

1. Inledning

Uppdraget

Syfte

För Statens beredning för medicinsk utvärdering (SBU) har det primära målet med rapporten om cytostatikabehandling vid cancer varit att bedöma om tillgänglig vetenskaplig litteratur ger stöd för att behandla ett antal viktiga tumörtyper med cytostatika (cellhämmande medel). Syftet har också varit att fastställa hur starkt detta stöd är. De olika typer av cytostatikabehandling som används rutinmässigt (se nedan) inkluderades i varierande omfattning efter tumörtyp. Bedömningen gällde förhållandena i svensk sjukvård. Projektet omfattade endast konventionell cytostatikabehandling, i standarddosering eller i högdos med olika typer av understöd, och avsåg inte att täcka andra nyligen upptäckta behandlingsprinciper som i och för sig kan komma att inkluderas under begreppet cytostatika. Dit hör t ex läkemedel som hämmar bildning av nya blodkärl och av metastaser samt sådana som förebygger cancer.

Litteraturgenomgången avsåg att besvara två huvudfrågor:

1. Hur starkt är det vetenskapliga stödet för att cytostatika är till nytta i olika behandlingssituationer vid de olika tumörtyperna?
2. Om ett vetenskapligt stöd för nytta kan konstateras, hur stora är behandlingseffekterna?

I ett andra led avsåg rapporten att kartlägga om den nuvarande användningen av cytostatika i svensk sjukvård överensstämmer med vad som enligt litteraturgenomgången har vetenskapligt stöd. Detta gjordes genom en prospektiv (framåtsyftande) undersökning av användningen av cytostatika under fyra veckor hösten 1997 i Stockholms- och Uppsala/Örebroregionen. Resultaten från denna kartläggning jämfördes sedan med slutsatserna från litteraturgenomgången.

Litteraturöversikten omfattade tumörformer för vilka tre undersökningar har visat att användningen av cytostatikabehandling har ökat under det

senaste decenniet. Den första undersökningen gjordes inför en konsensuskonferens organiserad av Medicinska forskningsrådet (MFR) och Sjukvårdens planerings- och rationaliseringsinstitut (Sprri) i Stockholm 1990 [3]. Den andra var en uppföljning 1993 för att studera om rutinerna hade påverkats av konsensusuttalandet [10]. Den tredje undersökningen gjordes 1996 under planeringen av projektet. Enkäter skickades då till alla onkolog-, kirurg- och internmedicinska kliniker, inklusive hematologiavdelningar och lungkliniker. I kirurgklinikerna inkluderades också urologi- och gynekologikliniker. Svaren visade att användningen av cytostatikabehandling ökat och att den bedömdes som mer värdefull än tidigare vid icke-småcellig lungcancer, magsäckscancer, bukspottkörtelcancer, tjock- och ändtarmscancer samt cancer i urinblåsan.

Litteraturgenomgången omfattade också de tumörgrupper som har det största antalet patienter som behandlas med cytostatika i Sverige, dvs bröstcancer och elakartade blodsjukdomar. Av de senare medtogs endast de vanligaste, dvs akut myeloisk leukemi, kronisk lymfatisk leukemi, Hodgkin's sjukdom, aggressiva, (högmaligna) non-Hodgkin lymfom, huvudsakligen storcelliga B-cellslymfom, och indolenta (långsamt växande) non-Hodgkin lymfom av follikulär typ. Sammantaget utgör dessa diagnoser minst 80 procent av alla elakartade blodsjukdomar.

Slutligen togs även äggstockscancer med, dels därför att den ofta behandlas med cytostatika, dels därför att en ny, mycket dyr grupp cytostatika, taxanerna, preliminärt hade visat lovande resultat när projektet förbereddes och användningen av dem hade börjat öka i Sverige.

Ovanliga kliniska syndrom och ovanliga morfologiska subtyper uteslöts för alla tumörtyperna. Dessa undantag anges mer i detalj i varje diagnoskapitel. Behandlingar vid vilka cytostatika kombineras med strålbehandling i syfte att förstärka stråleffekten medtogs endast undantagsvis.

Rapportens innehåll

Cytostatika tillförs oftast intravenöst men i en del fall via munnen, så kallade systemiska behandlingar. Cytostatika används i flera olika situationer. Vid några få typer av avancerad cancer kan behandlingen utplåna i stort sett samtliga tumörceller och bota patienten. Cytostatika används också alltmer som tillägg till kirurgi eller strålning, så kallad adjuvant

behandling. Avsikten är att ta bort mikroskopisk sjukdom och att öka chansen till bot i tidiga stadier av cancer. Det blir också allt vanligare att ge cytostatika före kirurgi eller strålterapi, så kallad preoperativ och neo-adjuvant behandling. Målet är då att minska tumörvolymen så att lokal behandling ska bli mer effektiv respektive förbättra den totala tumörkontrollen. Slutligen används cytostatikabehandling palliativt (i syfte att lindra) vid avancerad cancer som inte går att bota. Då är målet att minska tumören, lindra symtomen och förlänga livet.

Cytostatika kan också tillföras lokalt när cancern växer i kroppens hålrum eller loko-regionalt, t ex till levern via leverartären. Syftet är att uppnå högre cytostatikakoncentrationer och bättre effekt.

Utvärderingen av det vetenskapliga underlaget för de olika cancerformerna redovisas i Volym II. De olika typer av cytostatikabehandling som tas upp i diagnoskapiteln är de som har bedömts vara i rutinmässig användning i svensk sjukvård, står i fokus för pågående utveckling, t ex högdosbehandling, eller tidigare varit föremål för stort utvecklingsintresse men nu används mer sällan.

Vid många tumörtyper kan cytostatika enbart lindra symptom och ge viss förlängning av överlevnaden, men inte bota sjukdomen. Det sker i stort sett alltid på bekostnad av mer eller mindre svåra biverkningar. Därför har frågan om livskvalitet (QoL=quality of life) under och efter cytostatikabehandling tilldragit sig stort intresse under det senaste decenniet. I de diagnosspecifika kapitlen framgår tydligt att livskvaliteten har studerats i mycket varierande omfattning. Många kapitel avhandlar dock detta tema och många slutsatser dras utifrån presenterade data. Eftersom det är metodologiskt svårt att mäta livskvalitet innehåller rapporten ett separat kapitel som sammanfattar och kommenterar detta område.

