og være < 65 år.

Eksklusjonskriterier er :

\* Kontraindikasjoner mot intensiv cytostatikabehandling

- alvorlig hjertesykdom
- kronisk alvorlig lungesykdom
- annen aktiv malign sykdom



- annen alvorlig sykdom eller organdysfunksjon
- psykisk sykdom (slik at deltakelse og samarbeid er vanskelig)
- terminal syk
- graviditet

\* Tidligere cytostatikabehandling

\* Planlagt allogen stamcelletransplantasjon innen 6 måneder

\* Pasienter som ikke vil delta i studien

Protokollen inneholder også detaljerte retningslinjer for når behandlingen eventuelt skal avbrytes, som ved progressiv myelomsykdom, annen alvorlig sykdom eller dårlig stamcellehøst.

*Administrasjon:* Studien er et multicenter forsøk som involverer sykehus i Norge, Sverige og Danmark. Ansvarsfordelingen innenfor NMSG er som tidligere, med kontaktpersoner på lokale sykehus, regionale koordinatorene og sentralt sekretariat, nå i Göteborg. Ved det sentrale sekretariatet er det en forskningssykepleier som gjennomgår alle data og evt innhenter kompletterende opplysninger. Det blir også gjort stikkprøver av en uavhengig forskningssykepleier av blankettene mot journalene som sikkerhetskontroll.

**5) Studie #12/00: "Thalidomid til pasienter med ubehandlet myelomatose"**(15) er den siste protokollen fra NMSG jeg har valgt å gjennomgå.

Studieledelsen regner med en inklusjonstid på 40 måneder og at pasientene blir fulgt i 40 måneder. Målet er at 240 pasienter inkluderes årlig og at hovedanalysen skjer når 282 "events" (dødsfall) har inntruffet. Man planlegger å rapportere langtidsoverlevelse i 4 år etter hovedanalysen.

*Bakgrunn:* Thalidomid ble syntetisert i 1954 med tanke på anti-allergi effekt, men effekten viste seg å være svak. I stedet registrerte man en hypnotisk effekt og medikamentet ble derfor brukt som sovemiddel fra 1958. Dessverre viste det seg at thalidomid hadde en betydelig teratogen effekt og medikamentet ble trukket fra markedet på 1960-tallet. Senere har det vist seg at medikamentet har effekt på en rekke sykdommer, som aftøs stomatitt, lupus og myelomatose. (16). Effekten skyldes trolig en redusert produksjon av tumor nekrosefaktor (TNF) gjennom at mRNA for TNF degraderes raskere. Det er publisert flere ikke-randomiserte fase 2 studier, men antall deltakere har vært lavt og alle hadde refraktær myelomatose. (17,18,19). NMSG har også utført en studie, #10 (20), der 65 pasienter med avansert myelomatose fikk thalidomid. Den viste at 30 % fikk fall i immunoglobulinkonsentrasjonen og 25 % signifikant fall i plasmacelletall.

Dette var imidlertid en fase 2 studie som ikke var randomisert og thalidomid ble kun gitt til pasienter med avansert myelomatose. Blant de som svarte på behandlingen var det imidlertid flere som hadde overraskende gode resultater. Studien bekreftet at thalidomid kan være et nyttig medikament ved avansert myelomatose. Studier viser doseavhengig bivirkningsprofil hvor søvnighet, utslett, nøyropeni og perifer nevropati er de viktigste bivirkningene(17).

*Design:* Spørsmålet som søkes avklart i denne studien er om det er aktuelt å bruke thalidomid som en del av 1. linjebehandlingen av myelomatose. Det primære endepunktet blir da overlevelse med livskvalitet, tid til respons, frekvens av respons, tid til progresjon, tid til 2. respons, frekvens av 2.respons, tid til 2. progresjon, toksisitet og tid til definitiv svikt under Melfalan+Prednisolon behandling. Med den dokumentasjonen som finnes er det ikke etisk forsvarlig og ikke gi thalidomid til pasienter med avansert myelomatose.

Studien er en dobbeltblindet randomisert studie. Randomiseringen skjer blindet ved Copenhagen trial unit i København, mens sekretariatet befinner seg ved Kontor for Kreftforskning i Trondheim. Siden medikamentet ikke er registrert i de nordiske landene, leveres det fra ett apotek i hvert land til de enkelte sykehus.

Alle som oppfyller inklusjonskriteriene for studien vil få Melfalan-Prednisolonkurer hver 6. uke og randomiseres til om de i tillegg skal få placebo eller thalidomid. Melfalan+Prednisolon gis daglig i 4 dager til pasienten har oppnådd platåfase, dvs stabilisering av sykdom. Thalidomid gis som startdose 100 mg x 2 før dosen økes til 200 mg x 2. Etter oppnådd platåfase reduseres dosen til en vedlikeholdsdose på 100 mg x 2, som gis til 1 relaps. Thalidomid/placebo seponeres etter 2. relaps, da pasienten går ut av studien og behandlende lege stilles fritt i forhold til valg av videre behandling.

Pasientene undersøkes med røntgen, beinmargsutstryk, WHO performace og en rekke blodprøver som er angitt i protokollen. Man registrerer også antall innleggelsesdøgn, legebesøk og sykemeldingsdager m.m. Dette gjøres ved diagnosetidspunkt, 6,12,18,24 uker etter diagnosetidspunkt, og deretter hver 3. måned. Røntgen og beinmargsutstryk tas også på senere tidspunkt, angitt i protokollen.

Det gjøres også livskvalitetsundersøkelse v.h.a. spørreskjemaet EORTC QLQ-C30, som brukes i alle NMSG s livskvalitetsstudier. Det tas også serum, plasma og beinmargsprøver som fryses ned og er tenkt anvendt for å sammenlikne biologiske markører mot kliniske opplysninger.

*Inklusjonskriterier/eksklusjonskriterier:* Det er ingen øvre eller nedre aldersgrense for deltakelse i studien. Alle med behandlingstrengende, tidligere ubehandlet myelomatose kan inkluderes. Det er intet hinder at de er med i protokoll #8/00 (9).

Eksklusjonskriterier er:

- Tidligere behandlet myelomatose med cytostatika eller steroider
- Planlagt HMAS
- Kvinner i fertil alder
- Psykisk sykdom eller reduksjon som forventes å føre til manglende samarbeid
- Manglende samtykke
- Forventet levetid under 3 måneder
- Annen malignitet eller alvorlig sykdom

*Administrasjon:* Studien er et multicenter forsøk som involverer sykehus i Norge, Sverige og Danmark. Vanlige ansvarsfordelingen innenfor NMSG følges, med kontaktpersoner på lokale klinikker, regional koordinator og sentralt sekretariat, nå i Trondheim. Sekretariat for livskvalitetsundersøkelsen har som i tidligere studier vært på Ullevål sykehus. Det er også opprettet en Komité som skal gjøre en interimanalyse uavhengig av studien. Interimanalysen gjøres når 200 pasienter har vært fulgt i 6 måneder. Resultatet av analysen meddeles ikke studiens ledelsesgruppe eller deltagende sentra, med mindre resultatet kan motivere til en endring av timeplanen eller et tidligere avbrudd av studien. Komiteen bedømmer også hvorvidt en ny interimanalyse skal gjøres og tidspunktet for denne. Alvorlige hendelser, som død eller alvorlige bivirkninger, overvåkes kontinuerlig av komiteen.

Det skal bemerkes at inklusjonen i NMSG#12 går svært langsomt, sannsynligvis fordi pasientene vegrer seg for å takke ja til et medikament med relativt ubehagelige bivirkninger, samtidig som de får opplyst at de vil få thalidomid dersom sykdommen ikke skulle respondere på MP-behandling. (Inger M. Dahl, personlig meddelelse)

6) Denne protokollen kommer fra Nordic Myelodysplastic Syndrome (MDS) Group. Den kalles: **“Maintenance treatment with Decitabine vs. consolidation treatment with high-dose chemotherapy in patients with advanced MDS and MDS-AML, who have obtained CR with intensive chemotherapy”** (21)

Studien ble planlagt i 2002 , men alle formaliteter for oppstart var ikke klare før i 2005. Forventet slutt på innrulling er 24 måneder etter at studien startet, dvs sommeren 2007. Hele studien beregnes avsluttet 36 måneder etter studiestart. Det er beregnet å inkludere 120 pasienter per år slik at det totalt er med 240 pasienter i studien. I tillegg vil den franske MDS gruppen delta med pasienter.

*Bakgrunn:* Kurativ behandling av unge pasienter med MDS er stamcelletransplantasjon. Men de aller fleste pasientene er over 70 år og dermed ikke aktuelle for en slik behandling. Disse trenger en livsforlengende og livskvalitetshevende behandling. Det er færre som går i komplett remisjon med MDS enn med akutt myelogen leukemi (AML), trolig pga mange kromosomforandringer som indikerer dårlig prognose. Flere studier har vist at konvensjonell høy-dose kjemoterapi har gitt respons hos 40-60%. (22,23) Responstiden er også generelt kort. Til nå har standard behandling vært TAD-kurer, som består av Thioguanin+Cytosin arabinosid (Arabin®) og Danuorubicin, som induksjons og konsolideringsbehandling. I denne studien kuttet Thioguanin ut pga. leveringsproblem fra fabrikanten.

Det er kjent at ved MDS foregår det en DNA hypermetylering av det celle-syklus regulerende genet p15INK4B. En høy grad av hypermetylering er assosiert med en høyere frekvens av AML. Decitabine har vist at det induserer hypometylering av bla. p15INK4B. En fase 2 studie av Decitabine har vist en responsrate på ca 50 %. Den viste også at myelosuppresjon var vanlig og 7 % terapiinduserte dødsfall ble registrert. 20 % av pasientene kom i komplett remisjon.

Det MDS gruppen ønsker å undersøke med denne studien er om en kontinuerlig hypometylering kan gi lengre komplett remisjon. Slik gruppen ser det er det to mulige positive utfall av studien:

- 1) Decitabine gir forlenget tid i komplett remisjon i forhold til TAD.
- 2) Det er ingen signifikant forskjell mellom TAD og Decitabine, men Decitabine gir bedre livskvalitet og mindre behov for hospitalisering.

*Design:* Det primære endepunktet i studien er varigheten av komplett remisjon, det sekundære er overlevelse fra komplett remisjon (CR), livskvalitet fra CR og grad av hypermetylering under CR Dette er en åpen randomisert fase 3 studie. Pasientene randomiseres innen 1 uke etter at CRer oppnådd. Randomiseringen skjer ved sekretariatet i Huddinge. Den er stratifisert til Frankrike og de nordiske landene.

Studien er bygd opp slik at alle pasientene først gjennomgår en TAD-kur for å se om de kommer i CR. Alle som gjør det randomiseres til en av de to armene. I den konvensjonelle armen vil pasientene få konsolideringskurer med TAD, justert i forhold til sin helsetilstand og aktuelle

blodverdier. I den eksperimentelle armen vil de få lavdose Decitabine s.c. i sykluser på 6 uker over 24 måneder, 3 ganger i uka. Dosen justeres etter blodverdiene. Dette fortsetter de med så lenge de er i CR. Når pasientene i begge armene får tilbakefall er de ferdige med studien. Ved inklusjonstidspunktet tas beinmargsutstryk, kromosomanalyse, metyleringsanalyser og en rekke blodprøver som er angitt i protokollen. Det fylles også ut et livskvalitetsskjema type EORTC QLQ-C30. De samme prøvene tas ved CR. Livskvalitetsskjemaet fylles så ut 3,6 og 12 måneder etter CR. Beinmargsprøver tas som normalt ved MDS dvs 6,12 og 24 måneder etter CR. Ved avslutning av studien tas de samme prøvene som ved inklusjonen. Alle dødsfall og alvorlige bivirkninger meldes til sekretariatet.

*Inklusjonskriterier/eksklusjonskriterier:*

Inklusjonskriterier:

- Pasienter med MDS og > 10 % blaster + IPSS score på •INT2
- AML med foregående MDS

KMML

AML med foregående KMML

Eksklusjonskriterier:

- Pasienter med potensiell kurasjonsmulighet, som stamcelletransplantasjon
- Hjerte, lever eller nyresykdom som utelukker høydose kjemotreapi
- Pasienter som av en eller annen årsak ikke er i stand til å følge protokollen

*Administrasjon:* I denne studien vil sykehus fra Norge, Sverige, Danmark og Frankrike delta.

Den har et studieseekretariat ved Huddinge sykehus som bla. randomiserer pasientene. Det er også opprettet et eget senter Karolinska institutet som skal vurdere beinmargsutstrykene.

Kromosomanalyser og metyleringsanalyser gjøres i Aarhus.

Livskvalitetsanalysene gjøres på Ullevål sykehus.

Supergen, som produserer Decitabine, leverer medikamentet gratis til studien.

7) Den siste protokollen jeg vil gjennomgå ledes av Novartis, et kommersielt legemiddelfirma med base i Sveits. Studien kalles: ” **En fase 3 studie av STI571 versus Interferon- $\alpha$  kombinert med Cytarabin (Ara-C) til pasienter med nydiagnostisert tidligere ubehandlet**

**Philadelphia positivt kromosom kronisk myelogen leukemi i kronisk fase (KML-CP)."** Dette er en studie som inkluderte 1032 pasienter og disse følges nå i 5,25 år. Man regner med at studien varer i 5,5 år. Studien startet i første halvdel av 2000.

*Bakgrunn:* KML er en stamcelle sykdom hvor 95 % av pasientene er Philadelphia+ kromosom positive (25,26). Dette resulterer i at *abl* protooncogenet på kromosom 9 fusjonerer med *bcr* genet på kromosom 22. Det gir en aktivert form for *abl* protein-tyrosin kinase som bl.a fører til manglende apoptose. Sykdommen har tre distinkte faser med økende refraktærhet mot terapi. Den første fasen er den kroniske fasen som har en median varighet på 3-4 år. Deretter kommer den akselererte fasen med median varighet på 3-9 måneder og til slutt en blastkrise med median varighet på 3-6 måneder. Tallene for overlevelse er noe høyere enn varigheten av de ulike faser. Behandlingen av den kroniske fasen, dersom allogene stamcelletransplantasjon ikke er tilgjengelig, er hydroksyurea eller IFN- $\alpha$ . 70-80 % har fått komplett hematologisk respons (CHR) på sistnevnte. Senere er IFN blitt kombinert med Ara-C. Dette utgjør den konvensjonelle armen i studien. STI571, også kalt Glivec®, er en protein-tyrosin kinase hemmer. Den er relativ spesifikk for *bcr-abl* (27,28). Det ser ut til at medikamentet tolereres godt og kan tas p.o. Vanlige bivirkninger er kvalme, arthralgier, myalgi og periorbitalt ødem. GI-blødninger har vært observert. Medikamentet er mulig doserelatert levertoksisk. Dette er vist gjennom fase 1 og 2 studier av Novartis.

*Design:* I studien sammenliknes STI571 med IFN- $\alpha$  + Ara-C gitt til pasienter med KML i kronisk fase med Philadelphia+ kromosom. Det er en åpen randomisert multicenter studie. Det primære endepunktet er lengden av progresjonsfri overlevelse hos voksne pasienter. Sekundære endepunkter er livskvalitet, varighet- og antall komplette hematologiske responser(CHR) og major cytogenetisk respons(MCR), respons ved crossover, overlevelse, sikkerhet, toleranse og befolkningsfarmakokinetikk.

Studien er designet slik at pasienten randomiseres til 2 armer. Randomiseringen skjer hos Novartis. I den konvensjonelle armen får pasientene IFN- $\alpha$  og Ara-C, i den eksperimentelle armen 400 mg Glivec® daglig. Dosen er valgt fordi 300 mg i tidligere studier har vist at 100 % får CHR, men det er antatt at det trengs en doseøkning på 10 % pga ulike faktorer for å sikre samme resultat. Under studien er det ikke tillatt for pasientene å være med i utprøving av andre eksperimentelle medikamenter. Såkalt supportive care som transfusjoner er tillatt, men ikke cytostatika og vekstfaktorer, som G-CSF.

