

Värdering av effektivitet i klinisk vardag

Statistiska strategier för att hantera
skillnader i behandlingseffektivitet mellan
randomiserade kliniska prövningar och praktik

Ett regeringsuppdrag från Socialdepartementet

Aktivitet 1.2 i den Nationella läkemedelsstrategin 2016–2018

Rapportserie Denna rapport hör till serien SBU Bereder. Beredning av frågor för andra myndigheters eller uppdragsgivares verksamhet eller beslut, exempelvis till nationella riktlinjer. Vetenskapligt kunskapsunderlag tas fram av medarbetare på SBU i samarbete med ämnessakkunniga. Arbetsprocessen för att ta fram underlaget varierar beroende på frågeställning och behov.

ISSN 1400-1403

Innehållsdeklaration	Utvärdering av ny/etablerad metod Systematisk litteratursökning Relevansgranskning Kvalitetsgranskning Sammanvägning av resultat Evidensgradering gjord av SBU Evidensgradering gjord externt Baseras på en systematisk litteraturoversikt Konsensusprocess	✓ Framtagen i samarbete med sakkunniga Patienter/brukare medverkat ✓ Etiska aspekter ✓ Ekonomiska aspekter Sociala aspekter ✓ Godkänd av SBU:s prioriterings- och kvalitetsgrupp ✓ Godkänd av SBU:s nämnd
-----------------------------	--	---

Nyckelord Randomiserade kliniska prövningar, Register

Utgiven November 2016

Giltighetstid Resultat som bygger på ett starkt vetenskapligt underlag fortsätter vanligen att gälla under en lång tid framåt. Andra resultat kan ha hunnit bli inaktuella. Det gäller främst områden där det vetenskapliga underlaget är otillräckligt eller begränsat

Beställ Denna rapport (nr 256) kan beställas från Strömberg distribution.
Telefon: 08-779 96 85 • Fax: 08-779 96 10 • E-post: sbu@strd.se

Produktion Grafisk produktion av Anna Edling, SBU. Tryckt av Elanders Sverige AB, Mölnlycke, 2016. Omslagsfoto: Shutterstock

Diarienummer UTV2015/308

Citera denna rapport SBU. Värdering av effektivitet i klinisk vardag – Statistiska strategier för att hantera skillnader i behandlingseffektivitet mellan randomiserade kliniska prövningar och praktik. Ett regeringsuppdrag från Socialdepartementet Aktivitet 1.2 i den Nationella läkemedelsstrategin 2016–2018. Stockholm: Statens beredning för medicinsk och social utvärdering (SBU); 2016. SBU-rapport nr 256. ISBN 978-91-85413-99-7.

Innehåll

Sammanfattning	5
1 Bakgrund och syfte	9
2 Varför behövs uppföljning av behandlingseffekt?	13
Uppföljning i klinisk vardag	13
— Ordnat införande av nya läkemedel	14
— Myndigheternas behov av uppföljning av nya läkemedels effekt	15
— TLV:s utveckling av den värdebaserade prissättningen	15
— Läkemedelsverkets behov av uppföljningsdata	16
3 Begrepp och termer	19
Observationer och experiment	19
— RCT	19
— Observationsstudier	21
Kausal inferens	22
— The potential outcome model	22
Validitet	25
— Efficacy och effectiveness	26
4 Metod för översikten	29
Litteratursökning	29
Kvalitetsgranskning, sammanläggning och evidensgradering	30
5 Resultat	31
Grundantaganden	34
Skattning av PATE	35
— Ytterligare antaganden vid skattning av PATE	35
— Skattningsmetoder	36
Skattning av PATT	37
— Ytterligare antaganden vid skattning av PATT	37
— Skattningsmetoder	38
— Placebotest	39
Skattning av PATC	40
Övriga strategier	40
— Hierarkiska Bayesianska regressionsmodellerna	40
— Uppskattning av generaliserbarhetsbias	40
— Grafiska modeller	41
6 Datakällor	43
Vilka datakällor från klinisk vardag finns att tillgå?	43
— Läkemedelsregistret	43
— Övriga hälsodataregister	44
— Nationella kvalitetsregister	45
— Vård databaser	48
— Journalsystem	48
— Prövningsdatabaser	49