Rapporten innehåller inte någon systematisk översikt om biverkningar av cytostatika. I Volym I presenteras dock en relativt omfattande sammanfattning av denna aspekt på cytostatikabehandling. I Volym I redovisas också resultaten av den kartläggning av cancerbehandling med cytostatika i Sverige som genomförts inom projektets ram samt en jämförelse mellan det vetenskapliga underlaget och praxis.

Sjukvårdskostnader är ytterligare ett område som har fått ökat intresse i nuvarande ekonomiskt ansträngda situation. Det är svårt att mäta behandlingsskostnader, speciellt när de ska relateras till behandlingsvinster. Därför redovisas även detta område separat i kapitlet om kostnader och kostnadseffektivitet för cytostatikabehandling.

Arbetsprocessen

En nationell expert, i några fall flera experter, fick ansvaret för att söka, bedöma, sammanfatta och klassificera litteraturen för varje diagnoskapitel. Experten fick också i uppdrag att skriva en första version av kapitlet. Varje kapitel börjar med en introduktion som sammanfattar den nuvarande standardterapi för den aktuella tumörtypen. Därefter följer litteraturöversikten som är indelad i underavdelningar för olika typer av cytostatikabehandling. Varje underavdelning avslutas med en sammanfattning av vilka slutsatser som kan dras. I tabellform klassificeras och bedöms litteraturen inom respektive underavdelning med avseende på studietyp, vetenskaplig kvalitet och antal inkluderade patienter.

Avsikten var att behandlingsresultaten skulle beskrivas som punkttestimat för effekten samt att graden av statistisk signifikans skulle anges vid jämförelser. Men då de vetenskapliga originalrapporterna redovisar resultaten på olika sätt måste detta accepteras också i denna rapport trots att det försvårar jämförelser mellan studier och diagnoser. Målet var att redovisa behandlingsresultaten på samma detaljnivå för alla diagnoser. Detta gick inte heller att genomföra fullt ut, dels beroende på variation i resultatrapportering i litteraturen, dels beroende på författarnas preferenser.

Då första versionen av ett kapitel var färdig granskades den av en andra expert inom projektgruppen. Sedan skickades det till varje medlem i projektgruppen för kommentarer och ändringsförslag samt diskuterades vid återkommande möten inom gruppen. Kapitlen skickades därefter till internationella experter, i de flesta fall två stycken, inom det aktuella området. Några kapitel granskades även av medlemmar av SBU-Rådet. De mer allmänt hållna icke diagnosspecifika kapitlen skickades inte för internationell bedömning utan värderades och kommenterades inom projektgruppen och av experter inom SBU.

Hela denna granskningsprocess gav upphov till åtskilliga värdefulla synpunkter och kompletterande sakupplysningar, av vilka många inkluderades. Vissa kommentarer som uppfattades spegla enskilda uppfattningar eller som inte omfattades av utredningens målsättningar medtogs emellertid inte.

Kapitlen reviderades efter det att alla åsikter vägts samman, språkredigerades, diskuterades och godkändes slutligen av projektgruppen. Slutligen bedömdes den fullständiga rapporten av fyra externa experter samt av SBU-Rådet och SBU:s styrelse. Tillämpliga synpunkter inarbetades därefter i den föreliggande slutliga rapporten.

I allt väsentligt har processen fram till de slutliga kapitelversionerna syftat till att vara evidensbaserad och systematisk [8,13].

Litteratursökning

Litteratursökning, klassificering, tolkning och bevisvärdering utfördes i enlighet med riktlinjer från SBU [11].

Den medicinska databasen Medline samt i vissa fall Cancerlit användes för att söka kliniska studier att basera litteraturöversikten för varje tumörtyp på. Dessutom gjordes sökningar i Cochrane Library. Som sökord användes ”antineoplastic agents”, kombinerat med specifika begrepp för de olika cancerformerna.

Den tidsperiod som avsöktes varierade men omfattade minst de senaste 15 åren. För vissa tumörtyper var det viktigt att täcka hela tidsperioden – från 1950-talet – med cytostatikabehandling för att kunna bedöma den evidensbaserade grunden för t ex förlängd överlevnad.

Ytterligare studierapporter och studiedata identifierades av primärförfattarna och av de internationella granskarna via referenslistor i relevanta artiklar och i den ”grå” litteraturen [8,11], dvs i huvudsak konferenssammanfattningar. Databasen PDQ, utgiven av National Cancer Institute i USA, likväl som moderna utförliga onkologitextböcker användes också som källor för eventuella ytterligare studierapporter. I praktiken användes alltså flera sökstrategier [8]. Den slutliga versionen av denna rapport har

uppdaterats till att inkludera alla publikationer t o m 1998. Likaså medtogs rapporter om viktiga randomiserade studier som var kända 1998 men publicerades fram till och med september 2000.

Generellt uppfattade primärförfattarna att den datorbaserade litteratursökningen via de etablerade databaserna inte var så heltäckande som förväntat avseende erhållen nyckelinformation inom respektive fält. Det innebar att deras personliga kännedom om viktig litteratur, referenslistor i publicerade artiklar och den ”grå” litteraturen gav ytterligare information om viktiga studier. Inom den ”grå” litteraturen var mötessammanfattningar (abstracts) den viktigaste informationskällan. Abstracts accepterades om resultaten härrörde från etablerade forskargrupper, använd metodik bedömdes vara adekvat och om det kunde antas att resultaten snart skulle finnas tillgängliga i fullständiga rapporter. Trots de uppenbara svårigheterna att bedöma den begränsade informationen i ett abstract ansågs det oacceptabelt att helt utelämna studieresultat som endast hade publiceras på detta sätt. Skälet till detta är att resultaten från viktiga kliniska studier ofta rapporteras i denna form relativt lång tid före den slutliga publikationen. Så långt tidsramarna tillåtit har refererade abstracts ersatts med den fullständiga publikationen i den slutliga rapporten.

Klassificering och bedömning av litteraturen

I projektet användes SBU:s system [11] för hur vetenskapliga rapporter bör bedömas och klassificeras. Systemet består av två komponenter:

- klassificering av rapporten på basen av studietyp
- bedömning av rapportens vetenskapliga kvalitet för att vikta dess betydelse för slutsatserna.