Pasientene undersøkes ved diagnosetidspunktet og så hver måned de neste 6 månedene. Resten av studien undersøkes pasienten hver 6 uke. Da tar man en rekke hematologiske og klinisk kjemiske prøver som er angitt i protokollen. De cytogenetiske analysene gjøres 1., 2. og 3. måned det første året, så hver 6. måned. På pasienter som får Glivec® utføres også farmakokinetiske prøver 1. og 29. dag. Livskvalitetsundersøkelsene gjøres hver måned første 6 måneder og så 6, 9, 12, 18 og 24 måneder etter inklusjon. I livskvalitetsundersøkelsene anvendes FACT-BRM, som ved mange cancerforsøk. Studiegruppen mener den er bedre enn EORTC-QLQ-Leu fordi den er rettet spesielt mot KML, og ikke AML som EORTC-skjemaet. FACT-BMR fokuserer på IFN- $\alpha$  toksisitet, funksjonell og kognitiv velvære. Undersøkelsen består av 40 utsagn som pasienten graderer fra at det ikke passer i det hele tatt og til at det passer veldig mye. Effektparametere beregnes ut fra død, akselerert eller blastfase, tap av CHR, tap av MCR eller økende antall leukocytter. Tap av MCR er når det påvises > 35 % Ph+ metafaser i beinmarg. Tap av CHR er når man taper normale verdier i perifert blod og normal leukocyt differensiering og når det er tegn til ekstramedullær sykdom. Ut fra dette mener studiegruppen at dersom pasienter i den eksperimentelle armen har en signifikant forlenget tid i MCR og CHR så er det bevis på at Glivec® er en bedre behandling enn IFN- $\alpha$ +ARA-C. Dette knyttes selvsagt opp mot livskvalitetsundersøkelsene.

#### *Inklusjonskriterier/eksklusjonskriterier:*

##### Inklusjonskriterier:

- Alder mellom 18 og 70 år
- inkluderes innen 6 måneder etter diagnosen KML er stillet
- tidligere ikke behandlet for KML
- < 15 % blaster i perifert blod og beinmarg, <30 % blaster + promyelocytter i perifert blod og beinmarg, 20 % basofile, >100\*10<sup>9</sup>/L trombocytter.
- Ikke ekstramedullær leukemi

##### Eksklusjonskriterier:

- Fått annen eksperimentell behandling
- Tidligere cytostatika eller G-CSF behandling
- Aktuell for allogen stamcelle behandling
- ECOG performance status >3
- Bilirubin, levertransferaser eller kreatinin > 1,5 x referanseverdien

- INR over 1,5 x referanseverdi
- Ukontrollert diabetes, hjertesykdom eller infeksjon
- HIV
- Gravid eller ammende
- Annen malignitet innen 5 år
- Ikke forventet å følge protokollen

Det er også mulig og krysse armene en gang dersom visse kriterier er oppfylt, såkalt crossover.

Disse er:

- tap av MCR eller CHR
- økende leukocytter
- intoleranse
- ikke oppnådd MCR ved 12 måneder
- ikke oppnådd CHR ved 12 måneder

*Administrasjon:* Studien styres av Novartis og har deltagere fra en rekke land i Europa og USA. Det er opprettet en interimkomité eksternt til Novartis som skal evaluere sikkerhet og effektivitetsdata, spesielt ved crossover. Komiteen gjør også interimanalyser, den første etter 12 måneder. Man regner med en drop out på ca 10 %. Alle laboratorieverdier som brukes skal leveres av et sertifisert laboratorium og godkjennes. Verdiene og undersøkelsene føres i en pasientspesifikk perm. Denne skal alltid befinne seg på et fast sted og ved faste intervaller kommer det en representant fra Novartis og har et møte med alle som skriver i permene for å se på fullstendigheten, føring av journalen, oppbevaring av medikamentene og om "good clinical practise" bedrives. Dette er såkalt "site visits". Ingen av medisinene eller tomme bokser skal kastes, men returneres til Novartis, som også har levert medisinen kostnadsfritt. Pasientene har underskrevet en omfattende informert samtykke på åtte sider.



### Nøkkelartikler

Det å utføre klinisk kontrollerte studier sammenfaller med ønske om "evidence-based medicine (EBM)". Antes et al (29) har i sin artikkel sitert DL.Sackett som har definert EBM som : " *The conscientious, explicit and judicious use of current best evidence in making decisions about care of individual patients. The practice of evidence based medicine means integrating individual clinical expertise with the best available external clinical evidence from systematic research.*"(30) Det Sackett mener er at vi i behandlingen av pasienter skal fatte beslutninger som er basert på objektiv forskning og integrere disse resultatene med klinisk ekspertise og individets spesielle behov. Antes et al skriver i sin artikkel at målet for EBM er å tette gapet mellom forskning og klinisk praksis slik at alle beslutninger fattes på det best mulig vitenskapelig grunnlag. Det metodologiske rammeverket for å gjøre dette er organisert i fire trinn:

- overføre et klinisk problem til spørsmål som kan besvares
- lete etter best mulige bevis
- kritisk tolke resultatene
- implementere resultatene i klinisk praksis

Spørsmålet vi stiller bør inneholde fire elementer: Et veldefinert problem , selve intervensjonen, en sammenliknende intervensjon og en analyse av klinisk virkning. Et kjennetegn ved et godt formulert spørsmål er at dette skal inneha statistiske testegenskaper. Det å finne beviset er ikke alltid like lett. Per i dag betyr det for mange et søk på Medline, som ikke nødvendigvis er den beste kilden til kunnskap. Indeksstrukturen er komplisert, mye er feilindeksert, ikke alle tidsskrifter er med i databasen osv. Tilsvarende problemer er det med Embase og Cochrane, forhold man skal være obs på når man skal lete i litteraturen. Antes et al tar også opp problemet med at ikke alt publiseres på engelsk og derved er utilgjengelig for mange. Engelskspråklige tidsskrifter er overrepresentert både i antall på biblioteker, når det gjelder anseelse og tro på at kvaliteten av disse er bedre. Det er også slik at de artiklene som trykkes i stor grad har positive resultater og resulterer i stor publikasjonsbias, noe som lager problemer ved f.eks. metaanalyser. Antes et al graderer bevisene etter hvor sikre man tror de er for at et svar skal være korrekt (tabell 1). Nivå 1a regnes for å være det sikreste . Når vi skal tolke resultatene er det viktig å være kritiske da en feiltolking kan få store konsekvenser. I tolkingen tar man hensyn til nivået av undersøkelsen samt dominerende faktorer som randomisering, blinding og drop outs. Selve tolkingen gjøres med statistiske tester som X-kvadrat, t-test og lignende. Et grunnleggende problem er at man ikke har noen sikre indikatorer på kvaliteten på de forsøkene som er gjort, selv

om det gjøres forsøk på standardisering med guidelines, som CONSORT. CONSORT brukes av mer enn 70 tidsskrifter for å få til en kvalitetssikring.

Et annet aspekt er at selv om resultat er korrekt er det ikke sikkert at det kan brukes klinisk. For å undersøke om resultatet er anvendelig gjøres noen statistiske beregninger, som relativ risikoreduksjon (RRR). RRR er den proporsjonale reduksjonen av risiko mellom intervensjons- og kontrollgruppen, mens absolutt risikoreduksjon (ARR) er differansen mellom risikoen for en uheldig virkning i intervensjons- og kontrollgruppen. Fra ARR kan vi også regne ut "number needed to treat" (NNT) som er det inverse av ARR. NNT viser oss hvor mange intervensjoner som skal til for å forhindre en uheldig hendelse. Eksempelvis vil NNT på 10 si oss at vi må behandle 10 personer med hypertensjon for å forhindre ett hjerneslag. NNT brukes mye i helseøkonomisk planlegging for å avgjøre hvorvidt man skal starte profylakse mot ulike tilstander som i seg selv ikke er skadelige, men som kan få uheldige konsekvenser på sikt.

**Sauerland et al** (31) hevder i sin artikkel at det foreligger en verdensomspennende bekymring blant politikere og helseøkonomer vedrørende hvordan et stadig mer komplisert helsevesenet skal finansieres. Helsekostnadene kommer sikkert til å øke, også i fremtiden. Som Sauerland et al hevder er det derfor ikke lenger mulig å skille medisin fra økonomi. Per i dag, hevder de, finnes ingen konsensus på hvilke pasienter som bør motta behandling eller hvilken behandlingsform som bør velges. Det er en ubehagelig oppgave å forklare en pasient at vedkommende ikke har krav på behandling. Sauerland et al mener at EBM kan være en metode for å validere data og velge riktig behandling. Dette blir spesielt viktig med tanke på den endringen som har skjedd i befolkningen. Pasienten har fått stadig mer innflytelse over egen behandling, noe som og i og for seg er ikke det galt. Men når dette gir seg utslag i legeshopping, søksmål, oppslag i dagspressen og lignende blir det et problem. Dette skjer til tross for at behandlende lege mener behandlingen er uforsvarlig eller unødvendig. EBM kan her være en metode for at pasienten får den beste behandlingen. EBM kan også være med på å fjerne unødvendig og endog skadelig behandling og derved spare penger og lidelse. I tider med knappe ressurser må også ressursene fordeles riktig. EBM kan bli et hjelpemiddel for dette ved at pasienter som trolig får mest helse ut av hver krone får en prioritet. Man snakker ofte om cost-benefit og kvalitets-justerte leveår (QALY). Som en forenkling vil det si at de yngste skal prioriteres når andre parametere er like. EBM er også med på å gi likeverdig behandling slik at ikke sykehus og lokale autoriteter blir avgjørende for og behandling og derved prognose.

Likeverdig behandling betyr også at ikke noen blir prioritert fordi legen, pårørende eller noen andre er flinkere til å snakke for seg.

**Feinstein** (32) tar i sin artikkel opp de begrensninger som finnes ved å bruke kliniske forsøk. Han tar spesielt opp to problemer, nemlig bruk av kliniske forsøk ved profylakseforskning, og hvordan man i studiedesignet kan ha to helt likeverdige og like sanne måter å designe et klinisk forsøk på. Sistnevnte betegner han "pragmatic" og "fastidious", som jeg har valgt og oversette til pragmatisk og idealistisk. En pragmatiker vil ta hensyn til heterogeniteten i befolkningen, noe den idealistiske vil unngå ved å skape en homogen undersøkelsespopulasjon. Sistnevnte oppnås ved f. eks ved å bare bruke ett kjønn eller en rase som bortsett fra testsykdommen er friske og meget samarbeidsvillige. Dette vil skape et forsøk som er lett å administrere og som gir "klare og entydige" svar. En pragmatiker vil imidlertid hevde at disse resultatene ikke vil passe inn i den kliniske hverdagen pga den heterogeniteten som normalt finnes. Et annet stridsspørsmål går på valget av etablert standardbehandling. Pragmatikeren vil da velge et etablert behandlingsregime som han tester mot det nye, mens idealisten vil velge placebo som kontrollbehandling. Ulempen med placebo vil være at det ikke viser forskjellen mellom det eksperimentelle og det etablerte regimet. Dette avspeiles også i dosering. En idealist vil, av frykt for å ødelegge en dobbelblinding, velge faste doseringer, mens en pragmatiker heller vil titrere doseringen. Dette gjelder spesielt medikamenter som antikoagulantia. Pragmatikeren vil med dette opplegget lettere finne effektiv dose og trolig ha lavere drop-out pga bivirkninger av medikamenter. Idealisten og pragmatikeren vil også ha forskjellige endepunkter. En idealist vil velge harde endepunkter som død, manifest sykdom og lignende. Det betyr at man må velge større studiepopulasjoner, spesielt om hendingen er sjelden. Fordelen vil være at resultatene er "rene" og lett å tolke. Pragmatikeren bruker ofte myke endepunkter i tråd med den hippokratiske ed, slik at dersom behandlingen ikke gir bedring eller trøst så er den verdiløs.

Det er også en fordel at man kan identifisere type pasienter som behandles. Behandlende lege vil da kunne knytte grad av sykdom, sykdomsprogresjon, co-morbiditet og lignende til den aktuelle pasientgruppe. Dette er forhold som idealisten ikke i samme grad tar hensyn til da dette er myke endepunkter. Idealisten baserer seg mer på randomisering. Problemer oppstår når man skal foreta generalisering av pasientgruppen som helhet.

Det er også forskjell på hvordan de to forskertyper gjør sine analyser. Den idealistiske randomiserer etter "Intention to treat"-prinsippet. Det betyr at man teller opp alle i gruppen uavhengig av om de fikk behandlingen eller ikke. Man tar altså med drop-out, pasienter som kanskje aldri rakk å starte behandling. En pragmatiker vil ikke ta disse med. Idealisten tar heller

ikke hensyn til compliance. Argumentet er at dette skjer etter randomisering og at separate analyser gir bias. Alle som er randomisert til en gruppe er med i studien, uansett om man senere skulle finne ut at de ikke tilfredsstillt kravene. En pragmatiker vil ekskludere disse personene. Dette viser seg også i analysen ved at idealisten med sine harde endepunkter vil ta med alle som dør av studiesykdommen, mens pragmatikeren vil ta med alle uansett dødsårsak.

Feinstein presiserer at ingen av de nevnte tilnæringsmåter er riktige eller gale. Begge grupper har sine argumenter. Det er mulig å tilfredsstille begge leire ved å bruke en heterogen studiepopulasjon i tillegg til den "rene" homogene gruppen, flere myke endepunkter etc.

Feinstein påpeker at de største kontroversen er kommet i forbindelse med randomiserte forsøk hvor man har sett på profylaktisk virkning. Diskusjonen har som regel dreid seg om behovet for placebo og påvisning av effektivitet ved kombinasjonsbehandlinger. Dette har gitt stor usikkerhet omkring de fleste profylaktiske behandlinger, f.eks antikoagulasjon etter hjerteinfarkt. Mye av dette kan ikke unngås pga nødvendig kompleksitet i studiedesignet samt konflikten mellom pragmatiker og idealist.

Et annet problem er at mange medikamenter har en kortsiktig eller kun profylaktisk effekt mens den langvarige effekten er mye vanskeligere å påvise og fordi lang oppfølging koster. Feinstein hevder derfor at om man skal vise en profylaktisk effekt må man bruke andre verktøy enn randomiserte forsøk. Han mener det vil være uetisk å sette inn noe man tror vil virke profylaktisk mot placebo. Man får også problemer med å vise at det nye regimet er like bra som nye medikamenter som ikke testes i studien. Feinstein nevner også kort noen andre problemer. Man kan for eksempel ikke teste det nye medikamentet mot alle de gamle medikamentene, slik at man ofte må velge ut ett. Studiene passer heller ikke for små forskjeller mellom de ulike fabrikantenes medikamenter, for eksempel acetylsalisylsyre. Randomiserte studier passer heller ikke for etiologiske studier og studier av diagnostisk teknologi. Man kan her påføre pasientene skade samt få problemer med å rekruttere en representativ studiepopulasjon.

**Bonchek** (33) tar i sin lederartikkel i Heart opp problemer og fallgruber i randomiserte forsøk der nye prosedyrer testes. Han hevder at nye medikamenter ikke kan få klinisk godkjenning eller godkjenning av myndighetene uten å ha blitt testet i randomiserte forsøk. Det har ledet til den forestillingen at dette også gjelder nye prosedyrer og teknikker. Det er forskjell mellom medikamenter og prosedyrer. Det skyldes at medikamenter ikke forandres og bivirkninger oppdages underveis. Medikamenter er også uavhengig av legens dyktighet. Placebo kan anvendes og crossover er akseptabelt. Når det gjelder prosedyrer er de som regel inkomplette når

de introduseres. Indikasjonen er usikker og risikoen er stor. Etter en viss tid avtar risikoen dramatisk og indikasjonene blir mer sikre pga at man skaffer seg erfaring.

Bonchek hevder videre i sin artikkel at randomisering bør eliminere bias, men ofte er ikke det mulig, spes ved studier av prosedyrer. Den første bias skjer før pasienten inkluderes i studien. Det er slik at når to ulike medikamenter eller prosedyrer sammenliknes har legen ofte ingen preferanse for en av dem og randomisering er som regel uproblematisk. Men når man sammenlikner medikament mot prosedyre blir det annerledes. Da unnlater man ofte å inkludere høyrisikopasienter i studien. Det er observert at pasienter i randomiserte studier gjør det bedre enn historiske kontroller. Dette skyldes trolig at pasientene får mer oppmerksomhet / tettere oppfølging samt en mer standardisert terapi. At man ikke vil inkludere pasienter kan også henge sammen med subjektive inklusjonskriterier. En tredje biasmulighet er cross-over spesielt om det skjer flere ganger. For prosedyrer er kun en crossover mulig, mens det for medikamenter er mulig å gjøre dette flere ganger. Man kan også fjerne de medikamentene som ikke virker og til slutt sitter man igjen med de som virker.