7	Ekonomiska aspekter	51
	Modellanalyser av kostnadseffektivitet	52
	Nya effektdata från svensk klinisk praxis och modellanalyser	54
8	Etiska aspekter	57
	Behovet av prioriteringar	57
	Problem med att mäta nytta	58
	Hantering av olika kunskapsunderlag	59
	Rättviseproblem i ett övergångsskede	60
	Konsekvenser av ett ändrat kunskapsläge	60
	Att avsluta pågående behandling	60
	Rättviseproblem mellan olika branscher	61
9	Diskussion	63
	Varför behövs denna rapport?	63
	Olika angreppssätt	63
	Strategiernas antaganden, problem och begränsningar	65
	Alternativa datakällor	65
	Alternativa användningsområden och frågeställningar	67
	Resursåtgång	67
	Fortsatt arbete	68
10	Projektgrupp, externa granskare, råd och nämnd	69
	Projektgrupp	69
	— Sakkunniga	69
	— SBU	69
	Externa granskare	70
	Bindningar och jäv	70
	SBU:s vetenskapliga råd – Brage	70
	SBU:s nämnd	71
11	Extern samverkan och referensgrupp	73
12	Ordförklaringar och förkortningar	75
13	Referenser	79

Bilaga 1 Sökstrategier

tillgänglig på www.sbu.se/256

Sammanfattning

Bakgrund

Uppföljning av behandling i klinisk vardag handlar om att ta tillvara den information om användning, behandlingseffekt, säkerhet och kostnadseffektivitet som finns i hälsodata- och kvalitetsregister och i hälso- och sjukvårdens övriga dokumentation. Sådan uppföljning är särskilt viktig i samband med att nya behandlingsmetoder introducerats i vården. De data som finns tillgängliga när en behandlingsmetod introduceras baseras i bästa fall på randomiserade kontrollerade studier (RCT), som i flera avseenden skiljer sig från behandling i klinisk vardag. Patienterna i klinisk vardag har ofta mer samsjuklighet, annan ålder och sämre prognos i allmänhet än patienterna i RCT. Projektets frågeställning kan betraktas som ett överförbarhetsproblem där vi har en uppmätt behandlingseffekt från ett urval ur en population och där vi är intresserade av att testa hur väl denna effekt speglar behandlingseffekten i den bakomliggande populationen.

Syfte

Projektets syfte är att kartlägga vilka statistiska strategier som finns publicerade idag för att analysera överförbarheten av RCT-data till behandlingseffektivitet i klinisk vardag. Målet med dessa strategier är att förbättra bedömningen av behandlingseffektivitet i klinisk vardag.

Litteraturgenomgång

Inledningsvis genomfördes litteratursökningar med breda ansatser för att försöka ringa in ämnesområdet. Under processen med att identifiera potentiellt relevanta studier har kompletterande metoder såsom kontroll av referenslistor och forskningskontakter varit av stor betydelse.

Resultat

Litteratursökningen resulterade i nio relevanta artiklar. De strategier som identifierats i artiklarna handlar om att justera resultatet från RCT utifrån kända egenskaper hos patientgrupperna. Strategierna är främst inriktade på följande situationer:

1. Innan läkemedlet har börjat användas:

Läkemedlet är godkänt och genomgår hälsoekonomisk värdering inför förmånsbeslut. Baserat på registerdata om den tänkta svenska patientpopulationens egenskaper kan man med dessa strategier bättre uppskatta den förväntade effekten utifrån resultat i RCT:er.

2. Efter det att läkemedlet har använts en tid:

Det är då möjligt att undersöka om utfallet i klinisk vardag avviker från det förväntade utfallet som uppskattats utifrån resultat i RCT:er.