Klassificeringen omfattade följande studietyper:

- *Meta-analyser (M)*. En statistisk syntes av flera randomiserade studier. Meta-analysen kan baseras på originaldata från de ingående studierna eller enbart på den information som publicerats.
- *Fas III studier (C)*. Prospektiva studier som randomiserar mellan olika behandlingsalternativ. Kan genomföras med deltagande av flera centra, så kallade multicenter-studier, eller vid en enskild institution. Kvaliteten hos denna studietyp är starkt beroende av studiestorlek likväl som av resultatens externa och interna validitet.

- *Fas III studier (P)*. Prospektiva, vanligen icke-randomiserade studier syftande till att tidigt kartlägga effekt och säkerhet för ny behandling. Kan omfatta ett mycket litet antal patienter.
- *Retrospektiva studier (R)*. Studier som analyserar patientdata vilka ofta har samlats in under lång tid, vanligen av en institution eller av en internationell databas som samlar information från olika centra. Behandlingen har ofta ändrats under studieperioden. Stora, välövervakade studier, som har samlat data enligt ett i förväg uppställt protokoll och har lång uppföljningstid och adekvat statistisk bearbetning, kan vara informativa.
- *Review (L)*. Översiktsartiklar baserade på litteratursammanställning eller kapitel i textböcker. Kan innehålla detaljerade redogörelser av ingående studieresultat men kan också vara mer allmänna.
- *Övriga artiklar (O)*. Denna kategori omfattar t ex fallbeskrivningar, brev till tidskrifter, ledarkommentarer och sammanfattningar från vetenskapliga möten. Djurförsök och metodbeskrivningar hör också till denna kategori.

Förutom klassificeringen av studietyp bedömdes också publikationernas vetenskapliga värde på en tregradig skala: hög, medel eller låg. Tabell 1 ger exempel på bedömningar av vetenskaplig kvalitet för de olika typerna av studier.

Studiekategori, bevisvärde och antal patienter i studierna anges inom parentes för varje referens i de diagnosspecifika kapitlen. Dessa karaktäristika summeras också i slutet av varje kapitel vilket exemplifieras i Tabell 2. I tabellen klassificeras och värderas alla de 1 496 vetenskapliga artiklar som legat till grund för föreliggande rapport. De omfattar sammanlagt 558 743 patienter.

Åtskilliga problem observerades vid bedömningen av den vetenskapliga litteraturen exempelvis svårigheten att fullständigt täcka varje område genom litteratursökning via databaser.

Ett annat problem gällde litteraturklassificeringen. Det var vanligen okomplicerat att placera studierna i rätt kategori, medan kvalitetsbedömningen tenderade att bli subjektiv. Med hänsyn till det stora

antalet rapporter att klassificera gjordes inga ansträngningar att kvalitets-säkra denna bedömning. Enskilda exempel som togs upp till diskussion antydde dock inte några avgörande skillnader i hur primärförfattarna bedömde kvaliteten.

Ett tredje problem var att ange hur många patienter som hade inkluderats i enskilda rapporter. Denna svårighet berodde på komplex studie-design, ofullständig information och upprepad rapportering för delvis identiska patientgrupper. Detta innebär att de angivna antalen patienter inte kan garanteras vara korrekt. Det viktigaste med denna bedömning är dock att ange den ungefärliga storleken av de patientgrupper på vilka slutsatserna vilar.

Tabell 1 Principer för litteraturklassificeringen.

Typ av studie	Vetenskaplig kvalitet = bevisvärde (1 = högt, 2 = medel, 3 = lågt)	
M	Noggrann återanalys av väldefinierade originaldata från alla studier.	Sammanfattning baserad på ett fåtal studier och/eller endast på rapporterade data.
C	Stora välmonitorerade multicenter studier som noggrant beskriver protokoll, patienter och metoder inklusive behandling. Patientdatamängd tillräckligt stor för att besvara frågeställningen.	Randomiserade studier med för få patienter och/eller för många behandlingsarmar för att ge tillräckligt underlag för statistisk bearbetning. Dålig dataredovisning, stort patientbortfall. Dålig beskrivning av given behandling.
P	Väldefinierad frågeställning, tillräcklig mängd patientdata, adekvat statistisk bearbetning.	Så kallade "quick-and-dirty" studier.
R	Data från stora grupper konsekutivt följda patienter som är väldefinierade och där data analyserats med adekvat statistisk metodik (t ex multivariatanalys, fall-kontroll jämförelse etc). Lång uppföljningstid.	Begränsad mängd patientdata, dåligt definierade data, kort uppföljningstid, inadekvat statistisk metodik.
L	Noggrann litteraturgenomgång, välbeskrivna patientdata presenterade i tabellform.	Rapporter med inkomplett referensgivning och slutsatser med svagt stöd.
O	Textbokkapitel av hög kvalitet.	

Tabell 2 Artiklarna som medtagits i de diagnosspecifika kapitlens referenslistor klassificerades och graderade enligt nedanstående tabell (antal studier/antal patienter).

Typ av studie	1 = högt	2 = medel	3 = lågt	Totalt
M	27/142 118	11/52 546	1/18 000	39/212 664
C	309/125 664	267/47 698	220/43 193	796/216 555
P	109/15 709	194/11 490	133/6 120	436/33 319
R	63/42 203	45/20 944	14/10 563	122/73 710
L	44/15 048	18/7 447	1/–	63/22 495
O	32/–	8/–	–/–	40/–
Totalt	584/340 742	543/140 125	369/77 876	1 496/558 743

Att dra slutsatser ifrån publicerade data

Det finns modeller för att enhetligt formulera bevisvärdet och styrkan hos rekommendationer baserade på litteraturkvaliteten [11]. I föreliggande rapports diagnoskapitel används dessa principer även om varje kapitels slutsatser inte presenteras i tabellform utan som fri text. I Volym II, Section 1 tillämpas och beskrivs ett system för formulering av rekommendationer utarbetat av American Society of Clinical Oncology. Detta system bildade också basen för att jämföra det vetenskapliga underlaget med praxis (Kapitel 8).

Ett problem när slutsatser och rekommendationer ska utformas på basen av ett stort antal vetenskapliga rapporter av varierande kvalitet är hur man överför kvalitetsbedömningen av en studie till en vikt när slutsatsen formuleras. Det gäller inte minst när motsägande data har presenterats, vilket varit vanligt. Även om man har försökt att formalisera den tyngd en enskild studie ska tillmätas i slutsatsen ingår fortfarande subjektiva bedömningar. Då all litteratur som slutsatserna i föreliggande rapport grundas på ger detaljerad information om studietyp och kvalitet är det i princip möjligt för den intresserade läsaren att göra en egen bedömning.