**Kaptchuk** (34) påstår i innledningen til sin artikkel at det dobbel-blindede randomiserte kontrollerte forsøk (RCT) prøver å etterlikne objektiviteten i laboratorieforsøkene. RCT regnes som den mest pålitelige metoden til å finne "sannheten". Man forsøker å minimere de systematiske feilene ved metodologisk forsikring ved randomisering og blinding. Randomisering skal eliminere, bevisst eller ubevisst, menneskelig innflytelse på fordelingen av studiepopulasjonen. Blinding skal forhindre at behandling og analyse ikke påvirkes av fordommer. På bakgrunn av dette påstår man at RCT er uten menneskelig bias og derfor er akseptabel som gullstandard.

Kaptchuk stiller spørsmålet ved om Heisenberg prinsippet også gjelder i studiet av mennesket fordi vi ikke kan omgå menneskets subjektivitet og ubevissthet. Det empiriske beviset for objektiviteten i RCT ligger i de differensierte utfallene som RCT finner sammenliknende med andre studiedesign, ofte kalt "measure of bias". Når man har målt en diskrepans mellom resultatene ulike studiedesign er dette blitt tatt for å være et bevis på RCT's objektivitet. Man har derfor regnet RCT som best. Dette kalles det "discrepancy argument", lansert av Harry Gold i 1950 som introduserte blindingen. I tre senere studier er det funnet at diskrepansen mellom randomiserte og ikke-randomiserte studier har flere potensielle årsaker, som for eksempel tilfeldigheter, rapporteringsbias, heterogenitet og preferanse. Blinding kan introdusere bias i RCT på flere måter. Det å vite at man kan få placebo kan forandre responsen uansett om de får placebo eller et aktivt stoff. Deltakelse i RCT kan også øke sensibiliteten og oppmerksomheten slik at så

vel deltaker og kliniker oftere rapporterer positive effekter. Dette er forhold som ikke kan randomiseres eller blindes bort og som underminerer RCT. For å eliminere deler av dette introduserer Kaptchuk en platinastandard, der klinikere og deltakere er uvitende om at de deltar i RCT. Av etiske grunner er ikke dette tilrådelig, men det er imidlertid blitt forsøkt i to franske studier (35,36) som viser at placeboeffekten er sterkere i den informerte armen av studien. Kaptchuk tar også opp det han kaller "investigator self-selection" I RCT randomiserer vi pasientene, men ikke forskerne. Det er klart at forskere ikke er typiske leger, men en viss gruppe leger som har en spesiell interesse. De kan påvirke pasientene på forskjellige måter med sine spesielle holdninger og oppførsel.

**Braunholtz et al** (37) diskuterer i sin artikkel om RCT har en positiv korttidseffekt på pasientene og om det finnes en "forsøkseffekt". Dette er en litteraturstudie av 21 artikler.

De begynner med å presisere hva de mener med forsøkseffekt. Det kan være mange typer forsøkseffekter for de som deltar i en studie, for de som nekter å delta, de som ekskluderes osv. Effekten kan være gunstig for en gruppe og skadelig for en annen, for eksempel de som deltar kan få et nytt bedre medikament, mens de som nekter ikke får det. Vi kan si at forsøket påvirker seg selv. Forsøk påvirker også klinikerens. Forsøkseffekten har flere komponenter. Den består av en behandlingseffekt, som skyldes selve medikamentet eller prosedyren i studien. Videre finnes protokolleffekten fordi man har faste og bestemte behandlingsopplegg for velkjente behandlinger. Det er påvist at det bedrer utkommet å følge slike faste protokoller. "Care" effekt kan være vanskelig å skille fra protokolleffekten, men skyldes forhold som bedre pleie og oppfølging. Hawthorne effekten skyldes forandringer i pasientens og klinikerens oppførsel fordi de er med i RCT. Placeboeffekten er en velkjent del av forsøkseffekten. Spørsmålet blir hvordan disse effektene påvirker pasientene i studiene. Alt som systematisk skiller forsøksgruppen fra kontrollgruppen kan gi bias, også forsøkseffekten. Graden av bias vil variere. Man har bekymringer for systematiske forskjeller mellom klinikere som rekrutterer og de som ikke gjør det. Skulle de som rekrutterte være bedre klinikere vil det kunne skape bias. Sammenlikning mellom grupper av pasienter som har samme kliniker vil ikke gi noen Hawthorne effekt og studieindusert bias.

Braunholtz et al finner i sin gjennomgang av artiklene at pasienter som deltar i kliniske studier har et signifikant bedre klinisk utkomme. De fant ingen artikler hvor det var et negativt klinisk utkomme, se Figur 1. De konkluderer med at det er lite evidence på området og mye stammer fra cancerforsøk. Effekten kan derfor være noe forskjellig ved andre sykdommer, for eksempel innen psykiatrien. Likevel ser det ut til at RCT gir flere fordeler enn skade, spesielt når den nye

behandlingen er bedre. Etske implikasjoner kan derfor bli at alle egnede pasienter tilbys inklusjon i studien slik at dette blir rutine vel vitende om de fordeler og ulemper studien medfører.

**Angell et al** (38) har i en editorial i *New England Journal of Medicine* noen interessante poenger når det gjelder om hvorvidt tolking og formidling av resultatene påvirker vår oppfatning av en studie. Et meget godt poeng er at det som publiseres i de medisinske tidsskriftene ikke er visdom, men arbeidsskisser. En artikkel skal kommunisere med andre om resultatet av en studie. En studie blir derved bare del av det som bekrefter eller avkrefter en hypotese. Det er sjelden at en studie alene er nok. Dette skyldes bla mange uunngåelige bias, confounding faktorer med mer. Selv om slike faktorer er ekskludert kan resultatet være et slumptreff. Det er heller ingen automatikk i at et resultat kan overføres mellom ulike populasjoner. Som regel er leger klare over disse forholdene slik at de er motvillige til å endre sin praksis over natten, noe pasientene ikke alltid skjønner.

**Grover og Shoyer** (39) har definert klinisk forskning som den disiplin der leger og andre profesjonsgrupper anvender kunnskap fra basalforskningen til å utvikle preventiver, intervensjoner og for å bedre utkommet av sykdom for pasienter. Det fokuseres på utvikling av nye eller videreutvikling av eldre metoder for diagnostisering eller behandling. Forfatterne påpeker at forskningen fokuserer på effektivitet, kostnad-nytte effekt og utkomme for pasienten. Typiske endepunkter i studier er kostnad-nytte effekt, klinisk resultat (mortalitet og morbiditet), quality of life og tilfredshet. En av fordelene med klinisk forskning er muligheten for å studere kliniske spørsmål på en organisert måte. Dette krever ikke store laboratorier og lignende og kan utføres uten store økonomisk utlegg. Forsøkene kan gi svar på viktige hverdagslige kliniske spørsmål relatert til behandling, diagnostisering og helsepolitisk. Grover og Shoyer tar også for seg den generelle forskningsmetoden med hypotesedannelse, datainnsamling og hypotesetesting. Interessante poenger her er hvordan man bør definere de ulike rollene til det personell som deltar, hvem som skal delta og ansvarsområder. De tar også opp temaet interrimanalyser som en viktig kvalitetssikring for pasientene.

**Edwards et al** (40) tar for seg etikken i randomiserte kontrollerte studier. Informert samtykke er vidt og stort sett universelt akseptert. Forfatterne har imidlertid funnet to artikler (41,42) hvor konklusjonen er at en av fem leger regelmessig har meldt på pasienter i studier uten at informert samtykke er innhentet. Tre studier som refereres i samme artikkel har konkludert med at 47% av

legene i en multisenter studie mener at få pasienter visste at de var med i et forsøk, selv om skriftlig informert samtykke var innhentet. På tross av dette viser studier (43,44) at pasienter generelt er fornøyd med informasjonen de får. De føler også at de har gjort autonome valg. Mange av legene er bekymret over at de overlesser pasienter med informasjon og derved gir dem unødig engstelse og at det til og med kan føles slemt.

I artikkelen diskuteres også pasientenes motivasjon for å delta i RCT ut i fra et etisk standpunkt. Pasientens vanligste grunn til å delta beror på at de tror at dette er noe som vil gagne dem. Det er da troen på at en eventuell ny behandling kan hjelpe dem som er viktigst, ikke den økt oppfølgingen som et RCT automatisk vil medføre.

Et etisk problem er når en potensiell livreddende behandling ikke finnes etablert men at en mulig kurativ behandling er under utprøving. Skal man da kun gi behandlingen til dem som er med i studien, eller skal man også gi det til de pasienter som ikke deltar. I følge to studier (45,46) mener 79 % av legene at det skal man gjøre. Minouge (47) mener at behandling skal tilby også de utenfor studier, ellers kan det tvinge pasientene inn i studien.

**Parmar** (48) tar i sin artikkel for seg de problemene som finnes med tolking og rapportering av kliniske forsøk. Dette er forhold som lett vil påvirke metaanalyser, men også hvordan vi oppfatter resultatet av et klinisk forsøk og hvor troverdig dette virker.

Parmar hevder at målet med kliniske forsøk å besvare terapeutiske spørsmål på en unbiased og troverdig måte. Men selv om forsøket er designet slik at dette burde være mulig, er det en rekke problemer knyttet til tolking og rapportering.

Et av disse problemene er knyttet til randomisering. Vi randomiserer fordi vi ikke kjenner alle faktorer som påvirker utfallet av en sykdom. Når vi sammenlikner to grupper må vi derfor avgjøre om forskjellen mellom dem er pga ulik behandling eller andre ulike faktorene. Noen av dem kjenner vi kanskje ikke. Dette blir spesielt viktig i studier der vi ikke kan forvente mer enn en moderat effekt fordi observert forskjell kan skyldes ukjente prognostiske faktorer.

Randomisering har som hensikt å fordele alle prognostiske faktorer, unntatt den vi undersøker, likt i begge grupper. Ut fra dette har man konstruert en verdi, p-verdien, som sier oss at den observerte forskjellen med en sannsynlighet  $x$  skyldes maldistribsjon av de ukjente faktorene i de to gruppene. Dermed mener Parmar at man bør være forsiktig med å implisere ny terapi og behandling på bakgrunn av non-randomiserte studier, selv om resultatene er gode.

Det fokuseres mye på p-verdien i kliniske forsøk. P-verdien er en testverdi for hypotesen. Den gir ingen informasjon om behandlingseffekt, kun pålitelighet. Parmar mener derfor at p-verdien alene, eller overvektning av denne, kan gi et inntrykk av at den er viktigere enn den egentlig er.



Man kan ikke sette kliniske signifikans lik statistisk signifikans. Statistisk signifikans er når p-verdien er under 0,05. Klinisk signifikans er avhengig av behandlingseffekt, toksisitet og kostnader. Den forventede p-verdien er relatert til antall pasienter i studien. Jo flere pasienter desto mindre blir p-verdien. Dette gjelder når relativ risiko er lik. Størrelsen påvirker også konfidensintervallet ved at det blir smalere når størrelsen øker. Se tabell 2.

Det snakkes ofte om positive og negative forsøk. Negativt er det forsøket der det ikke finnes en statistisk signifikant forskjell. Men ut i fra resonnementet omkring p-verdien ser vi at små forsøk ofte er inkonklusive i stedet for negative. Et positivt forsøk er de forsøkene der vi påviser en statistisk signifikant forskjell mellom to behandlingsarmer. Men tolkningen av positive resultater har vært viet liten oppmerksomhet. For eksempel har man ikke diskutert hvor positive resultatene er før man konkluderer med at behandlingen er effektiv og inkluderer denne i klinisk praksis. Det inkluderer også elementer som fordommer hos klinikeren mot forsøket, representativiteten blant forsøkspersonene, konfidensintervall osv. Det er også slik at klinikere tolker de positive resultatene forskjellig. Parmar mener derfor at alle anbefalinger som stammer fra et forsøk bør diskuteres ut i fra de forskjellige måtene å tolke dem på.

Parmar tar også for seg problemer med å stoppe et klinisk forsøk før tiden. Ettersom mange forsøk kan ta tid, vil man få data underveis. Skulle disse vise en markert forskjell i behandlingsresultatet, både i positiv og negativ retning, må forsøket stoppes. Dette er en beslutning som krever nøye overveielser. Man risikerer et falskt positivt resultat (Type I feil). Dette gir også bias i estimatet for effekt, som uriktig forsterkes, samt et konfidensintervall som er for smalt. Selv når man vet dette blir resultater publisert. Tar man hensyn til denne effekten er denne trolig overestimert. Problemet er større når man ikke tar noen hensyn.

**Hjort et al** (49) har i sin artikkel tatt for seg legers holdninger til kliniske forsøk og hvordan disse påvirker påmelding til kliniske forsøk i NMSGs regi. Å få nok deltagende pasienter til et forsøk er den store bøygen. Ved lave innmeldingsrater risikerer man også en skjevfordeling av pasienter. Hjort et al har i sin studie tatt for seg de 107 sentrene som er med i NMSG forsøk # 4/90. Det ble sendt ut et spørreskjema til legene ved disse sentrene som inneholdt 66 spørsmål, hvorav 32 gikk på generelle holdninger som kunne påvirke innmeldingsraten. Hjort et al fant i sin studie at 72 % av de beregnete nydiagnostiserte ble rapportert til studien og at 94 % av dem ble inkludert i studien. Legenes hovedbegrunnelse for å inkludere pasienter i studien var å bidra til medisinske fremskritt, deretter læringsfordeler. Dette er også vist i andre studier(50,51). De økonomiske aspektene kom langt ned på lista. Det var viktig for legene at studieprotokollen var enkel og meningsfylt. Dette er trolig den viktigste faktoren for å øke inklusjonsraten. Også

kommunikasjonen mellom ledergruppen og de deltagende sykehus er viktig. Her er en klar og lettfattelig studieprotokoll, følelse av deltakelse i studien og møter med ledergruppa viktige elementer. Hjort et al mener derfor at studien må legges opp slik at disse forholdene ivaretas for å få en best mulig inklusjonsrate. Viktig er også de etiske aspektene. Mindre viktig er arbeidsmengde og deltakelse i beslutningsprosessen. Minst viktig er de økonomiske aspektene. NMSG #4/90 hadde en høyere inklusjonsrate enn det som er vanlig, selv om det er store forskjeller mellom de ulike sykehusene. Hjort et al presiserer imidlertid at den fundamentale forskjellen mellom Skandinavia og USA er at man i Skandinavia har statlige ansatte leger, mens man i USA har et mer privat sykehussystem. Dette er en faktor som kan spille inn på inklusjonsraten ved at man i USA har mindre tid til forskning i sykehus.

**Hjort et al** (52) har også skrevet en artikkel som omhandler inklusjonsraten og kvaliteten ved små og store sentra som deltar i en studie. Artikkelen tar utgangspunkt i NMSGs protokoll #4/90. Med kvalitet menes her inklusjonsrate, oppfølging av pasientene og doseintensiteten av forsøksmedisinen. I #4/90 deltok sykehus av forskjellig størrelse. Hjort et al viser at det ikke er noen signifikant forskjell i påmeldingsraten mellom små og store sykehus. Et lite sykehus er definert som et sykehus med et nedslagsfelt på under 100.000. Små sykehus hadde en inklusjonsrate på 40 % mot 38 % ved universitetssykehus. Når det gjelder studiepopulasjonen er alderen litt lavere på universitetssykehusene, ellers var det ingen forskjeller mellom store og små sykehus. Drop outs var også relativt likt mellom de ulike sykehustypene.

I sin diskusjon mener Hjort et al at man ikke direkte kan ekstrapolere disse resultatene fra hematologisk cancerbehandling til generell klinisk praksis, pga de ulike seleksjonsmekanismene som finnes. De anfører videre på at det finnes artikler som peker på at mindre sentra har en lavere kvalitet. Hjort et al mener også at det er viktig at mindre sentre er med i studier da det vil øke representativiteten i studiepopulasjonen, at man raskere får det nødvendige antallet pasienter i studien og at det har en lærings- og vitenskapelig effekt på de deltagende sentrene. De konkluderer med at kvaliteten er lik slik de har definert den. De spekulerer videre på hvorfor det ikke ble vist noen forskjell og mener at flere faktorer har spilt inn, som en aktiv holdning ved de sykehus som deltok i studien. Videre har nok en desentralisering av ansvar fra et sentralt sekretariat til ni regionssentra, slik at informasjon og kommunikasjon har gått lettere, spilt en rolle. De administrative belastningene er også forsøkt holdt på et minimum. Hjort et al konkluderer derfor med at små sykehus kan gi et viktig og kvalitativt godt pasientmateriale til kliniske studier, uten at studien settes i fare.