Strategierna bygger på etablerade statistiska metoder, däribland olika regressionsmodeller, matchningsmetoder och viktning utifrån förekomst av egenskaper hos patienterna.

Det finns dock en rad antaganden, problem och begränsningar som ökar osäkerheten i uppskattning av den förväntade effekten av en behandling i klinisk vardag:

- Ett antagande är att man känner till alla faktorer av betydelse för selectionen till RCT och utfallen. Om detta antagande inte stämmer finns en risk för antingen över- eller underskattning av effektiviteten i klinisk vardag.
- Ett problem är tillgången till den nödvändiga informationen om de betydelsefulla faktorerna, både för kliniska prövningspatienterna och för patienterna i klinisk vardag.
- De strategier som tas upp i rapporten kan inte, utan osäkra antaganden och extrapolering av data, användas för att generalisera resultat utanför den typ av patienter som undersöks i den RCT som använts. Detta innebär i princip att generaliserbarheten begränsas av de inklusions- och exklusionskriterier som används i RCT:n.

Denna rapport visar att det finns strategier eller tillvägagångsätt som skulle kunna användas för att möjliggöra en bättre uppskattning av vilken nytta en behandlingsmetod ger i klinisk vardag. Det återstår dock att testa om denna typ av analys är genomförbar i praktiken och vilka praktiska eller legala problem som behöver lösas.

Etiska aspekter

Etiska problem har identifierats inom två områden:

- Omvärdering av ett läkemedels effektivitet i klinisk vardag och därmed dess kostnadseffektivitet kan föranleda att pågående behandlingar kan behöva avslutas. Detta kan innebära en förlust för patienter och anhöriga samt en påfrestande situation för behandlaren.
- Ytterligare krav på dokumenterad effekt av läkemedelsbehandling kan ses som en orättvisa gentemot berörda företag om inte motsvarande krav ställs på andra insatser i vården.

1 Bakgrund och syfte

Kunskap om behandlingseffekter är en förutsättning för en ändamålsenlig användning av insatser i hälso- och sjukvården. När det gäller läkemedel är denna kunskap vid godkännandetillfället begränsad till vad som framkommit i framför allt kliniska läkemedelsprövningar. Dessa är omgärdade av betydande osäkerheter. I takt med att ett läkemedel används av patienter i klinisk vardag finns viktiga möjligheter till fortsatt generering av kunskap genom dokumentation och analys av behandlingsutfall.

Förutsättningarna för att studera läkemedelseffekter inom ramen för en klinisk prövning respektive i klinisk vardag¹ skiljer sig emellertid markant.

I en tillräckligt stor, väl designad randomiserad kontrollerad prövning (RCT) finns en jämförelsegrupp (kontrollgrupp) som tack vare en slumpmässig tilldelning av behandling är likvärdig den behandlade patientgruppen med avseende på observerbara och icke-observerbara faktorer. Behandlingseffekten av ett läkemedel motsvaras då av skillnaden i behandlingsutfall mellan den grupp som fått läkemedlet och jämförelsegruppen.

I klinisk vardag kan det visserligen vara möjligt att identifiera jämförelsegrupper, men i och med att behandlingsalternativen inte tilldelas slumpmässigt, kan man inte utgå ifrån att patienter i en behandlad grupp respektive en jämförelsegrupp är likvärdiga. Detta för med sig ett antal metodologiska problem när det gäller att värdera behandlingseffekter på populationsnivå. Det kan också vara så

¹ Läkemedel i klinisk vardag är läkemedel (som ges på sjukhus eller på recept) som inte ingår i kliniska kontrollerade studier

att det nya läkemedlet ännu inte används i någon större omfattning, alternativt att i princip alla patienter som bör behandlas får behandling, vilket innebär att det inte finns några patienter i realtid att jämföra med. Detta kan exempelvis vara fallet när ett nytt läkemedel mot en livshotande sjukdom har visats bota eller förlänga överlevnaden i väsentlig grad. Samtidigt finns en samsyn kring vikten av att följa upp behandlingseffekter i klinisk vardag, eftersom det är effekten som bör ligga till grund för det pris samhället är redo att betala och de rekommendationer kring hur läkemedlet bör användas. Metoder för uppföljning av behandlingseffekter i klinisk vardag är relativt väl etablerade, men metoder för att generalisera RCT-resultat, och jämföra behandlingseffekt från RCT med observationella data, är däremot mindre väletablerade.