I flera av diagnoskapitlen refereras till data från meta-analyser. Vidare bildar dessa ofta basen för de slutsatser som formuleras i dessa kapitel, inte enbart om kvalitativa utan även om kvantitativa behandlingseffekter. Även om meta-analysen är en kraftfull metod är skillnader mellan meta-analyser och stora högkvalitativa prospektiva, randomiserade studier inte ovanliga. Det gäller framför allt hur stor effekt en behandling har, men också dess riktning [15,20]. Sålunda bör även meta-analyser tolkas försiktigt. Avgörande slutsatser för ett terapiområde kan alltså med fördel också baseras på omdömesgilla värderingar av resultat från enskilda noggrant gjorda och kontrollerade studier.

Diagnoskapitlen

Rekommendationer om cytostatikabehandling baseras ofta på litteratur med högt bevisvärde. Inom många områden, även sådana som omfattar många patienter och resurskrävande behandlingar, är emellertid förvånansvärt få kontrollerade och bekräftade data tillgängliga. Frånvaron av bra data om betydelsen av cytostatikabehandling för överlevnad och livs-

kvalitet vid metastaserad bröstcancer och om nyttan av högdoskemoterapi med stamcellsstöd vid flertalet tumörtyper utgör exempel. Det medför med nödvändighet att läkarens rekommendationer om behandling till en enskild patient påverkas av mer eller mindre välgrundade subjektiva överväganden. Det var ibland initialt möjligt att spåra sådana överväganden också i kapitlens sammanfattningar och slutsatser för denna rapport. Efterhand bearbetades emellertid slutsatserna i projektgruppen med avsikt att de skulle grundas på studier av tillräcklig omfattning och kvalitet.

Det är uppenbart att datakvaliteten skiljer sig avsevärt mellan olika diagnoser men också att traditioner kan påverka tolkningen av data. Vid adjuvant behandling av bröstcancer och vid behandling av metastatisk icke-småcellig lungcancer finns nu data från ett flertal högkvalitativa randomiserade studier. Däremot måste slutsatserna vid indolenta lymfom och kronisk lymfatisk leukemi ofta baseras på data från fas II-studier. Terapeutisk nihilism vid en diagnos verkar kunna ge upphov till en mer konservativ bedömning än inom traditionellt mer behandlingsinriktade områden, vilket antyds då t ex cancer i bukspottkörteln jämförs med kronisk lymfatisk leukemi. Sådana skillnader mellan kapitlen är svåra att undvika men ett försök gjordes att väga in detta när de övergripande slutsatserna formulerades. Vidare är slutsatserna i de enskilda kapitlen ofta ganska allmänt hållna och tar inte direkt upp svårigheterna att översätta resultat från studier till den kliniska rutinsituationen så som diskuterades ovan.

Återgivningen av data om livskvalitet i diagnoskapitlet följer i flertalet fall den tolkning som gjorts i de publicerade studierapporterna.

Diskussioner inom projektgruppen ledde till en mer kritisk bedömning av metodiken för att mäta livskvalitet vilket framgår av det särskilda kapitlet om sådana mätningar (Kapitel 5). Livskvalitetsavsnitten i de diagnos-specifika kapitlen bör läsas i ljuset av de metodologiska svårigheter som beskrivs i kapitlet om livskvalitet.

Cytostatikabehandling i ett framtida perspektiv

Vissa framtidsaspekter på utvecklingen av cytostatika för cancerbehandling diskuteras i Kapitel 2 ”Vad innebär cytostatikabehandling?”. Här ska bara en del mer generella aspekter tas upp.

Kunskapen om genetiska faktorer som orsakar cancer och får den att växa har exploderat under det senaste decenniet. Det finns stora förhoppningar om att denna kunskap också ska öppna för helt nya behandlingar som direkt påverkar olika genetiska defekter. Detta område, dvs genterapi, befinner sig redan i sin tidiga kliniska utveckling men det är ännu för tidigt att bedöma huruvida dessa behandlingsprinciper kommer att revolutionera cancerbehandlingen. Tills detta är känt kommer nuvarande icke-specifika cytostatika fortsätta att vara ett huvudinslag i behandlingen av avancerad cancer, liksom i flera adjuvanta terapier. Även om genterapi och andra tillvägagångssätt visar sig användbara kanske cytostatika förblir en viktig del av cancerbehandlingen som en komponent i kombinationsbehandling.

Ett annat högtintressant område som håller på att utvecklas och grovt sett kan inkluderas under rubriken cytostatikabehandling är användningen av en brokig samling substanser som t ex antas hämma bildning av nya blodkärl och av metastaser eller som gör tumörtillväxten långsammare. Dessa substanser är nu i olika faser av kliniska studier. Läkemedel baserade på dessa principer kommer troligen inte att kunna ta bort tumörerna helt utan snarast framkalla ”tumörvila”. Det innebär en konceptuell förändring genom att cancer då kan förvandlas till en kronisk sjukdom förenlig med långtidsöverlevnad. En sådan utveckling kommer att kräva att vår syn på cancer ändras, men också att studier utformas på ett nytt sätt såväl i fråga om definition av lämplig dos som bedömning av effekt [9]. Även i detta fall kan den lämpligaste behandlingen vara att kombinera de nya läkemedlen med cytostatika med avsikt att tumörens volym ska minska.

En annan intressant utveckling är behandling med så kallade minitransplantationer där man kombinerar helkroppsstrålbehandling i låg dos eller icke-ablativ cytostatikabehandling (cytostatikabehandling som inte slår ut benmärgen) med tillförsel av stamceller för vita blodkroppar från en besläktad givare. Syftet är att åstadkomma en immunologisk reaktion riktad mot tumörcellerna.

Mot bakgrund av den pågående högtintressanta utvecklingen och i ljuset av den ofta svåra och långsamma vägen från att bevisa en farmakodynamisk princip till att etablera en terapi verkar det rimligt att cytostatika kommer att kvarstå som en nödvändig komponent i cancerbehandlingen ytterligare en tid.