**Kerridge et al** (53) har skrevet en artikkel om EBM og etikken rundt den. Som de sier er EBM grunnlagt på idealet om at beslutninger i behandlingen av enkeltindividet bør bygge på bruk av beste tilgjengelige bevis. De hevder EBM har økt klinikerens generelle kunnskap, leseevner, datakunnskap, lager et rammeverk for læring, gitt bedre kommunikasjon med pasienter og mer effektiv bruk av ressurser. Fra et etiske standpunkt er det sterkeste argumentet for EBM at pasient og doktor kan gjøre informerte beslutninger, og at man anvender de best evaluerte helsemetodene/behandlinger. Men medisinske beslutninger fattes på bakgrunn av EBM, personlig erfaring, personlig bias, personlige verdier og politiske og økonomiske overveielser. EBM er et praktisk eksempel på det som i filosofien kalles konsekvensialismen, som sier at verdien av en handling måles mot handlingens konsekvenser. Kritikken mot konsekvensialismen, som også blir en konsekvens mot EBM, er tredelt. For det første kan ikke alle resultater måles eller defineres. For det andre er det uklart hvem sin interesser som skal telle i utkommet, og for det tredje kan konsekvensialismen lede til konklusjoner som er uetiske fra andre synspunkter. Førstnevnte kritikk skyldes at EBM hevder å gi en enkel, logisk prosess for å ta avgjørelser. Men for å ta balanserte avgjørelser må alle relevante konsekvenser vurderes. I medisinen er ikke alle utkomme klare, f.eks. grad av smerte. Noen kan ikke måles og noen er mer komplekse, som livskvalitet. Den andre kritikk begrunnes med at i EBM har pasientene fortsatt liten innflytelse over hva som skal prioriteres innen medisinen. EBM forkaster ekspertens mening, men det er fremdels eksperten som avgjør hva som skal forskes på, tolker data og implementerer resultatene. Mange ønsker mer brukermedvirkning i forskningsagendaen. Vedrørende det tredje kritiske punkt kan EBM lede til aktiviteter som trosser vanlig moral siden man kun vurderer intervensjoner ut i fra bevis for effektivitet. Et eksempel på dette er Nazi-eksperimentene i konsentrasjonsleire under andre verdenskrig. Ut fra konsekvensialismen var dette bra siden man har fått viktige forskningsresultater, men måten det er gjort på et totalt uetisk. Dett er også resultater som brukes i metaanalyser og nå har vi et problem med hvordan vi skal håndtere dette. RCT er bra som grunnlag for kliniske beslutninger, men det har også sin verdi ved prioritering av ressurser innenfor helsevesenet. Dette appellerer også til politikerne. Men det er også mulighet for bias fordi det kreves store mengder data for å tilfredstille standardene for EBM, og det finnes kun noen få behandlinger som har dette. Dermed vil ressursene rettes mot disse, og det blir mindre til de delene av medisinen som ikke har evidence eller hvor dette ikke er mulig å oppnå, f.eks. i palliativ medisin. Man må derfor gjøre avveininger. Kerridge et al hevder at EBM konsentrerer seg om effekten på individet. Leger kan ikke bare konsentrere seg om individet, men også samfunnets helse. Av og til blir det konflikt mellom disse interessene, og EBM er ikke midlet til å løse denne konflikten.

**Campbell** hevder i sin artikkel at klinikerer også må forske(54). Det begrunner han med at selv om klinikerer og forskerer samarbeider godt, så trengs "intimiteten i et sinn" for å komme fram til nye revolusjonerende tanker. Klinisk forskning har gitt flere viktige bidrag til medisinen. Det har vært direkte fordeler, som forbedring av kliniske teknikker som hjerteangiografi. Det har også vært indirekte fordeler, som bidrag til biologisk forskning. Mange biologiske fag er opprettet av klinikerer, som fysiologi. Også medisinsk utdannelse har dratt veksler på klinisk forskning. Klinisk forskning har bidratt til å forklare hvordan kroppen fungerer på en litt annen måte. Man har ellers ofte en litt mer praktisk tilnærming. Gjennom klinikken har man også funnet kliniske spørsmål og sett sykdomsmønstre. Klinisk forskning har også bidratt til vår forståelse av hvordan organer og organsystemet fungerer funksjonelt.

Klinisk forskning har også opplevd problemer og fått kritikk. Fra et etisk standpunkt har det vært hevdet at klinikerer bruker pasienten som forsøkskanin. Som et vern mot dette har vi Helsinki deklarasjonen. Campbell er imidlertid mer urolig for skadene prematur bruk av forskningsresultater og behandlinger kan påføre pasientene. En annen kritikk går ut på at de fleste kliniske forsøk konsentrerer seg om endestadier av sykdom. Man har også hatt mest fokus på de alvorligste sykdommene og glemt de mindre alvorlige, selv om de kanskje medfører et like stort helsetap. Man kan også lett få triviell forskning. En siste kritikk har vært at grensen mellom forskning og praksis er vag og at man kan fristes til å implementere sin egne resultater i praksis. Campbell gjør et poeng ut av at det er et slags hierarki innen medisinsk forskning, hvor det som står på toppen er basalfagene. Og jo mer basal forskningen er, jo bedre. I en slik sammenheng taper den kliniske forskningen. Dette er Campbell i mot. Han mener den kliniske forskningen skal fortsette fordi det er det rette å gjøre. Ennå er det mange kliniske spørsmål som må besvares som f. eks årsak til hoste og diaré som vi har overraskende lite forskning på. Campbell ser for seg i fremtiden at man vil samarbeide mer med basalfagene og epidemiologene. Klinikerer blir da et bindeledd mellom praktisk medisin og basalfagene.

Helsevesenet har ikke ubegrensede resurser og derfor må krav til utnyttelse av disse skjerpes. Prioritering av pasientbehandling er vanskelig, men RCT kan hjelpe oss med dette. I helsedirektoratets utredningsserie IK-2381: "**Benmargstransplantasjon i Norge**"(55) og NOU 1997:18 "**Prioritering på ny**"(56) er det i offentlig regi laget noen prinsipper for hvordan man skal prioritere i det offentlige helsevesenet i Norge.

I NOU 1997:18 skiller man mellom første- og annenordensbeslutninger. Dette gjelder prioriteringsbeslutninger. Førsteordensbeslutninger er politiske beslutninger som kan skje på statlig, regionalt eller lokalt nivå. Dette er beslutninger/prioriteringer som ofte dreier seg om budsjett, bemanning og kapasitet. Det finnes for eksempel begrensinger ved organdonasjon, der man har en begrenset mengde organer til disposisjon. Annenordensbeslutninger bestemmer hvem som skal få disse godene og er derved kliniske beslutninger. Det dreier seg altså om individuelle tildelinger. I tillegg er disse også overordnet ved at det lages faglige standarder for diagnostikk, utredning og behandling. Første- og annenordensbeslutninger er derfor ikke helt uavhengig av hverandre.

Figur 2 illustrerer behovet for å prioritere. Vi har et helsegap som kommer til å bli enda mer tydelig i tiden framover. Man regner med 20 % flere krefttilfeller de nærmeste år fordi vi blir eldre og utsettes for økt eksponisjon av risikofaktorer. Dette vil øke behovet for prioriteringer ved at kravet til utnyttelse av ressursene skjerpes.

NOU 1997:18 slår fast at man på mange områder innen medisinen mangler dokumentasjon på helseeffekt. Det er derfor ønskelig med mer klinisk forskning for bla. å vurdere kostnad-nytte og kostnads-effektivisering som brukes i prioriteringer.

Utvalget legger opp til at man skal prioritere etter alvorlighetsgrad av sykdom og forventet nytte av behandlingen. I alvorlighetsvurderingen skal man vurdere risiko for død eller funksjonstap, grad av fysisk og psykisk funksjonstap, smerter og ubehag. Man skal også vektlegge nytten av behandlingen mer enn man har gjort tidligere. Nyttene skal dokumenteres ved hjelp av studier.

Utvalget graderer også graden av sikker dokumentasjon etter type studiedesign slik som Antes et al. Se figur 3

Utvalget har fastsatt en nedre grense for tiltak som anses for nyttige. For helbredende behandling skal man oppnå 5-10 % bedret langtidsoverlevelse i forhold til ingen behandling. For livsforlengende behandling skal mediane overlevelse forlenges med 20 %, minimum tre måneder.

Utvalget legger videre opp til likebehandling når man prioriterer. Det betyr at dersom en pasient får et tilbud, så skal alle i samme situasjon få et likeverdig tilbud. Med situasjon menes alvorlighetsgrad, nytte av tiltak og kostnadseffektivitet. Alle med høyere prioritet skal også få et likeverdig tilbud.

Prioriteringseffekten foreslås i tre trinn:

- Faggruppene anbefaler klassifisering og sortering av tilstanden og tiltak innen sitt fagfelt.
- Politiske og administrative beslutningstakere treffer beslutning om kapasitet og ressurser, det vil si at de prioriterer gruppene.

- Faggruppene formulerer kliniske retningslinjer for valg av pasienter til riktig behandling. Utvalget anbefaler dermed at faggrupper innen de ulike feltene, for eksempel kreftomsorg, opprettes. De anbefaler at både allmennlegetjenesten og spesialisthelsetjenesten er representert i tillegg til tilsynsmyndigheter og evt. fagspesifikke grupper som fysioterapeut og logoped i slaggruppe.

Utvalget foreslår en inndeling av prioritetsgrupper:

I Grunnleggende helsetjenester

II Utfyllende tjenester som screening og utredning

III Lavt prioriterte

IV Tilbud som ikke hører hjemme innenfor offentlig finansiert helsetilbud.

Grunnleggende tjenester er å tilfredsstille grunnleggende behov som pleie, omsorg, palliasjon og livreddende behandling. Dette skal ha høy prioritet. Utprøvende behandling kommer i en særstilling, og det innser utvalget. Det bør derfor skilles ut fra denne prioriteringslisten.

Utprøvende behandling har også fordelingspolitisk relevans ved at man er med på å framskaffe dokumentasjon på effekt.

I helsedirektoratets utredningsserier, IK-2381: "Benmargstransplantasjon i Norge" er det et eget kapittel om prioritering. Rådet setter opp tre spørsmål som bør besvares, når man skal prioritere pasienter: Er transplantasjon er *eneste* sjanse til å redde livet, hvor stor er nytten og hvilke lidelser påføres pasienten. Rådet legger opp til fire prioriteringsgrupper for benmargstransplantasjon (på bakgrunn av Lønningutvalgets tidligere arbeid):

I: Indikasjoner der transplantasjon er eneste mulighet for helbredelse og man har god sjanse for å lykkes.

II: Indikasjoner der transplantasjon gir beste mulighet for helbredelse med god livskvalitet, og anbefales.

III: Indikasjoner der vi ikke kjenner behandlingens verdi, men hvor behandlingen kan utføres for å avklare nytteverdi.

IV: Indikasjoner der dagens teknikk ikke har vist seg å være til nytte.

En utfordring blir å avgjøre hvor de ulike behandlingalternativene hører til. Vi ser at de kliniske forsøkene hører hjemme i gruppe III.

## Diskusjon

Over 100 år er gått siden man gjennomførte det første kontrollerte kliniske forsøket (57), og siden den gang har det vært gjort store fremskritt innen det medisinske fagfelt. Det stilles stadig større krav til dokumentasjon på alle områder. Bakketeig og Magnus har i sin temaartikkel i Tidsskriftet (58) gitt en kort oversikt over hva en forskningsprotokoll skal besvare og forslag til disposisjon av denne. Se fig. 4.

Ideen om kliniske forsøk skyldes ønsket om evidence-based medisin (EBM), hvor all utredning og behandling skal uføres på et vitenskapelig grunnlag. Det kontrollerte randomiserte forsøket (RCT) søker å etterligne de kontrollerte betingelsene som finnes ved laboratorieforskning.

### Høyeste bevisnivå

Antes et al har i sin artikkel laget en tabell ( tabell 1), der de ut i fra studiedesign graderer de bevisene vi oppnår i form av forskningsresultater (29). Der settes en metaanalyse av randomiserte kontrollerte forsøk på nivå 1, dvs det er de svarene man er mest sikker på er riktige. En slik gradering kan nok være riktig, men bør ikke vektlegges for sterkt. Det er ikke slik at konsensusrapporter, klinisk erfaring og sammenlikningsstudier er verdiløse selv om de stilles nederst i graderingen. En slik studiedesign gir riktignok ikke den største korrekthet, men den er med på å danne et grunnlag for videre studier på et høyere nivå. Man stiller da et klinisk spørsmål på bakgrunn av et mønster man har oppdaget i sin kliniske hverdag og fått verifisert gjennom etablerte metoder. Resultatene av slike studier kommer raskere og kan gi en betydelig forandring av kliniske beslutninger, uten at dette er galt. Faktisk er det en viss sannsynlighet for at det klinikere erfarer empirisk er riktig. Mye av årsaken til at RCT regnes for å ha det høyeste bevisnivået ligger i det som Kaptchuk påstår (34), nemlig at man forsøker å etterligne objektiviteten i laboratorieforskningene. Videre forsøker man å minimere de systematiske feilene ved å randomisere og blinde forsøkene. Ut fra dette har man akseptert RCT som en slags gullstandard. Det er en stilltiende enighet om at RCT er helt uten menneskelig bias, men dette stemmer ikke helt. Det er et kjent faktum at vi mennesker ofte forandrer oppførsel og handlinger etter hva som er sosialt akseptabelt, uten at dette alltid skjer bevisst. Trolig finnes Heisenberg-prinsippet også i studiet av mennesket slik som Kaptchuk hevder (34). Bare kunnskapen om at man er med i en studie vil øke sensitiviteten hos pasient og lege for symptomer, bivirkninger og

lignende. Pasientene vil kjenne etter symptomene og trolig rapportere de negative sidene i mindre- og de positive i større grad i den tro at de mottar den nye behandlingen. Legene vil på sin side vektlegge symptomer og graderinger, for eksempel funksjonsgraderinger, mer enn ellers. Symptomer man vanligvis ikke ville ha vektlagt får nå mye større betydning eller de ignoreres totalt. Man kan si at man lettere rapporterer positive og underslår negative effekter. Dette kalles rapporteringsbias og er ikke et element man kan blinde eller randomisere bort. For å få bukt med dette har Kaptchuk lansert ideen om en platinastandard hvor både kliniker og pasient er uvitende om at de deltar i en studie (34). To franske studier (35,36) har vist at placeboeffekten er større i den armen der pasient og kliniker har fått informasjon. Slik Kaptchuk beskriver platinastandarden er den imidlertid et klart brudd på Helsinkideklarasjonen der det heter at deltakelse i en klinisk studie krever et informert samtykke fra pasienten. Dermed skulle platinastandarden være umulig å gjennomføre uten å bryte et viktig prinsipp for pasientenes autonomi. Hvordan man skal gjennomføre dette praktisk er vanskelig å forstå. Hvordan skal man distribuere forsøksmedisinen og hvordan skal man få gjennomført den tette oppfølgingen som er nødvendig? Litt ironisk kan man jo spørre om hvorfor man trenger en platinastandard da en av fem leger melder pasienter til kliniske studier uten at informert samtykke er innhentet. (41,42) og 57 % av pasientene ikke visste at de var med i en studie, selv om skriftlig informert samtykke var innhentet. I forlengelsen av tanken om en platinastandard snakker Kaptchuk om "investigator self-selection" (34), dvs. at de legene som forsker er en gruppe som har en spesiell interesse og slik er blitt selektert ut. Det kan påvirke pasientene ved at de har en spesiell holdning og oppførsel. Dette kan unngås ved at flere klinikere får interesse og mulighet til å forske med for eksempel deltakelse i kliniske studier. Dette er gunstig for klinikerens ( og pasienten) ved at, som Campbell hevder, må holde seg mer oppdatert innen sitt fagfelt (54). Forskning har også en læringseffekt for klinikerens ved at man må gå dypere inn i materien og lærer seg å tolke vitenskapelige artikler på en god kritisk måte. Dette er evner som på sikt vi gagnar alle pasienter.

### *Bruk av metaanalyser*

Antes et al setter en metaanalyse av RCT som det høyeste bevisnivået (29). Dette er ikke helt problemfritt. Det er riktig at RCT gir den høyeste graden av bevis, men det er noen feller man kan gå i. Hvor skal man hente artiklene og hvilke kritiske krav skal man ha til artiklene? Det er i dag et stort antall tidsskrifter innenfor alle medisinske fagfelt. Mange av disse er tilgjengelige, ofte via Internett. Men det er slik at artikler som ikke finnes på engelsk eller som ikke publiseres



i de mest anerkjente tidsskriftene, som NEJM, ikke får den oppmerksomheten de fortjener. Dersom artikler ikke publiseres i anerkjente tidsskrifter er de nærmest verdiløse. Man har også den forestilling at artikler publisert i tidsskrifter med høy "impact factor" er av høyere kvalitet, noen som ikke behøver å være tilfelle. Et ytterligere problem i artikkelsøkingen er hvilke søkemotorer som brukes, og problemer som knyttes til disse. Et søk på Medline eller Embase er nødvendigvis ikke det beste. De har en komplisert indeksstruktur og mange artikler er feilindekserte (29). Det er også velkjent at det eksisterer en publiksjonsbias, dvs at positive resultater oftere publiseres og får mer betydning enn negative. Fordi et negativt resultat av en studie er like viktig som et positivt, kan metaanalyser være villedende da vi derved ikke får kunnskap om hva som ikke virker. Man gjør ikke alltid en avveining av hvor gode de artiklene er som anvendes i metaanalyser. Spesielt blir små studier, som ofte er positive, vektlagt like mye som større. Dette får en betydning for p-verdi og konfidensintervall, noe Parmar diskuterer i sin artikkel(48). Til forsvar for metaanalysen har Angell et al (38) et godt poeng når de sier at en studie ikke er visdom, men en arbeidskisse som sammen med flere skisser er med på å avkrefte eller bekrefte en hypotese. Det finnes få hypoteser innen medisinen hvor kun en studie har forandret den medisinske tankegang og behandling. Dette er en fordel fordi opp gjennom historien har det vært mange store veldeignede studier som senere har vist seg å ha gitt et resultat som ikke kan reproduseres. Dette kan skyldes unngåelige bias, confounding eller rett og slett en tilfeldighet (38) Det er heller ingen automatikk i at resultatene fra en populasjon, som i studiesammenhenge er homogen og "renset", kan overføres til den generelle befolkningen. Vi har alle hørt om den hvite, middelaldrende mannen som dominerende studieobjekt. Ut fra disse synspunktene er en metanalyse et godt hjelpemiddel for å lage en syntese av kunnskap innen et fagfelt. Det krever imidlertid et man har de nevnte mulighetene for feilkilder i mente.

I Nylenna-utvalgets forslag til ny lov vedrørende regler for forskning, NOU-2005 "God forskning - bedre helse", foreslås bla. at Helsinkideklarasjonen lovfestes og at det å publisere resultater ikke bare er en rett men også en plikt (Nylenna, personlig meddelelse). Dette kan motvirke publikasjonsbias.

### *Hva med prosedyrer og operasjoner ?*

Det viser seg at RCT ikke er like velegnet når det gjelder forskning på prosedyrer og operasjoner. (33). Færre prosedyrer (11-24 %) er dokumentert gode i RCT enn i generell

medisinsk klinisk forskning (31 –53 %) (tabell 3). Resultatene spriker, men andelen av overbevisende ikke-eksperimentelle bevis er større for prosedyre-forskning. Den lave andelen av RCT- bevis innen prosedyrer kan lett bli et problem dersom politikere krever økt bevisførsel før det bevilges økonomisk støtte. Bonchek hevder i sin artikkel at det er fundamentale forskjeller mellom medikamenter og prosedyrer når det gjelder RCT (33). Medikamenter er uavhengige av legenes dyktighet, man har placebo og crossover er akseptabelt. Disse fordelene finnes ikke ved prosedyreforskning, som har en læringseffekt knyttet til seg ved at operatøren blir bedre for hvert inngrep pga erfaring. Ut fra dette synspunktet kan det være fornuftig ved RCT for prosedyrer å kreve et visst volum per operatør samt forlange at operatørene bør ha god erfaring innen lignende prosedyrer. Ellers risikerer man å forkaste nyttige prosedyrer fordi operatøren mangler erfaringen. Bonchek hevder videre i sin artikkel at randomisering og blinding er vanskelig med tanke på prosedyrer (33). Det vil være vanskelig å blinde ved å gjøre en "dummy" operasjon hvor man bare legger et snitt i pasienten uten å gjøre noe annet. Dette vil kunne være en risiko for pasienten med bla. infeksjonsfare. En dobbeltblinding vil være umulig. Vanskene med randomisering begrunner Bonchek med at leger ofte har preferanse for medikamentell konservativ behandling eller en prosedyre. De vil for eksempel fjerne høyrisikopasienter fra studien ved at disse ikke blir inkludert eller ledet mot medikamentell behandling. Dette gjelder når man skal sammenligne medikamenter mot prosedyrer. Ved sammenligning av to medikamenter eller to prosedyrer gjelder ikke dette. Personlig synes jeg denne påstand er å undervurdere klinikerens vurderinger og tiltro til å gjøre en rettferdig fordeling av pasientene. Problemet kan løses ved å ha flere og mer eksakte inklusjons- og eksklusjonskriterier, og la noen som ikke kjenner pasienten randomisere. De fleste klinikere er så profesjonelle at de kan forholde seg til dette, og vil ikke ta det som et nederlag om de ikke får gjøre en bestemt prosedyre, men heller må gi medikamenter. Likevel vil det kunne være en gruppe pasienter som defineres som høyrisikopasienter og som ikke vil ha nytte av en prosedyre pga risiko, men da må disse pasientene ikke ledes over til den medikamentelle armen av studien da dette vil kunne gi bias.

### Muliggjøring av multicenter forsøk

I dag anvendes multicenter forsøk i stadig økende grad for å få et stort nok antall forsøkspersoner med i studien. For at et slikt forsøk skal være mulig og for at alle som deltar er med på de samme premissene, er det viktig at alle er med på like vilkår. For å få dette til utformes en forskningsprotokoll, der premissene er klare og et viktig redskap i gjennomføringen av studien.

Da er man enige om de ulike kriterier og i hvordan man skal utføre forsøket. En fordel ved multicenterstudier er at man kan oppnå en heterogenitet i studiepopulasjonen ved at ulike sykehus deltar i studien. Man må likevel passe på at alle har en felles grunnoppfatning og medisinsk tradisjon. Det vil for eksempel være vanskelig for et norsk sykehus å delta i en russisk studie fordi de medisinske tradisjoner er så ulike.

Det er blitt hevdet at små sykehus med et lavt pasientvolum bidrar lite til en multicenterstudie og at kvaliteten på det de leverer er dårlig (52.) Hjort et. al har i sin artikkel sett på kvaliteten, målt ved bla. inklusjonrate, hos store og små sykehus som deltok i NMSG #4/90. Der viser de at det ikke er noen kvalitetsforskjell. Det konkluderes med at små sykehus tvert i mot gir et viktig bidrag ved å øke representativiteten og at man raskere oppnår det nødvendige antallet forsøkspersoner i studien. Dette er konklusjoner som jeg støtter. Dersom protokollen er klar og utvetydig er det ikke noe galt i at små sykehus skal delta i kliniske forsøk. Det vil kanskje stille litt større krav til de sentrale sekretariater, som får flere å forholde seg til og må være villige til å håndtere flere spørsmål. Men dette er ulemper som klart oppveies av de fordeler Hjort et al (52) nevner. Ellers kan man også argumentere med at mange pasienter ikke behandles på de store sykehusene, men på de små. Ved å inkludere små sykehus i multicenterforsøk skaper man også tillit til hverandre og bedrer kommunikasjonen mellom sykehusene, noe som trengs. Kommunikasjon er også viktig for å opprettholde en høy inklusjonsrate (52), noe som en viktig kvalitetsindikator.

### Tolking av kliniske forsøk

Resultatene av kliniske forsøk, spesielt RCT, er et kraftig instrument for medisinske beslutninger. Dette forutsetter imidlertid at resultatene tolkes riktig og på en god måte. Det er viktig at man er klar over at det finnes bias og at man forsøker å minimere dem med blinding og randomisering. Man må heller ikke legge for mye vekt på p-verdien, slik Parmar forklarer i sin artikkel (48). Man skal også være klar over at man regner med sannsynligheter. Det er et ordtak som sier at all statistikk lyver og det kan være sant. Mye handler om hvordan man ser på de forskjellige tall og verdier, jfr. Parmars diskusjon om relativ og absolutt risiko (48). Disse tallene, spesielt p-verdien, er avhengig av antall studiepersoner slik at denne blir mer usikker i små studier. Klinisk signifikans er ikke lik statistisk signifikans, slik Parmar også hevder (48). Dette er et godt poeng, da man lett vil kunne finne eksempler på at legemidler med litt bedre effekt i henhold til statistisk signifikans, ikke har klinisk signifikant effekt. Man skal derfor lese en studie

på en kritisk måte og stille spørsmål ved de konklusjoner som trekkes. Det er slik at selv om kliniske forsøk, og spesielt RCT, gir de beste bevisene, så kan disse misbrukes. Misbruket skjer spesielt blant ukyndige og personer med kommersielle interesser. Man hører stadig om at studier har vist at noe er farlig, mens man senere erfarer at dette er det sunneste man kan gjøre eller innta.(38). Det er ikke for ingenting at mange kalle avisene i Akersgata for de største medisinske tidsskriftene. Dette er til tider meget frustrerende for deler av befolkningen. Det er imidlertid lite de medisinske fagmiljøene kan gjøre, annet enn å være faglig oppdatert og delta i debattene, da opplysninger i pressen kan bero på mistolking av kliniske data. Alvorligere er et mer bevisst misbruk, som spesielt bedrives av miljøer med kommersielle interesser. Legemiddelindustrien er pålagt å dokumentere effekt med kliniske forsøk før et preparat kan bli registeret. Det forekommer dessverre kliniske utprøvinger som ikke holder mål, for eksempel sammenlikning av to eksperimentelle armer (Dahl, personlig meddelelse). Manipulasjon av statistikk forekommer også. Reduksjon eller bedring i relative tall fremheves, mens det er minimale utslag i absolutte tall. Dette er uheldig fordi man stoler på klinisk forsøk. Jeg vil derfor konkludere med at siden kliniske forsøk, spesielt RCT, oppfattes å ha en så høy bevisgrad og impact på våre holdninger, må man være på vakt i forhold til misbruk.

### *Like riktig og likeverdige?*

Som Feinsein tar opp i sin artikkel finnes det to likeverdige og like sanne måter å gjennomføre et klinisk forsøk på (32). Det er pragmatikeren som tar hensyn til heterogeniteten i befolkningen, sammenligner eksisterende behandling mot ny, dosetitrering og mykere endepunkt mens idealisten på mange måter står for det motsatte. Ut i fra et umiddelbart inntrykk kan det synes som om pragmatikeren designer kliniske studier som er mer tilpasset en klinisk hverdag der man bruker mer skjønn. Idealisten på sin side ser ut til å designe studier som er mer rigide og gir mindre plass for individuelle vurderinger. Pragmatikeren er nok den som appellerer til klinikerer flest nettopp fordi han anvender nødvendig skjønn i arbeidet med mennesker. Personlig tror jeg vi trenger begge typer forskere fordi de vektlegger litt forskjellige endepunkter med ulik styrke, f.eks. livskvalitetsstudier. Ved å ha en sunn blanding av disse designmåtene og ulike måter å tolke resultatene på kan man få belyst ulike aspekter ved sykdommen. Med en slik tilnærming til et klinisk problem vil man få et bedre resultat når man skal gjøre opp resultatene. Man kan også argumentere med om at dersom en ny behandling skal implementeres bør den være signifikant bedre hos både idealist og pragmatiker.

### Kokebokmedisin

En kritikk mot EBM, og derved kliniske forsøk, er at det fremmer utviklingen av det som kalles kokeboksmedisin. (31) Kokebokmedisin innebærer at man slavisk behandler alle pasienter med samme sykdom likt. Det blir som om en diagnose utløser et helt bestemt og uforanderlig behandlingsregime og derved lite rom for klinisk skjønn, hevder kritikerne. Personlig deler jeg ikke helt dette syn. Det er nok slik at en diagnose vil utløse bestemte behandlingsregimer, men dette er ingen ulempe. Man har jo nettopp gjennom forsøk, ofte flere, vist at den aktuelle behandling er til det beste for pasienten i form av lavere mortalitet og morbiditet. Påstanden om at det kliniske skjønnet forsvinner er overdrevet. Det vil fremdeles være behov for dette i diagnostisering og ikke minst i gjennomføring av behandlingen. Mange av de kliniske forsøkene er gjort på yngre ellers friske mennesker, mens den kliniske hverdagen er noe helt annet. Der møter man også eldre pasienter som er multisyke og har ulike krav og forventninger. Ikke alle vil passe inn i de behandlingsregimene som de kliniske forsøkene har fremskaffet, nettopp pga. heterogeniteten. For disse pasientgruppene vil man være tvunget til å bruke klinisk skjønn og vurdere standardbehandlingen eller annen terapi bør utprøves. Det finnes også eksempler på at resultatene av ulike kliniske forsøk er motstridene, eksempelvis for hypertensjonsbehandling. I slike situasjoner vil det klinisk skjønn være viktig. I argumentet for kokebokmedisin ligger også en av de største fordelene til kliniske forsøk, nemlig at man får en standardisert behandling som beviselig virker. Dette sikrer en likebehandling av pasienter slik at alle får god behandling, uansett hvordan de er som personer og hvor i verden de bor. Tidligere har det forekommet små miljøer med egne rutiner og meninger om behandling innen et fagfelt. Disse miljøene har vært bygd opp rundt en autoritetsperson, ofte en professor, som på bakgrunn av empiri og anekdotisk kunnskap har innført sine egne behandlingsprinsipper. Ved å bruke EBM kan man unngå dette og ved at kunnskap blir vitenskapelig basert (54) tillates også yngre leger uten tilsvarende erfaring å delta i diskusjonen.

### Administrasjon og kostnader

En klar ulempe med de kliniske forsøkene er at det er mye administrasjon rundt disse, som tar tid. En konsekvens av dette er at de koster en del penger. Årsaken til at det er mye administrasjon

er at man forsker på mennesker. Dette medfører at mange forskrifter og regler må tas hensyn til, bla Helsinkideklarasjonen. Med dette følger et skjemavelde og byråkrati. Man må for eksempel søke forskningsetisk komité. Det er nok viktig at man skal ha god kontroll, spesielt når mennesker anvendes som forsøksobjekter. Likevel må man holde byråkrati og skjemavelde under kontroll og forsøke å forenkle dette. Ellers kan man risikere at dette blir et hinder som mange forskere ikke klarer å penetrere slik at de til slutt gir opp. Det vil være meget synd og som ytterste konsekvens være skadelig da man ikke får rekruttert nok fagfolk til forskningsvirksomhet.. Et eksempel i denne oppgaven på det store byråkratiet er Glivec® studien (29). Her er det ikke myndighetene som kompliserer arbeidet, men selskapet som utfører studien. Når skriftlig informert samtykke kreves får pasienten åtte tettskrevne sider med informasjon. Dette blir voldsomt for både pasient og den legen som skal rekruttere pasienter til studien. Man nærmer seg da situasjoner som Edwards beskriver som "slemt og engstelsesskapende"(40). I tillegg stilles det store krav til de klinikerne som deltar vedrørendel dokumentasjon og andre plikter. Jeg er ikke uenig i at pasienter skal informeres grundig, men dette bør gjøres på en mer klar og konsis måte slik at de lettere forstår essensen av forsøket.

En annen årsak til at det ofte blir mye administrasjon er at man trenger et bestemt antall forsøkspersoner for å få nok "power" og signifikant p-verdi slik Parmar skisserer i sin artikkel (48). Det er ikke uvanlig, spesielt innen kardiologien, at man hører om kliniske studier med flere tusen forsøkspersoner. Årsaken er at man har sjeldne endepunkter. Det sier seg selv at dersom alle disse skal følges opp på en god måte kreves det mye administrasjon. Det koster penger og krever personalressurser. Man kan diskutere om det å drive med klinisk forsøk er verdt kostnadene når man tenker på hvor mye billigere mange andre forskningsmetoder er. På den annen side er kliniske forsøk den metoden som gir de sikreste resultatene (29) og spørsmålet blir hva samfunnet er villig til å betale for kunnskap. Personlig tror jeg at selv om det er relativt dyre metoder er det noe som må ses på som en investering for framtiden ved senere å hindre tidlig død, lidelse og unyttige behandlinger. Det er ikke til å se bort i fra at forskere som driver denne type forskning kan bli mer økonomisk bevisste enn tilfelle er i dag. Samfunnet må på sin side bør også bidra ved å minske byråkrati og papirmøller slik at man kan spare penger på unyttig administrasjon. I tillegg bør de økonomiske rammene for klinisk forskning, og derved stabiliteten omkring denne bedres. Som sagt er forskning en investering for framtiden som på sikt vil føre til kostnadsreduksjoner.