Utvärdering av cancerstudier

Utvärderingsmått

En klinisk studies utvärderingsmått kan definieras som sådant som kan mätas för att uppfylla studiens målsättningar. Utvärderingsmått för effekt ska direkt eller indirekt spegla meningsfull nytta för patienten [9]. Ett stort antal slutvariabler används för att mäta effekt, som ofta kan uttryckas på flera olika sätt. Det kan därför vara svårt att jämföra olika rapporter och därför har olika sätt att uttrycka t ex överlevnadsvinst inom och mellan olika kapitel fått accepteras i föreliggande rapport.

Då det är viktigt att förstå de vanligaste slutvariablerna i denna rapport sammanfattas de nedan. Därefter följer en kritisk genomgång av deras innebörd. Det bör observeras att även om innebörden av dessa slutvariabler ofta anges som ”standard” används ibland olika definitioner [19]. Idealet är alltså att varje studierapport definierar innebörden av de slutvariabler som används.

- *Andel tumörrespons* – andel av patienterna vilkas tumörer svarar på cytostatikabehandlingen, dvs minskar signifikant i volym (se vidare nedan). Tumörrespons bör numera bekräftas vid en andra utvärdering efter minst fyra veckor. Utvärderingen av tumörrespons bör helst vara objektiv, dvs i princip grundas på radiologiska metoder. Bedömning av tumörrespons enbart utifrån förändringar i tumörmarkörer är ännu inte accepterat.
- *Tid till tumörrespons* – tid från behandlingsstart till objektivt observerad tumörrespons.
- *Duration av tumörrespons* – tid från behandlingsstart eller observerad tumörrespons till observation av återfall eller tillväxt av tumören (tumörprogress).
- *Progressionsfri överlevnad* – tid från behandlingsstart till tidpunkten för observation av sjukdomsprogress. Likvärdigt med tid till progress. På populationsnivå uttrycks detta utvärderingsmått ofta som median-tid för progressionsfri överlevnad och/eller som andelen progressionsfria patienter vid en specifik tidpunkt.

- *Sjukdomsfri överlevnad* – vid adjuvant behandling (se vidare Kapitel 2) tid från behandlingsstart till tumörens återkomst (relaps). På populationsnivå uttrycks detta utvärderingsmått ofta som mediantid för sjukdomsfri överlevnad och/eller som andelen sjukdomsfria vid en specifik tidpunkt. Vid avancerad sjukdom tid från tumörens försvinnande till dess återkomst.
- *Total överlevnad* – tid från behandlingsstart till död. På populationsnivå uttrycks detta utvärderingsmått ofta som mediantid för total överlevnad och/eller som andelen överlevande vid en specifik tidpunkt.
- *Symtomförbättring/stabilisering* – t ex andelen patienter med behandlingsrelaterad förbättring av specifika symtom eller tid till uppkomst eller försämring av ett sjukdomsrelaterat symtom. Detta utvärderingsmått kan mätas och rapporteras på flera olika sätt.
- *Förbättring av livskvalitet (QoL)* – patientens värdering, genom olika typer av frågeformulär, av behandlingsrelaterad förbättring av faktorer som antas spegla livskvalitet. Rapporteras vanligen som förändring över tid i olika variablers skalmått.
- *Tolerans* – frekvensen (per patient och/eller behandlingscykel) och graden av biverkningar, med specificering av de som bedömts som allvarliga. Graden uttrycks i en etablerad toxicitetsskala och biverkningarnas relation till behandlingen anges. Detaljbeskrivning av viktiga biverkningar avseende tidpunkt när de inträffade, varaktighet, reversibilitet, utseende, dosberoende, behandling osv, ger ytterligare information. Uppkomst på lång sikt av sekundära maligna tumörer, annan behandlingsrelaterad sjukdom och försämrade funktionsnivå kan också inkluderas även om sådana aspekter vanligen behandlas i separata undersökningar av långtidseffekter.

Vissa av dessa slutvariabler, t ex andelen tumörrespons och progressionsfri överlevnad, kan betraktas som surrogatvariabler, dvs något som är enklare att mäta än det sanna utvärderingsmålet som den antas avspegla (se vidare nedan).

Med beaktande av antalet utvärderingsmått som används och antalet statistiska varianter av datapresentation är det uppenbart att studiedata

kan redovisas på många olika sätt. Tolerans inkluderas nästan alltid och presenteras vanligen i enlighet med principerna ovan. De vanligast förekommande principerna för att rapportera effekt exemplifieras nedan. Vid adjuvant behandling är sjukdomsfri och total överlevnad de viktigaste slutvariablerna. De uttrycks ofta som andelen patienter i olika behandlingsarmar som är sjukdomsfria och lever vid en specifik tidpunkt, t ex fem år efter behandlingen. Då två behandlingar jämförs bör även hela överlevnadskurvorna bedömas och skillnader anges, t ex med logrank statistik och hazardkvot med konfidensintervall. Långtidsnyttan av en behandling jämfört med en annan kan uttryckas som absolut eller relativ behandlingseffekt.

En behandlingsstudie vid tidig bröstcancer med två grupper om vardera 1 000 patienter, där den ena gruppen får adjuvant behandling, den andra inte, kan tjäna som exempel på hur effekterna redovisas. Om total överlevnad efter fem år i kontrollgruppen är 60 procent och i behandlingsgruppen 75 procent är den absoluta överlevnadsvinsten 15 procent, dvs 150 patienter. Den relativa behandlingseffekten kan antingen uttryckas som 25 procent ($150/600$) ökad överlevnad eller som 37,5 procent ($150/400$) minskad dödlighet. Antalet patienter som behöver behandlas, dessvärre ett ganska lite använt mått, för att en ska ha räddats till livet efter fem år blir 6,7 (1/absolut riskreduktion). Det kan också sägas att 600 patienter i adjuvantgruppen behandlas i onödan eftersom de botades med enbart kirurgi.

I den palliativa situationen är andelen tumörrespons fortfarande ett vanligt utvärderingsmått. Skillnader i progressionsfri och total överlevnad, uttryckt i vinst i medianöverlevnad och hazardkvoter, som baseras på överlevnadskurvor kan emellertid anses bättre reflektera patientnyttan. Vidare rapporteras ofta skillnader över tid i livskvalitetsmått medan effekt på symtom anges som andelen patienter som uppnår en viss grad av lindring eller som mediantid till uppkomst eller försämring av specifika symtom.

Vid preoperativ cytostatikabehandling slutligen används inte bara sjukdomsfri och total överlevnad som utvärderingsmått utan ofta också andelen patienter som blivit operabla eller andelen som är möjliga att operera i syfte att bota, dvs att all synlig tumör avlägsnas.

En kritisk bedömning av utvärderingsmått

Tumörrespons har sedan länge definierats som en påtaglig, minst 50-procentig, minskning av summan av tumörvolymerna och fastställs oftast genom objektiv bedömning med radiologiska metoder. Det är vanligt att man delar upp tumörrespons i partiell respons (PR), dvs en påtaglig minskning men inte fullständigt försvinnande av tumören, och komplett respons (CR), dvs att tumören helt försvunnit. Även om responskriterier i studierapporter anges som standard är det nödvändigt att de ytterligare definieras, t ex att de anges enligt WHO [16].

Vidare ska principer för val av vilka delar av den totala tumörmassan man mäter respons på och krav på responsduration anges. Definitioner av respons och metoder för responsbedömning varierar fortfarande mellan olika studieprotokoll. Det finns också en variation över tid i det att antalet tumörresponser vanligen är lägre i moderna studier än i de tidigare. Det finns flera anledningar till detta; de radiologiska metoderna är nu mer exakta, t ex datortomografi istället för slätröntgen eller scintigrafi, det ställs krav på att tumörrespons ska bekräftas och, i synnerhet, att det görs en oberoende värdering av responsdokumentationen.

Principerna för bedömning av tumörremission är för närvarande under förnyad utvärdering inom en arbetsgrupp med deltagare från de nationella cancerinstituten i USA och Canada samt från den europeiska samarbetsorganisationen EORTC [1]. Arbetsgruppen föreslår sammanfattningsvis att tumörvolymen ska skattas på basen av endast mätning av tumörens längsta diameter, istället för som tidigare på produkten av den längsta diametern och dess vinkelräta diameter.

Vid elakartade blodsjukdomar kan man numera med molekylärmedicinska tekniker urskilja molekylär komplett remission, dvs man kan inte finna något uttryck av t ex en sjukdomsspecifik genförändring. Det är ännu oklart vilken betydelse denna fingergradering av den antitumorala effekten kan få i rutinsjukvården.

Även om en tumörrespons i vissa fall kan vara till direkt nytta för patienten, t ex genom minskning av det tryck tumören kan orsaka på viktiga vävnader, kan den inte betraktas som ett fullt validerat (giltigt) surrogatmått för patientnytta [4,9]. Vid t ex avancerad icke-småcellig lungcancer

tyder randomiserade studier på att relationen mellan tumörrespons och överlevnad är mycket svag. Vid avancerad tjock- och ändtarmscancer är sambandet höggradigt statistiskt signifikant men överlevnadsvinsten per definierad ökning av responsfrekvensen mycket liten [6]. Generellt kan sägas att sambandet inte har undersökts särskilt grundligt. Ur biologisk synvinkel motsäger ett dåligt samband mellan tumörrespons och överlevnad inte nödvändigtvis ett samband mellan tumörrespons och symptomlindring.

Det föreligger snarast konsensus bland erfarna cancerläkare om det positiva och starka sambandet mellan tumörrespons och nytta för patienten i form av t ex symptomlindring och förbättrat allmäntillstånd. Faktum är att symptomlindring rapporteras också utan objektiv tumörremission [7]. Även om dessa observationer kan antas vara korrekta föreligger metodologiska svårigheter att bedöma de palliativa effekterna och därför behöver också dessa samband studeras mer innan det går att dra några generella evidensbaserade slutsatser [2]. Tumörmarkörer är inte etablerade surrogatvariabler varken för överlevnad eller symptomlindring [2]. Tumörremission bör således i första hand betraktas som en indikator på antitumoral aktivitet som kan vara värd att studera ytterligare i jämförande studier där man använder slutvariabler som mer direkt mäter patientnyttan.

Det kan också vara avsevärda skillnader i hur forskare samlar in och rapporterar data om andel tumörrespons. Detta bidrar till de vida intervall för antitumoral effekt som rapporteras för i princip identiska behandlingar [19]. Trots att definitionerna av PR och CR är hyggligt standardiserade påverkar en rad andra faktorer de responsandelar som rapporteras [19]. Det sätt som används för att med ”objektiva” radiologiska metoder bedöma tumörvolymen kan uppenbarligen också påverka observerad remissionsandel [12].

När man till detta adderar att tumörrespons är tveksamt som surrogatvärde för patientnytta blir det uppenbart att data från fas II-studier vanligen inte kan visa vilket värde en behandling kan få när den tillämpas i klinisk rutin [9]. Ett undantag från denna regel är när en ny behandling är mycket mer effektiv än en tidigare använd. I ett sådant fall måste det också vara allmän kunskap att sjukdomen inte kan förbättras med existerande terapi [18].

Tid till respons, responsduration, progressionsfri och sjukdomsfri överlevnad beror på hur intensivt patienter följs upp, varför jämförelser mellan studier är svåra. Tid till progress fastställs vanligen med hjälp av objektiva, oftast radiologiska mätningar av tumörvolymen. En förlängning av tiden till progress anses ofta motsvara klinisk nytta. Även om detta troligen är riktigt i flertalet fall, är det nödvändigtvis inte sant för alla tumörtyper och tumörlokaler varför sambandet behöver motiveras för de behandlingssituationer som undersöks. I kliniska studier av tumörtyper där effektiv ”nästalinjes” cytostatikabehandling finns tillgänglig, har tid till progress ofta tillmätts större betydelse eftersom den totala överlevnaden kan påverkas av den behandling som ges efter studien varvid överlevnad blir ett osäkert studiemått.

Medan sjukdomsfri överlevnad speglar den botande potentialen hos en behandling inkluderar total överlevnad dödsfall av alla orsaker, alltså även möjliga, svåra biverkningar av behandlingen. Total överlevnad kan påverkas av behandling som har getts efter att sjukdomen återkommit, varvid slutsatsen om en överlevnadsskillnad för en viss behandling blir osäker. Generellt kan progressionsfri, sjukdomsfri och total överlevnad anses vara utvärderingsmått med god förmåga att spegla patientnytta. Vid t ex indolenta (långsamt växande) lymfom uppfattas, på basen av långvarig erfarenhet, även tumörrespons vara starkt kopplad till patientnytta, t ex i form av symtomlindring.