### Fordel ved deltakelse for forsker og pasient

Selv de leger som ikke er med i planleggingen og styringen av en klinisk studie vil ha fordel av å delta. Det har en læringseffekt å være med i kliniske studier. Man lærer seg vitenskapelig tankegang og nye metoder for diagnostikk og behandling. Den er viktig som en faglig oppdatering, noe som på sikt vil gagne klinikerens pasienter. Læringseffekten er den nest viktigste grunnen til at leger inkluderer pasienter i kliniske studier (49) og bør dyrkes ved at studiegruppen jevnlig sender ut informasjon om studiens progresjon, resultater og ved å holde ulike kurs. Dette vil, i tillegg til å ha en økt læringseffekt, også virke kvalitetsfremmende ved at deltagende klinikere føler de gjør noe nyttig. Slik vil flere pasienter inkluderes i studier, noe som øker representativiteten i studiematerialet. Å ha en stor andel av legestanden som forskere er også viktig for å forhindre bias av forsøkene. Dette fordi man lett ellers kan komme i en situasjon der man får en selektering av forskende leger. En selektering i forskningsøyemed vil alltid føre til en skeivfordeling og mulighet for bias. Det er dette som Kaptchuk kaller for "investigator self-selection" ved at man får selektert ut en spesiell gruppe leger med spesielle holdninger og oppførsel (34). Dette vil i sin tur kunne påvirke den behandling pasientene får. Av den grunn bør så mange som mulig forske. Dette gjelder spesielt dersom de legene som forsker skulle vise seg å være bedre klinikere, noe jeg personlig tror kan variere.. Det er også viktig at flest mulig av legene forsøker for å unngå de forsøkseffektene som Braunholtz et al (37) nevner i sin artikkel, som care-effekten og Hawthorne effekten. Som en konsekvens av dette bør man heller ikke selektere ut spesielle sykehus som skal forske for bla. å unngå "care" effekten.

Det viser seg at også pasientene har en fordel av å delta i kliniske studier (37) Da ser man bort i fra den opplagte nytten de måtte ha av en ny metode eller et bedre medikament. Pasienten opplever en positiv korttidseffekt og et signifikant bedre klinisk utkomme enn de pasienter som ikke deltar, selv om behandlingen ikke viser seg å være bedre. Dette skyldes bla. "care" effekten og at pasientene som deltar i kliniske studier følges bedre opp og etter rigide skjemaer. Dette gjelder spesielt for de maligne sykdommene der man er redd for å overse noe (Dahl pers. med.) Her ligger også noe av problemet med å bruke historiske kontroller i kliniske forsøk fordi man da overser bla. "care" effekten. De historiske kontrollene vil derfor ofte ha et dårligere utfall enn pasienter som deltar i kliniske studier.

### Interrimanalyser

Som en kvalitetssikring av kliniske studier har man interrimanalyser. Dette anvendes spesielt i multisenter forsøk, men er ikke begrenset til denne type studier. Da oppnevner man før studien begynner en uavhengig komité som skal foreta visse analyser underveis i studien. De har myndighet til å stanse studien dersom den skulle være skadelig for deltakerne eller dersom den nye behandlingen skulle vise seg å være overlegent bedre enn etablert behandling. Dette betyr en kvalitetssikring for deltakerne. Det er intet krav om at alle kliniske studier skal ha interrimanalyser, men bør gjennomføres i store studier som går over lang tid slik at disse kan avsluttes straks hypotesen er besvart. Et problem, spesielt i Skandinavia, er at fagmiljøene er så små at alle forskerne kjenner hverandre slik at det kan være vanskelig å få et uavhengig og uhildet interimsstyre. Det kan også oppstå uenighet når man må stoppe en studie og studieledelsen ikke er enig i vurderingene. For dem er jo studien deres "barn". Av disse årsakene kan det derfor bli vanskelig å stoppe en studie. Kritikkk mot interrimanalyser er at man ikke får fullføre studien og dermed ikke har fullstendige resultater. Det kan være slik at man pga. erfaring skaper bedre resultater. Uakseptable bivirkninger som kan skade pasientene bør imidlertid alltid føre til at en studie avbrytes. Det kan være betenkelig dersom man påviser en forbedring med den nye behandlingen og derfor velger å stanse studien fordi alle kontrollene skal få den nye behandlingen. Ut i fra et etisk ståsted er tankegangen riktig. Det man glemmer er at man kan ha ufullstendige resultater fordi oppfølging over en lengre periode mangler. Man risikerer å gi mange pasienter skader fordi langtidsbruk ikke er undersøkt godt nok eller fordi en såkalt "banan-effekt" ikke er påvist, d.v.s. at den initiale effekten av et medikament (eller prosedyre) avtar (Dahl pers. meddelse). Derfor bør alle studier ha en viss varighet, i alle fall når det gjelder oppfølging. Derfor er jeg også litt i mot at pasienter som ikke deltar i kliniske studier skal få forsøksmedisinen slik som Minouge (47) foreslår. Man kan risikere iatrogene skader på en større gruppe enn det som er nødvendig, Men alt i alt er det riktig å ha en interrimanalyse nettopp for å gjøre slike vurderinger underveis.

### Helhetsforskning

En av de kliniske forsøkernes største fordel er, som Kaptchuk hevder (34), at de (spesielt RCT), forsøker å etterlikne objektiviteten i laboratorieforsøkene. Samtidig som man søker dette får man brukt mennesket som helhet i forsøket. I et laboratorieforsøk vil man plukke ut en del av



mennesket, eks. levercelle, og teste medikamenter eller funksjoner på disse. Man gjør da en generalisering ut fra disse cellene om at medikamentet vil fungere på en bestemt måte i kroppen. Det man underslår er at leveren består av flere ulike celletyper som sammen fungerer i en funksjonell enhet og at menneskekroppen ikke alltid oppfører seg slik som man forventer. De kliniske forsøkene ser individet som en helhet, og tar høyde for at mennesket virker annerledes enn hva summen av alle enkeltkomponentene skulle tilsi. Det er også derfor alle nye medikamenter skal igjennom kliniske forsøk før de godkjennes. Dette gir de kliniske forsøkene en unik stilling innen medisinen. En annen fordel ved å bruke kliniske forsøk er at mange symptomer og årsaker kun kan forskes på ved å bruke helhetsforskning slik som de kliniske forsøkene representerer. Eksempler på symptomer som sjelden er gjenstand for forskning er hoste og diaré. Dette er vanlige tilstander som vi ikke vet så mye om (54). Campbell tror at kliniske forsøk er måten å belyse disse spørsmålene på og derved sørge for livskvalitetsøkning til mange pasienter ved å lindre disse plagsomme symptomene.

En kritikk mot de kliniske forsøkene er at man bruker mennesker som "forsøkskaniner", men alle som deltar gjør jo dette på frivillig basis. Dessuten opplever de som deltar en positiv effekt (37). Pasientene er også beskyttet av titalls ulike lover og regler bla. Helsinki-deklarasjonen. Jeg tror ikke et vesentlig antall pasienter føler seg som "forsøkskaniner", men heller er stolte av å delta i vitenskapens framskritt.

Det har også vært kritikk mot de kliniske studiene ved at de bruker for mange harde endepunkter (54), men det er vanskelig å se noen annen forskningsmetode. Ved å fokusere kun på myke endepunkter får man ofte mer usikre svar. Det bør i tilfelle bli et fokusskifte mot å legge til myke endepunkter, noe som griper rett inn i debatten mellom den idealistiske og kresne forskeren som Feinstein skisserer (32). En klar fare med å bruke for mange myke endepunkter, noen få er greit, er at man risikerer å få utolkbare/uklare og utydelig resultater (32). For pasienter fleste er kanskje de myke endepunktene vel så viktig som de harde. Mange er av den oppfatningen at et par ekstra måneder å leve ikke er verdt smerter, kvalme eller injeksjoner flere ganger daglig. Derfor bør man også vektlegge de myke endepunktene i en viss grad, nettopp for å få en avklaring av slike spørsmål.

En annen mulighet er å gjennomføre sideprosjekter, som NMSG har gjort i for eksempel protokoll # 12/00 (15)

### Noen etiske betraktninger

Til kliniske studier, kanskje spesielt RCT, er det knyttet noen etiske betraktninger og bekymringer. EBM er, som Kerridge hevder (53), et praktisk eksempel på konsekvensialismen, dvs at en handlings etiske verdi måles ut i fra konsekvensene den skaper. Kritikkk mot konsekvensialismen er at mange resultater ikke kan måles eller defineres. Hvem sin interesser skal telle og hvilke mulige uetiske konklusjoner blir trukket, sett fra andre ståsteder ?. Ikke målbare eller definerbare resultater er jo typisk for de myke endepunktene man kan ha i kliniske studier. De myke endepunktene er likevel viktige jfr. diskusjonen tidligere, spesielt for pasientene. Dermed kan EBM, ut i fra et etisk standpunkt, være litt problematisk. Man kan jo også hevde at konsekvensen for pasienten er en bedre behandling som har et vitenskapelig fundert grunnlag, og at EBM dermed ikke kan rammes av kritikken mot konsekvensialismen på punktet om udefinerbare endepunkter. Et mer problematisk område er hvem sin interesser som skal telle i resultatet. Umiddelbart vil de fleste svare pasientens. Det er imidlertid ikke helt i overensstemmelse med virkeligheten.. Kerridge hevder at pasientene har liten innflytelse over hva som skal forskes på (53). Det er fremdeles eksperten og den farmasøytiske industrien som på vegne av pasientene bestemmer hva de mener er til pasientens beste. Pasientene er heller ikke helt uskyldig når det gjelder hvem sin interesser som skal telle. Mange av dem handler ut fra egeninteresse når de godtar å bli med i en klinisk studie. (40). Det er faktisk deres hovedmotiv. Det er også et etisk problem at man bruker mennesker som "forsøkskaniner" for å utvikle nye behandlingsmetoder. Menneskeverdet reduseres og man mister fokus på enkeltpersonene som er med i studien. Man har ikke lenger med individer å gjøre, men forsøkspersoner. Dette blir ekstra komplisert siden man ofte er behandlende lege for pasienten. Men det å være behandlende lege kan også være en fordel ved at man ser individet bak forsøkspersonen.

### Til hjelp ved prioriteringer

Man kan ikke lenger skille medisin og økonomi (31). Prioritering av pasienter er vanskelig og de kliniske forsøkene kan hjelpe oss. Ved å ha et vitenskapelig grunnlag kan man lettere argumentere for sine prioriteringer. Prioriteringer er helt nødvendig da man ikke har ubegrensede ressurser i det norske helsevesenet. (56). Dette gjelde både 1. og 2. ordens prioriteringsbeslutninger. Likevel er det mange områder innen medisinen hvor man mangler dokumentasjon (56). Dette gjelder spesielt innen kirurgi (31). Dette er et problem da

myndighetene ikke vil finansiere behandling som ikke er dokumentert helsenyttig. Men siden behandlingen er så etablert vil ikke pasientene godta å ikke få tilbud om denne. Derfor tyr man ofte til historisk kontroll og tall, jfr. min tidligere diskusjon om positive effekter.

Dokumentasjon er et redskap til å vurdere kostnad-nytte og kostnadseffektivisering, som i sin tur kan anvendes i prioritering. Med begrensede ressurser er man nødt til å sette et tak på hvor mye man kan betale for en relativt lite nyttig behandling. NOU har satt en nedre grense på 5-10 % bedret langtidsoverlevelse(56). Ved å bruke dokumentasjon vil man kunne likebehandle ulike grupper, selv om sykdom og prognose er forskjellig, noe som er et viktig prinsipp. Det betyr at alle pasienter med samme diagnose og prognose skal behandles likt over hele landet. Oppslag i dagspressen har imidlertid vist at det forekommer forskjeller mellom de ulike helseregionene. NOU-rapporten legger opp til at de ulike fagmiljøene skal danne egne prioriteringsgrupper som skal prioritere innen sine fagfelt. Dette er fornuftig da man har ekspertpanel som på bakgrunn av vitenskapelige bevis, der kliniske forsøk spiller en meget viktig rolle, kan prioritere sine pasienter. I slike faggrupper er det også plass til representanter for pasientene og lekfolk. Faggruppen anbefales å prioritere fire prioriteringsgrupper, slik det er skissert for beinmargstransplantasjon (55). Da slipper man å sette grupper opp mot hverandre og unngår at ressurssvake grupper nedprioriteres. Dette er likebehandling i praksis. Ved å følge anbefalingene fra faggruppen unngår man også forskjellsbehandling i forhold til hvilken helseregion man tilhører, noe som er viktig for befolkningen og et prinsipp helsevesenet må ta på alvor. Ved å kunne vise til kliniske studier er det også lettere for pasienter å forstå hvorfor vedkommende ikke kan få en spesiell behandling han mener kan være nyttig og derved å har krav på. . I fremtiden vil Internett bli en utfordring for legene da pasientene vil lese om eksperimentelle og nye og behandlingsformer og kreve å få disse. Da vil godt gjennomførte kliniske studier bli viktige i argumentasjonen, ekspertkunnskap blir lettere akseptert og "legeshopping" blir vanskeligere.

Økonomisk støtte til kliniske studier kan være problematisk da man initialt ikke har dokumentert nytte av behandlingen. Dette har stamcelletransplanasjonsgruppen løst ved plassere pasienter som deltar i studier i prioriteringsgrupper III, altså relativt langt nede. Dette er rasjonelt da man ikke kjenner nytteverdien av den behandling som gis til disse pasientene, men mener det er viktig å finne dette ut v.h.a. vitenskapelige metoder. Personlig synes jeg det er litt merkelig at kliniske studier og standardbehandling skal finansieres fra samme budsjett. (Helsebudsjettet). Det bør lages et skille da viktige studier lett blir nedprioritert. Etablert og utprøvende behandling er gjensidig avhengig av hverandre, men bør ikke likestilles i prioriteringssammenheng. Til

forsvar for denne måten å gjøre det på er at selv eksperimentell behandling legger beslag på store ressurser. Dette tas ofte fra pasienter som får etablert behandlingen, d.v.s. som har dokumentert nytte av behandlingen. Et godt eksempel på dette er stamcelletransplantasjon (SCT) ved KML, som havner i prioritetsgruppe I og altså er regnet som meget god og kurativ standardbehandling, mens SCT ved myelomatose regnes som utprøvende behandling(55). På dette grunnlag blir det naturlig å prioritere standard- og eksperimentell behandling ulikt.

### *Hematologi og kliniske studier*

Et viktig kjennetegn ved de maligne hematologiske lidelsene er at de er relativt sjeldne og har en lav langtidsoverlevelse. Myelomatose som er den vanligste maligne hematologisk sykdommen har en innsidens på 6/100.000 (1). Dette betyr at relativt få pasienter rammes og at man har problemer med å få inkludert nok pasienter til en klinisk studie. Dette illustrerer behovet for multicenter forsøk, også på tvers av landegrensene, slik NMSG arbeider. På denne måten kan man få nok deltakere til at studien gir et sikkert resultat. Et samarbeid over landegrensene vil også øke representativiteten i studiematerialet (52) og dermed kan man lettere generalisere, selv om man skal være kritisk i forhold til dette. Hematologi er også et lite fagfelt innen indremedisinen med et begrenset fagmiljø. Samarbeid mellom helseregionene og intraregionalt er derfor viktig. Dette blir ytterligere synliggjort når man vet at mange av pasientene har hyppige innleggelse på lokalsykehus som nødvendigvis ikke har hematologisk spesialkompetanse. (49). Noen maligne hematologiske sykdommer behandles også på lokalsykehus noe som kan skape problemer i gjennomføringen av kliniske studier fordi pasientene er så spredt. Uten et godt samarbeid mellom sykehusene blir dette en umulig oppgave. Det illustrerer også behovet for en klar og utvetydig protokoll. Det er også viktig at man ikke overlesser lokalsykehuset med oppgaver knyttet til studier, men forsøker å ha en liten administrasjon, enkle skjemaer og et fornuftigutvalg av undersøkelser, som i de NMSG studier jeg har beskrevet.(52) Personlig mener jeg Glivec® studien, med sin kompleksitet, hyppige "site visits" og meget rigide rapporteringssystem (Dahl pers.med.) skyter over mål og kan skremme vekk leger fra å delta. .