Symtomlindring är ett konceptuellt enkelt och rimligen viktigt utvärderingsmått i studier av palliativ cancerbehandling. Emellertid används detta utvärderingsmått förvånansvärt sällan i kliniska studier. För att den ska vara ett informativt mått bör man motivera vilken typ av symtom och vilka grader av förbättring som bedöms som signifikanta. Det ställer i sin tur krav på bedömningsskalan och dess validitet, dvs förmågan att spegla kliniskt viktiga förändringar. Symtomskalor ingår ofta i frågeformulären för mätning av livskvalitet. Ibland bedöms förändringen av symtomen objektivt, t ex av läkaren eller av en forskningssköterska. Liksom när patienten själv rapporterar om förändrade symtom måste man beakta risken för att kännedom om behandlingen påverkar bedömningen. Extern och blindad utvärdering skulle kunna vara ett sätt att undvika detta problem men används sällan.

Ett antal validerade frågeformulär om livskvalitet finns tillgängliga och denna typ av mätningar används i ökande omfattning i cancerstudier. Detta förefaller principiellt riktigt eftersom cytostatikabehandling för stora grupper av patienter knappast leder till påtagligt längre överlevnad. Då blir förbättrad livskvalitet till följd av behandlingen viktig. Även om det intuitivt är lätt att förstå innebörden av livskvalitet lider mätningar av denna variabel fortfarande av konceptuella och metodologiska problem. Sålunda är det svårt att definiera vad som korrekt speglar livskvaliteten. Vidare är mätningarna även i väl kontrollerade och väl utförda kliniska studier förenade med flera metodologiska problem, t ex det stora antalet frågor, öppen studiedesign, avsaknad av i förväg uppställda hypoteser om förväntade skillnader, svårigheter i bedömningen av klinisk relevans samt minskande mängd data efterhand som studien pågår. Dessa problem diskuteras mer detaljerat i kapitlet om livskvalitet. Även om livskvaliteten alltså intuitivt är ett starkt utvärderingsmått som förefaller direkt kopplat till klinisk effekt, tillåter informationen från mätningarna dessvärre inte otvetydiga slutsatser om den kliniska nyttan. Detta område är uppenbarligen i behov av ytterligare metodologisk och konceptuell utveckling för att kunna ge kliniskt användbar information.

Ett surrogatmått är lättare att mäta än det egentligen mest kliniskt relevanta utvärderingsmättet. Det är en prognosfaktor och skillnader mellan behandlingsgrupper i surrogatmättet ska översättas till skillnader i det egentliga utvärderingsmättet. I idealfallet bör det dessutom vara möjligt att uppskatta effekten på det egentliga utvärderingsmättet med hjälp av effekten på surrogatmättet [5,14]. Surrogatmått används ofta i kliniska studier, som t ex blodtryck vid kardiovaskulär sjukdom, ögontryck vid glaukom och CD4-antal i HIV-studier. Inom cancerområdet är andelen tumörresponser och förändringar av tumörmarkörer som PSA och CA-125 exempel på surrogatmått. Dessvärre är väldigt få av de surrogatmått som används inom medicinsk forskning tillfredsställande validerade.

Sammanfattningsvis kan man säga att rapporteringen av klinisk nytta av cytostatikabehandling inte bör bero på ett enda utvärderingsmått. Snarare bör värderingen av klinisk nytta grundas på en totalbedömning baserad på flera utvärderingsmått. För behandlingar med goda möjligheter att bota är sjukdomsfri och total överlevnad viktigast. Vid palliativ

behandling måste förlängning av överlevnad eller tid till progress, symtomlindring och livskvalitetsaspekter vägas samman. I den slutliga bedömningen av vilken nytta behandlingen gör måste detta ställas mot vilka biverkningar den ger och vilka uppföringar den kräver av patienten. Enstaka patienter kan dock ha starka preferenser för en viss behandling. Det kan då få större betydelse för valet av terapi än rekommendationer baserade på den vetenskapliga litteraturen.

Studiedesign och tolkning av studier av cytostatikabehandling

I stort sett alla bekräftande studier inom onkologin, dvs de som ger data med sådant bevisvärde att de kan användas för rekommendationer om behandlingar, behöver vara jämförande. Det betyder att två eller flera behandlingsalternativ jämförs i parallella grupper prospektivt och randomiserat. Studier av cytostatikabehandling är dock, till skillnad från andra studier om effekten av andra läkemedel, sällan blinda, dvs både vårdgivare och patient har vanligen kännedom om vilket behandlingsalternativ som ges.

Även om denna öppna studiedesign knappast påverkar objektiva slutvariabler som progressions- och sjukdomsfri överlevnad eller total överlevnad, kan inte oavsiktlig påverkan helt uteslutas. Sålunda kan den öppna studiedesignen t ex ge en omedveten tendens att tidigare dra bort patienter från en gammal behandling än från en ny vilket kan påverka resultatet. För subjektiva slutvariabler som symtomförbättring och livskvalitet kan en öppen studiedesign påverka utfallet oavsett om utvärderingsmättet skattas av vårdgivaren eller patienten.

Icke jämförande fas II-studier kan ytterst sällan ge data som tillåter en bedömning av verklig klinisk nytta av ett nytt läkemedel eller en ny behandling, inte ens om de utförts noggrant och har rekryterat många patienter. Orsaken till detta är att de effektdata som erhålls från sådana studier beror på vilken typ av patienter som inkluderats. Denna patientselektionsfaktor är svår att kvantifiera. Den är ofta inte alldeles uppenbar ens efter en detaljerad genomgång av de kriterier som används för att ta in och utesluta patienter och beskrivningen av studiepopulationens karaktäristika. Patientselektion är troligen en viktig faktor bakom de ofta dramatiska skillnaderna i t ex andelen tumörrespons som observeras i olika fas II-studier med i stort sett identisk behandling och likartade

kriterier för rekrytering av patienter. Sålunda bör även mycket lovande observationer från fas II-studier vanligen bekräftas i jämförande fas III-studier [17].

Nya cytostatika kan vanligen väntas ge bara en lite förbättrad effekt jämfört med den behandling som är standard. Därför kan i princip endast prospektiva, jämförande, randomiserade studier ge data som tillåter rimligt säkra slutsatser om effekt och säkerhet. Vidare måste dessa studier vanligen vara stora, vilket gör det nödvändigt att institutioner i flera länder deltar.