Noe som også er spesielt med de maligne hematologiske lidelsene er at de er potensielt dødelige. Pasientene krever en tett oppfølging og det er vanskeligere å få dem med i studier fordi de er redde for å få noe som kanskje ikke virker. Placebobasert forskning er spesielt vanskelig å drive på kreftpasienter. På den annen side finnes det også pasienter som har tro på det nye og derfor er

villige til å delta. Denne tankegangen er forståelig da det er pasientens eget valg å delta i en studie. Problemet er om det skulle oppstå en skeivfordeling av pasienter og derved bias. Her skiller ikke forskning innen hematologi seg fra annen forskning på maligne sykdommer. Et annet kjennetegn ved hematologisk malignitet er at man ikke har noen solid svulst man kan fjerne kirurgisk. I stedet er man helt avhengig av cytostatika, eventuelt i kombinasjon med stråler. Fordelen er at man slipper diskusjonen om prosedyrer ved kliniske studier, der erfaring er et viktig element. Nå har man mulighet for en standardisering og utprøving av ideelle doser. Det er også en mulighet for en enkel oppfølging med blod- og beinmargsprøver. Selv om dette også er invasivt gir undersøkelsene presise resultater, uavhengig av prøvetaker.. Det er lett å måle respons. Ved å bruke cytostatika har man også større muligheter til utvikling av ny behandling. For solide svulster er behandlingen ofte kirurgi evt. med stråling og/eller cytostatika som adjuvans.

I utviklingen av ny behandling er også forbedring av eksisterende behandling et viktig element. Et godt eksempel er protokoll #11/00 fra NMSG (14). Der studerer man om Cyclofosamid og Dexametason er bedre og mildere induksjonsbehandling til HMAS enn Vinkristin, Dexametason og Doxorubin (VAD). Skulle det vise seg å være slik er det en stor forbedring for pasientene da induksjonsbehandling med Cy+Dex tar kortere tid og trolig har mindre bivirkninger. Ved å bruke cytostatika har man også muligheten til å legge til medikamenter til standardbehandlingen slik som NMSG gjør i protokoll #8/00 (9) og #12/00 (15). Dette er kliniske studier som har til hensikt å forbedre overlevelse og livskvalitet.

Man er langt fra målet om kurativ behandling, med disse medikamentene kan være et skritt i riktig retning. Av og til gjør man kvantesprang, også innen hematologien. Et eksempel er HMAS, som radikalt har forbedret levetiden for myelomatosepasienter. Dette er bevist med hjelp av kliniske studier, flere av dem kontrollerte. Altså kan vi med bakgrunn i Antes et als tabell (29), se tabell 1, konkludere med at et høyt bevisnivå også er nødvendig i hematologisk forskning.

Et annet særtrekk innen hematologien er at pasientene ofte har mange liggedøgn i året. Det skyldes behandling med cytostatika og lange perioder med cytopeni. Pasientene er sårbare for infeksjoner, blødninger og utvikling av anemi og det er ofte nødvendig med isolasjon. Dette betyr at forsøkseffektene, som Braunholtz (37) nevner i sin artikkel blir viktigere, bla. "care" effekten fordi man får et nærere forhold til pasientene som kan påvirke utkommet i en positiv retning. Som motargument mot dette kan man hevde at disse effektene "rammer" alle hematologiske pasienter, uansett om de er med i en studie eller ikke. Selv om dette er delvis

riktig vil pasienter i kliniske studier få bedre oppfølging i form av hyppige kontroller og lavere terskel for innleggelse.

Hematologien står også i en særstilling når det gjelder prioriteringer. Det kommer av at pasientgrunnlaget er lite og man derfor har vansker for å bygge opp kompetanse ved flere sykehus. Det medfører at man må sentralisere tilbudet, noe som medfører økte kostnader til bla. reise. Dette er en utfordring for samarbeidet mellom helseregionene. Det er ikke nok pasienter til at man for eksempel kan utføre allogene beinmargstransplantasjoner på mer enn et sted i Norge (Rikshospitalet). En annen årsak til særstillingen er at behandlingene er meget kostbare. Det kommer av at medikamentene som brukes er dyre og at man bruker dem over tid. Også de støttende medikamentene som tilvekstfaktorer (G-CSF, Epo) er dyre. Man har forsøkt med Interferon  $\alpha$ -2b som vedlikeholdsbehandling ved myelomatose, men NMSG har i en av sine tidligere protokoller ikke kunnet påvise nytteeffekt av dette cytokin.. Interferon  $\alpha$ -2b er en behandling som koster ca. kr. 300.000 i året (Dahl pers meddelse). Dette er et godt eksempel som viser hvor nyttig de kliniske studiene er ved at man får fjernet unyttige behandlinger. Et annet element i prioriteringsdebatten og hematologi er bruken av beinmargstransplantasjoner. En beinmargstransplantasjon koster gjennomsnittelig mellom kr 500.000 og 600.000 . I tillegg kommer det mange medisinske risikofaktorer, som fare for infeksjon og graft versus host disease (GVHD) samt at behandlingen er meget krevende for pasienten. Det er også en begrenset tilgang til donorer. En beinmargstransplantasjon er eneste kurative behandling for mange blodsykdommer, som for eksempel alvorlig aplastisk anemi. Der man ikke har andre tilbud blir ofte stamcelletransplantasjon siste utvei, selv som utprøvende behandling. Da er det viktig å ha klare prioriteringer, slik det er anført av et rådgivende organet i Helsedirektoratets utredningsserie IK-2381(55). Pasienter som prioriteres i prioriteringsgruppe III, skal ikke ta ressursene fra pasienter i prioriteringsgruppe I og II der kliniske studier har vist god nytte av behandling. Dette viser også behovet for kliniske studier som hjelp i prioriteringer, slik at pasienter med dokumentert nytte prioriteres foran pasienter der nytteverdien er liten eller fraværende. Men det kan være frustrerende for pasienter som kjenner til stamcelletransplantasjon som behandling med et mulig kurative potensialet. Frustrasjonen er ekstra stor når pasienter får avslag på søknad om transplantasjon og har lest på Internett eller i annen litteratur at det har hjulpet andre i liknende situasjon. Det hjelper ikke at pasientene fores med informasjon fra mange kanaler og kommer i villrede, slik som Angell et al tar opp i sin editorial (38). Et element som ytterligere øker frustrasjonen er at mange har maligne sykdommer og vet at tiden er i ferd med å renne ut for dem. Situasjon er da vanskelig både for pasient og kliniker. Da kan det være godt, i alle fall for klinikerne, å ha klinisk forskning å støtte seg til i beslutningsprosessen.

## Konklusjon

Som en konklusjon kan vi si at de kliniske forsøkene har mange fordeler, men også noen ulemper. De største fordelene som de kliniske studiene har er at de gir oss de sikreste medisinske bevisene vi kan oppnå. På dette kan man bygge opp veiledende retningslinjer som kan hjelpe oss i det vanskelige prioriteringsarbeidet. Dette ved at man kan ta vekk unyttig behandling og rette ressursen mot de behandlingene som beviselig er effektive og nyttige. Dette vil også sikre en rettferdig likebehandling av pasientene ved at alle får den mest optimale behandlingen som finnes, uansett personlige forhold eller bosted.

De kliniske forsøkene har også den fordelen at de ser mennesket som en funksjonell enhet, og ikke bare summen av mange små komponenter. Dermed har man et ypperlig verktøy ved utprøving av nye medikamenter og behandling ved at man ser *hele* mennesket slik at uforutsette faktorer ikke overses. I de kliniske forsøkene har man også et verktøy for å undersøke mange lidelser på en ny måte, som kan gi annerledes kunnskap enn det som var tilsiktet.

Kliniske forsøk har også den fordelen av at de krever en klar og konsis forskningsprotokoll. Dette muliggjør samarbeid mellom mange sykehus. Slik kan man oppnå større representativitet i studiematerialet og man bedriver etterutdanning av de deltagende legene.

Selv om de kliniske forsøkene har mange fordeler og er et kraftig vitenskapelig instrument så har de også sine ulemper. De kan lett misbrukes, bevisst eller ubevisst. Dette henger bla. sammen med hvordan man tolker resultatene og hvilke verdier i resultatene man velger å vektlegge. Spørsmålet om vi trenger en platinastandard er spennende da det viser oss hvordan vi forandrer våre holdninger og hvordan det kan påvirke utfallet av en studie. Det viser også behovet for randomisering og dobbeltblinding for å øke sikkerheten av resultatene.

De kliniske forsøkene har vist seg å ikke være like gode når det gjelder prosedyrer og operasjoner av årsaker som tidligere er nevnt. Man må nok derfor vurdere andre instrumenter enn de kliniske forsøkende for dette. De kliniske forsøkene innebærer også en del kostnader pga. en del administrasjon og at man trenger et stort studiemateriale. Det er også en del etiske betraktninger knyttet til bruk av kliniske forsøk. Man bruker rett og slett mennesker som forsøkskaniner. Dette er et faktum man ikke kan komme bort i fra, selv om man forsøker å øke

sikkerheten for deltagerne med bla. Helsinki-deklarasjonen. Andre etiske bekymringer er knyttet til hvilke utkomme man måler og hvems inntresser som skal telle.

Hematologien skiller seg ut fra resten av medisinen ved at det er en liten pasientgruppe med ofte sentralisert behandling. De maligne hematologiske lidelsene ikke er solide og det åpner for større andel av klinisk bevis ved bruk av kliniske forsøk på medikamenter. Dermed kan man oppnå en større andel av randomiserte kontrollerte studie bevis enn i resten av medisinen. Man kan derfor konkludere med at klinisk studier passer ypperlig for hematologien, også med tanke på at man har et større behov for multicenter forsøk (pga lave insidenser), enn i mange andre deler av medisinen.

Konklusjonen blir derfor at det dobbelblindede randomiserte kliniske forsøk er vårt beste instrument når det gjelder utvikling og evaluering av nye behandlinger, også innen hematologien.



**Tabell 1**

Nivå	Type undersøkelse
1a	Metaanalyse av randomiserte kontrollerte studier.
1b	Randomisert kontrollert studie
2a	Veldesignet kontrollert studie
2b	Veldesignet quasi-eksperimentell studie
3	Ikke eksperimentelle studier som sammenlikningsstudier, case-control studier og lignende
4	Konsensusrapporter, klinisk erfaring fra eksperter og lignende

Tabellen viser graden av bevis etter hva slags studie som gjøres, det høyeste bevisgraden øverst. Fra Antes et al(29).

**Tabell 2**

Stu-di-estørrelse	Relativ risiko	95 % konfidensintervall	P-verdi
20	0,82	(0.12,3.79)	0,33
200	0,82	(0.47,1.43)	0,24
2000	0,82	(0.69,0.98)	0,01
20 000	0,82	(0.77,0.87)	>0,00001

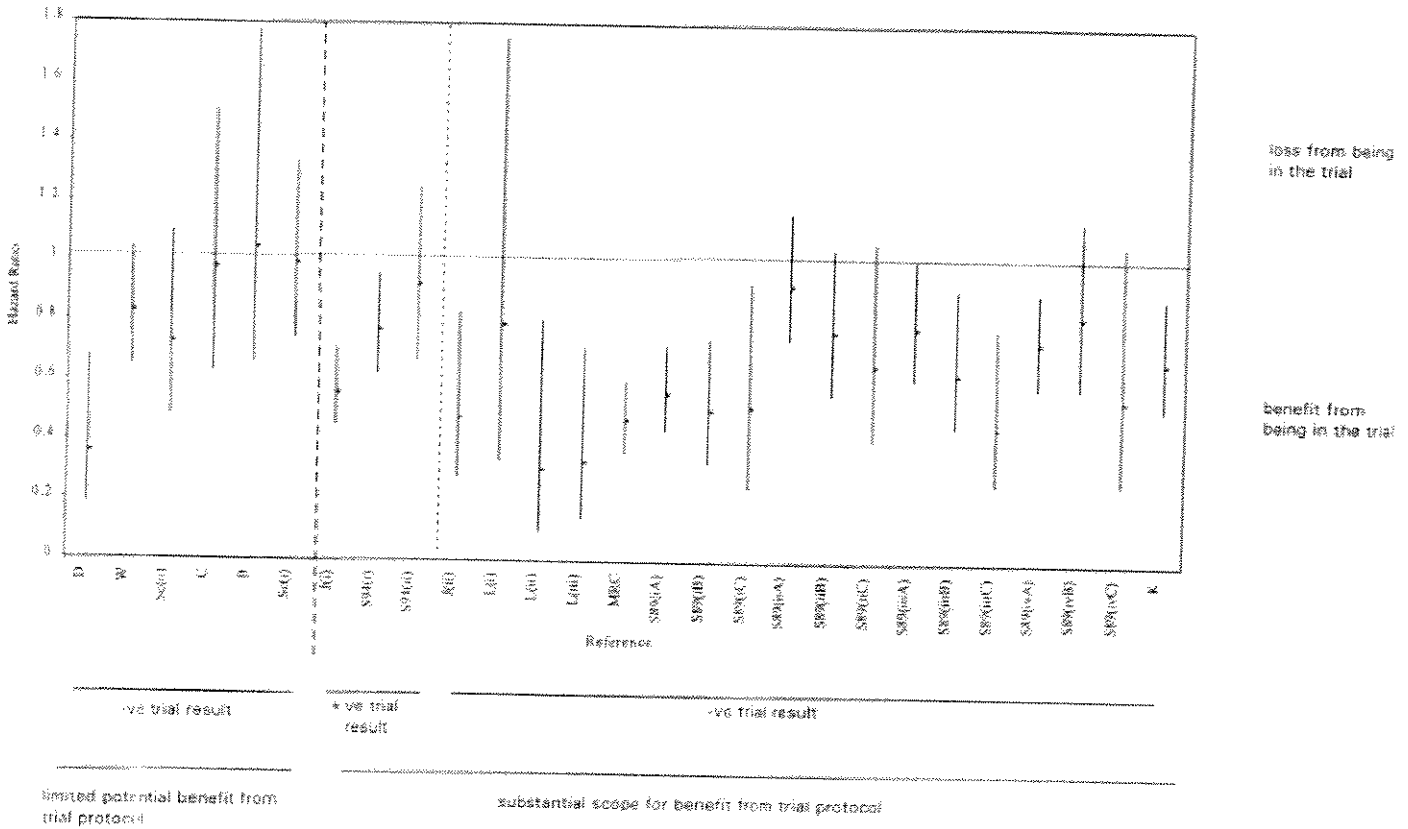
Tabellen er hentet fra Parmars artikkel(48). Tabellen viser hvordan studiestørrelsen påvirker konfidensintervall og p-verdi ved at konfidensintervallet blir smalere og p-verdien sikrere.

**Tabell 3**

Type bevis	Generell medisin		Kirurgi	
	Ellis et al (59)	Gill et al (60)	Howes et al(61)	Kenny et al (62)
Randomiserte kontrollerte studie bevis	53	31	24	11
Overbevisende ikke-eksperimentelle bevis	29	50	71	66
Inget bevis	18	19	5	23

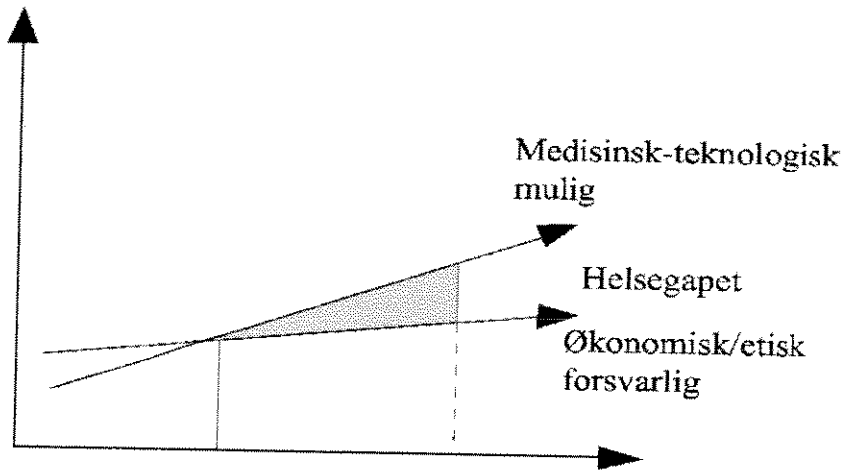
Tabellen er hentet fra Sauerland et al (31). Den viser hvor stor grad av bevis og dokumentasjon man har innen generell medisin og kirurgi for behandling. Angitt i prosent.