Beslut om rutinbehandling kommer alltså att baseras på data om patientgrupper som inte nödvändigtvis är fullt representativa för svenska förhållanden. En annan tänkbar nackdel med de stora studierna är att det under lång tid kan vara svårt att få tillgång till patientgrupper som är viktiga för mindre fas II-studier, vilka teoretiskt skulle kunna peka ut vägen till genombrott i ett terapiområde. Samtidigt ger tillvägagångssättet med multicenter- och multinationella studier troligen upphov till slumpmässiga fel som tenderar att minska skillnaderna mellan behandlingar. Om en signifikant skillnad i effekt ändå noteras i en stor studie så är detta troligen en tämligen robust observation.

Patientselektionsfaktorn är också av stor betydelse i retrospektiva analyser av patientmaterial insamlade från en eller flera institutioner. I retrospektiva studier är ofta data av varierande kvalitet och resultaten av sådana studier kan vanligen endast betraktas som hypotesgenererande. Detta gäller t ex för de stora internationella databaserna över patienter med olika tumörtyper som behandlats med högdosbehandling med stamcellsstöd. Fallkontrollmatchning inom kohorter kan emellertid öka möjligheterna till en korrekt jämförelse varvid resultaten kan utgöra stöd till de från randomiserade studier.

Strikt populationsbaserade patientmaterial minskar problemet med patientselektion avsevärt. Även om det i dessa fall fortfarande är svårt att veta om ett förändrat utfall är resultatet av en speciell behandling, andra interventioner eller andra inte orsaksrelaterade faktorer, kan sådana studier ge viktig kunskap. Om sådana studier prospektivt följer ett protokoll, noggrant registrerar samtliga fall och använder god statistisk analys kan värdet vara nästan i nivå med det av randomiserade studier. Möjligheterna till sådana studier är särskilt goda i de nordiska länderna. För

ovanliga tumörtyper eller ovanliga former av vanliga tumörer kan väl genomförda observationsstudier vara det enda sättet att nå ökad kunskap.

Dessvärre baseras slutsatser om effekter av cytostatikabehandling fortfarande ofta på andelen tumörresponser i icke-kontrollerade fas II-studier. Även om högkvalitativa fas-III studier utförts finns fortfarande frågan om resultatens generaliserbarhet, dvs i vilken omfattning en observerad effekt faktiskt kan uppnås då behandlingen används rutinmässigt. Troligen blir förväntad nytta mindre om den nya behandlingen används hos patienter med sämre prognos, t ex pga hög ålder, mer avancerat tumörstadium, sämre allmäntillstånd eller mer omfattande tidigare behandling. Detta blir mer kritiskt ju mindre en behandlingseffekt är; om cytostatikabehandling ger en median överlevnadsförlängning på en till tre månader, som vid icke-småcellig lungcancer, kanske detta kan observeras endast om behandlingen ges till patienter med gynnsamma prognosfaktorer. Idealet är att expertpaneler beaktar detta när de formulerar rekommendationer för cancerbehandling. Sådana rekommendationer bör alltså inte bara ange vilka patienter som ska få behandlingen utan också vilka som inte bör få den.

Referenser

1. American Society of Clinical Oncology. New international unidimensional response criteria for solid tumors: simpler and standardised. 35th annual meeting; special session 1999.
2. American Society of Clinical Oncology. Outcome for technology assessments and guideline development. http://www.asco.org/prof/pp/html/m_outcome3.htm 1999.
3. Medicinska forskningsrådet, Sjukvårdens prioriterings- och rationaliseringsinstitut. Cytostatikabehandling vid cancer (Glimelius B; Hafström L-O, eds); i Konsensuskonferens Stockholm 1990, 146.
4. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. *Statistics in Medicine* 1996; 15:2797-2812.
5. Buyse M, Molenberghs G. Criteria for the validation of surrogate endpoints in randomised experiments. *Biometrics* 1998;54:1014-29.
6. Buyse M, Thirion P, Carlson RW, et al. Relation between tumour response to first-line chemotherapy and survival in advanced colorectal cancer: a meta-analysis. *Lancet* 2000;356:373-378.
7. Constenla DO, Hill ME, A'Hern RP, et al. Chemotherapy for symptom control in recurrent squamous cell carcinoma of the head and neck. *Ann Oncol* 1997;8:445-449.
8. Deeks JJ. Systematic reviews of published evidence: Miracles or minefields. *Ann Oncol* 1998;9:703-709.
9. Eisenhauer EA. Phase I and II trials of novel anti-cancer agents: Endpoints, efficacy and existentialism. *Ann Oncol* 1998;9: 1047-52.
10. Glimelius B, Brorsson B. Cytostatikabehandling. Ändrad praxis tack vare ny forskning – ej som följd av konsensusuttalande. *Läkartidningen* 1993:1937-41.
11. Goodman C. Literature searching and evidence. Interpretation for assessing health care practices. SBU (The Swedish Council on Technology Assessment in Health Care). Stockholm 1993.
12. Gwyther S, Bolis G, Gore M, et al. Experience with independent radiological review during a topotecan trial in ovarian cancer. *Ann Oncol* 1997;8:463-468.
13. Haward RA. Preparing guidelines and documented clinical policies. *Ann Oncol* 1998;9:1073-78.
14. Jonsson B. Problems relating to difficult medical areas, eg cancer and AIDS drugs; Preclinical and Clinical Evaluation of Medicinal Products, Nordiska Läke-medelsnämnden 1996;41:47-55.
15. LeLorier J, Grégoire G, Behhaddad A, et al. Discrepancies between meta-analyses and subsequent large randomized, controlled trials. *N Engl J Med* 1997;337:536-542.

16. Miller AB, Hoogstraten B, Staquet M, et al. Reporting results of cancer treatment. *Cancer* 1981;47:207-214.
17. O'Connell MJ. Etoposide, doxorubicin, and cisplatin chemotherapy for advanced gastric cancer: an old lesson revisited. *J Clin Oncol* 1992;10:515-516.
18. Staquet M, Dalesio O. Designs for phase III trials; in Buyse ME, Staquet M, Sylvester RJ (eds), *Cancer clinical trials*, Oxford, Oxford University Press 1992: 261-275.
19. Tonkin K, Tritchler D, Tannock I. Criteria of tumour response used in clinical trials of chemotherapy. *J Clin Oncol* 1985; 3:870-875.
20. Villar J, Carroli G, Belizán JM. Predictive ability of meta-analyses of randomised controlled trials. *Lancet* 1995;345: 772-776.