Figur 1.



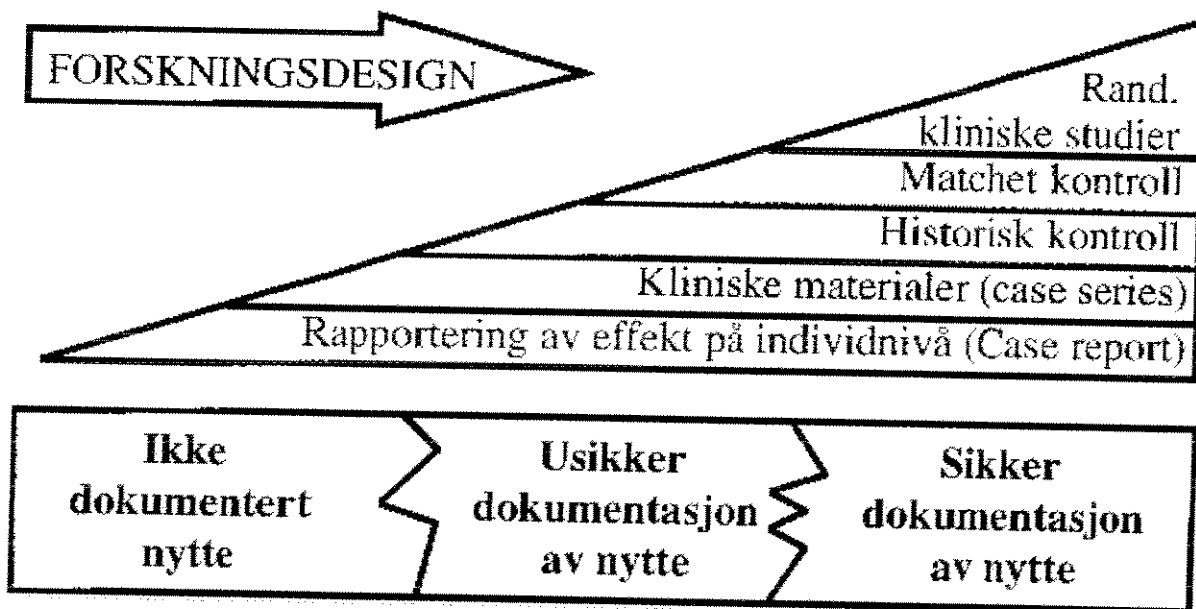
Hentet fra Braunholz et al (37). Er en figur som viser en gjennomgang av 21 artikler som viser en positiv korttids effekt av å være med i kliniske studier.

Figur 2



Hentet fra Nou 1997:18 Prioritering på ny (56). Den viser helsegapet og behovet for å prioritere innen helsevesenet som har begrensede ressurser.

Figur 3



Hentet fra NOU 1997:18 Prioritering på ny (56). Figuren viser hvordan graden av sikker dokumentasjon av nytte er avhengig av studiedesign.

**Figur 4**

**Spørsmål som forskningsprotokollen skal besvare**

Hva er bakgrunnen for at vi ønsker å utføre undersøkelsen?  
Hva er gjort tidligere?  
Hva er formålet med undersøkelsen?  
Hvordan skal vi gå frem?  
Hva eller hvem skal undersøkelsen omfatte?  
Hva skal telles, måles, observeres eller forstås?  
Hvordan skal dataene innhentes?  
Hvordan skal dataene lagres, kvalitetskontrolleres og analyseres?  
Hvordan skal prosjektarbeidet styres og organiseres?  
Hva slags personell, utstyr og andre ressurser er nødvendig?  
Hva vil undersøkelsen koste, og hvordan skal vi skaffe penger?  
Hvordan skal tidsplan og fremdrift være?  
Hvordan og for hvem skal resultatene presenteres?  
Hvilke godkjenninger er nødvendige?  
Er undersøkelsen etisk forsvarlig?

**Forskningsprotokollens disposisjon**

Tittel  
Bakgrunn  
Formål og problemstilling  
Design  
Utvalg  
Variabler  
Datainnsamling  
Analyse  
Prosjektorganisasjon  
Personell, utstyr, ressurser  
Kostnader og finansieringsplan  
Tidsplan  
Publisering  
Etikk

Hentet fra Bakketeig og Magnus (58). En oversikt over hva en forskningsprotokoll skal besvare og forslag til disposisjon.

**Bildene på forsiden** er hentet fra:

[www-edlib.med.utah.edu/WebPath/HEMEHTML/HEMEIDX.html#1](http://www-edlib.med.utah.edu/WebPath/HEMEHTML/HEMEIDX.html#1)

[www.kalpesh.itgo.com/MDS.html](http://www.kalpesh.itgo.com/MDS.html)

## Referanser

- 1 Evensen SA, Brinch L, Tjønnfjord GE, Wisløff F. Blodsykdommer 5. utgave. Universitetsforlaget 1999.
- 2 The Nordic Myeloma Study Group. Protokoll #5/94 "Intensivbehandling med stamcellestøtte av yngre myelompasienter".
- 3 Selby, et al. Multiple myeloma treated with high dose intravenous melphalan. Br. J Haemat. 1987; 55: 55-62.
- 4 Barlogie, et al. High dose chemoradithery an autologous bone marrow transplantatinon for resistant multiple myeloma. Blood 1987; 70: 869-872.
- 5 S.Lenhof, M. Hjorth, E. Holmberg et al. Impact on survival of high-dose therapy with autologous stemcell support in patients younger than 60 years with newly diagnosed multiple Myeloma: A population-based study. Blood 2000; 95: 7-11.
- 6 M.Hjorth, S.Lenhoff et al. Högdos melfalan med stamcellsstöd nu etablerad terapi vid myelom. Läkartidningen 2000; 97: 4585-4592.
- 7 F. Wisløf, J.Tangen et al. Høydosebehandling med autolog stamcellestøtte- endelig et fremskritt ved myelomatose. Tidsskr Nor Lægeforen. 2001;121: 2402-2406.
- 8 The Nordic Myeloma Study Group. Protokoll #7/1998 "Intensivbehandling av yngre myelompasienter".
- 9 The Nordic Myeloma Study Group. Protokoll #8/00 "Effekten av i.v. Bisfosfonat på helse relatert livskvalitet og skjelettmorbiditet ved nydiagnostisert behandlingkrevende myelomatose. En prospektivt randomisert dobbeltblindet dose-effective studie med cost-utility analyse".
- 10 Coleman RE, Purohit OP. Osteoclast inhibition for the treatment of bone metastases. Cancer Treat. Rev. 1993; 19: 79-103.
- 11 Belch AR, Bergsagel DE, Wilson K, et al. Effect of daily etidronat on the osteolysis of multiple myeloma. J.Clin Oncol 1991; 9: 1397-1402
- 12 Lahtinen R, Laakso M, Palva I, et al. Randomised, placebo-controlled multicenter trial of clodronat in multiple myeloma. Lancet 1992; 340: 1049-1052
- 13 McCloskey EV, MacLennan ICM, et al. A randomised trial of the effect of coldronat on skeletal morbidity in multiple myeloma. Brit J Haematol 1998; 100: 317-325

- 14 The Nordic Myeloma Study Group. Protokoll # 11/00 "VAD mot CY+Dex.
- 15 The Nordic Myeloma Study Group. Protokoll #12/00 "Thaliodomid til pasienter med ubehandlet myelomatose".
- 16 Singhal S, Metha J, Desikan R, et al. Antitumor Activity of Thaliodomide in refractory multiple myeloma. *N Engl J Med* 1999; 341: 1565-1571.
- 17 Juliuson G, Celsing F, Tureson I, et al. *Br J Haematol* 2000; 109: 89-96
- 18 Moehler TM, Egerer G, Punzel M, et al. Thaliodomide Plus Ced-Chemotherapy in patients with poor prognosis multiple myeloma. *Blood* 94(suppl.1); abstr. 548, 1999.
- 19 Sabir T, Raxa S, Anderson L, et al. Thaliodomide is effective in the treatment of recurrent refractory multiple myeloma. *Blood* 94(suppl. 1); abstr 549, 1999.
- 20 The Nordic Myeloma Study Group. Protokoll #10/99 "Thaloidonid til pasienter med avansert myelomatose".
- 21 The Nordic Myelodysplastic Syndrome Group. "Maintenace treatment with Decitabin vs consolidation treatment with high-dose chemotherapy in pasients with advanced MDS and MDS-AML, who have obtained CR with intensive chemotherapy"
- 22 De witte T, Aul C, Selleslag D, et al. Intensive chemotherapy followed by allogenic or autologous stemcell transplantation for pasients with myelodysplastic syndromes and acute myeloid leukemia following MDS. *Blood* 2001 Oct15; 98(8): 2326-2331.
- 23 Bernell P, Hast R, Hellstrom Lindberg E, et al. GM-CSF in combination with chemotherapy does not improve the remission rate in acute leukemia after myelodysplastic syndromes- a randomized study. *Blood* 1999, suppl. 1, abs. 2536.
- 24 Novartis®. "En fase 3 studie av STI571 mot Interferon- $\alpha$  kombinert med Cytarabin (Ara-C) til pasienter med nydiagnostiert tidligere ubehandlet Philadelphia positivt kromosom kronisk myelogen leukemi i kronisk fase(KML-CP)"
- 25 Nowell PC, Hungerford DA. A minute chromosome in chronic granulocytic leukamia. *Science* 1960; 132: 1497-1501.
- 26 Rowley JD. A new consistant abnormality in chronic myelogenous leukemia identified by quinacrine, fluorescence and Giemsa staining. *Nature* 1973; 243: 290-293.
- 27 Buchdunger E, Muller M, et al. A potent protein-tyrosin kinase inhibitor which inhibits PDGF receptor and c-Kit mediated in vitro signal transduction and in vivo tumor growth. Submitted
- 28 Druker BJ, Tamura S, Buchdunger E, et al. Effects of a selective inhibitor of the Abl tyrosine kinase on the growth of Bcl-Abl positive cells. *Nature Med* 1996; 2: 561-566.
- 29 Antes G, Galandi D, Bouillon B. What is evidence-based medicine?. *Langenbeck's Arch. Surg.* 1999; 384: 409-416.

- 30 Sackett DL, Rosenberg WMC, Gray JAM, et al. Evidence based medicine: what it is and what it isn't. *BMJ*; 312: 71-72.
- 31 Sauerland S, Lefering R, Neugebauer EAM. The pros and cons of evidence-based surgery. *Langenbeck's Arch. Surg.* 1999; 384: 423-431.
- 32 Feinstein AR. An additional basic science for clinical medicine: II The limitations of randomized trials. *Ann Intern Med.* 1983; Oct;99(4): 544-550.
- 33 Bonchek LI. Randomised trials of new procedures: problems and pitfalls. *Heart* 1997; Dec;78(6): 535-536.
- 34 Kaptchuk TJ. The double-blind, randomized, placebo-controlled trial. Gold standard or golden calf?. *J of Clin. Epidemiology* 2001; 54(6): 541-549.
- 35 Dahan R, Caulin C, Figea L, et al. Does informed consent influence therapeutic outcome? A clinical trial of the hypnotic activity of placebo in patients admitted to hospital. *Br. Med. J* 1986; 293: 363-364.
- 36 Bergmann JF, Chassany O, Gandiol J, et al. A randomized clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. *Clin. Trials Meta-anal* 1994; 29: 41-47.
- 37 Braunholtz DA, Edwards SJL, Lilford RJ. Are randomized clinical trials good for us (in short term)? Evidence for a "trial effect". *J of Clin. Epidemiology* 2001; 54(3): 217-224.
- 38 Angell M, Kassirer JP. Clinical Research --What should the Public Believe? *New Engl. J Med* 1994; 331: 189-190.
- 39 Grover FL, Shroyer AL. Clinical Science research. *J Cardiovasc. Surg.* 2000; Apr;119(4 pt 2): 11-21.
- 40 Edwards SJL, Lilford J, Hewison J. The ethics of randomised controlled trials from the perspectives of patients, the public, and healthcare professionals. *BMJ* 1998; 317: 1209-1212.
- 41 Williams CJ, Zwitter M. Informed consent in European multicenter randomized clinical trials - are patients really informed? *Eur. J Cancer* 1994; 30: 907-910.
- 42 Taylor KM, Kelner M. Interpreting physician participation in randomized clinical trials: the physician orientation profile. *J Health Soc Behaviour* 1987; 28: 389-400.
- 43 Harth SC, Thong YH. Parental perceptions and attitudes about informed consent in clinical research involving children. *Soc. Sci. Med* 1995; 40: 1573-1577.
- 44 Lynoe N, Sandlund M, Dahlqvist G, Jacobsson L. Informed consent: study of the quality of information given to participants in a clinical trial. *BMJ* 1991; 303: 610-613.
- 45 Tindall B, et al. Effects of two formats of informed consent on knowledge amongst persons with advanced HIV disease in a clinical trial of didanosine. *Patient Educ Counselling* 1994; 24: 261-266.



- 46 Twomey JG. Investigating pedantic HIV research ethics in the field. *Weston J Nurs. Res.* 1994; 16: 404-413.
- 47 Minouge BP, Palmer-Ferandez G, et al. Individual autonomy and the double-blind controlled experiment: the case of the desperate volunteer. *J Med Philosophy* 1995; 20: 43-55.
- 48 Parmar KBM. Pitfalls and biases in the reporting and interpretation of the results of clinical trials. *Lung Cancer* 1994; 10(suppl. 1): 143-150.
- 49 Hjorth M, Holmberg E, Rodjer S, et al. Physicians' Attitudes Toward Clinical Trials and Their Relationship to Patient Accrual in a Nordic Multicenter Study on Myeloma. *Contolled Clinical Trials* 1996; 17: 372-386.
- 50 Pocock SJ. *Clinical trials: A practical approach.* John Wiley and Sons, New York, 1983.
- 51 Shuster JJ, Krischer JP, Boyett JM. Ethical issues in cooperative cancer therapy trials from a statistical viewpoint. I. A general overview. *Am j Pediatr Hematol/oncol.* 1985; 7: 57-63.
- 52 Hjorth M, Holmberg E, Rodjer S, et al. Patient accrual and quality of participation in a multicentre study on myeloma: a comparison between major and minor participating centres. *Br. J of Heamatology* 1995; 91: 109-115.
- 53 Kerridge I, Lowe M, Henry D. Ethics and evidence based medicin. *BMJ* 1998; 316: 1151-1153.
- 54 Capbell EJM. *Clinical science.* Ciba 1976; 44: 41-52.
- 55 Innstilling fra rådgivende organ nedsatt av Helsedirektoratet. Benmargstransplantasjon i Norge. Helsedirektoratets utredningsserie 6-92, IK-2381.
- 56 Norges offentlige utredninger(NOU). Prioritering på ny. Gjennomgang av retningslinjer for prioriteringer innen norsk helsetjeneste. Sosial og helsedepartementet NOU 1997: 18.
- 57 Lindbæk M, Skovlund E. Kontrollerte kliniske forsøk- jakten på sann effekt av behandling. *Tidsskr Nor Lægeforen.* 2002; 122: 2631-2635.
- 58 Bakketeig LS, Magnus P. Hva er en forskningsprotokoll og hvorfor er den nødvendig?. *Tidsskr Nor Lægerforen* 2002; 25: 2464-2467.
- 59 Ellis J, Mulligan I, Rowe J, Sackett DL, et al. Inpatient general medicin is evidence based. *Lancet* 1995; 346: 407-410.
- 60 Gill P, Dowell AC, Neal RD, et al. Evidence based general practice: a retrospective study of interventions in one training practice. *BMJ* 1996; 312: 819-821.
- 61 Howes R, Chagla L, Thorpe M, et al. Surgical practice is evidence based. *Br J Surg* 1997; 84: 1220-1223.

62 Kenny SE, Shankar KR, Rintala R, et al. Evidence-based surgery: interventions in a regional paediatric surgical unit. *Arch Dis Child* 1997; 76: 50-53.