Nationella läkemedelsstrategin (NLS) är socialdepartementets plattform för att samverka med myndigheter, intresseorganisationer och industri inom läkemedelsområdet. I strategin, som uppdateras regelbundet, finns ett antal aktiviteter. Behovet av att ta fram metoder för att förbättra möjligheten att följa upp läkemedels effekter i klinisk vardag identifierades tidigt. Detta utmynnade i ett regeringsuppdrag till SBU att granska den vetenskapliga litteraturen och ta fram en metod som gör det möjligt att använda data från kliniska läkemedelsstudier som referens vid uppföljning av behandlingseffekt i klinisk vardag. Uppdraget utgör en första del av aktivitet 1.2 ”Värdering av behandlingseffekt i klinisk vardag” inom målområde ”Effektiv och säker läkemedelsanvändning”.

Syftet med utredningen är att med hjälp av en litteratursökning identifiera en eller flera metoder som kan användas för att översätta effektdata från kliniska prövningar till motsvarande effekt när behandling används i klinisk vardag. De metoder som projektet tagit fram är främst inriktade på situationer som skiljer sig något från den ursprungliga frågeställningen. Projektgruppen har formulerat en frågeställning som de eftersökta metoderna ska bidra till att besvara:

Givet en effekt av en behandling från en eller flera RCT, vilken effekt kan man förvänta sig när vi använder behandlingen i klinisk praxis hos en patientgrupp med en annan förekomst av betydelsefulla egenskaper?

Frågeställningen handlar om att betrakta populationen i den randomiserade studien som ett urval från den population som finns i klinisk praxis, men där patienter har olika sannolikhet att finnas med i RCT:n som således inte är ett vanligt enkelt slumpmässigt urval från populationen. Frågeställningen kan ses som ett generaliseringsproblem där vi har en skattad behandlingseffekt från ett urval ur en population och där vi är intresserade av att generalisera resultatet till den bakomliggande populationen.

De tillgängliga strategierna handlar i princip om att korrigera resultatet från RCT utifrån uppskattade patientspecifika urvalssannolikheter med vissa variationer utifrån vilka data som finns tillgängliga. Man kan tänka sig att, redan innan en kohort är behandlad i klinisk praxis, utifrån RCT-data uppskatta den förväntade effekten av behandlingen i klinisk praxis, vilket kan vara av värde vid en tidig värdering av behandlingens nytta. Alternativt kan man använda data från en behandlad kohort med data från RCT för att se om utfallet i klinisk praxis motsvarar förväntningarna.

Mottagare av denna rapport är, förutom uppdragsgivaren socialdepartementet, myndigheter, hälso- och sjukvårdens huvudmän och företagen. Av myndigheter är det Tandvårds- och läkemedelsförmånsverket (TLV), Läkemedelsverket (LV) och Socialstyrelsen som kan ha störst användning av rapporten. Huvudmännen representeras av landstingen och SKL (NT-rådet och andra delar av samverkansmodellen). Läkemedelsföretag och företag inom andra delar av life science sektorn, är mottagare och potentiella användare av rapportens resultat. Rapporten torde också ha intresse för hälsoekonomer, läkemedelsansvariga och professionella med sjukvårdsansvar. Resultaten av rapporten skulle även kunna användas i SBU:s arbete med att förbättra sina utvärderingsmetoder.

2 Varför behövs uppföljning av behandlingseffekt?

Uppföljning i klinisk vardag

I bred bemärkelse innebär uppföljning av läkemedel i klinisk vardag att studera användning och behandlingsutfall efter marknads godkännandet, när läkemedlet börjat användas av patienterna. På individnivå är detta något som varje behandlande läkare ägnar sig åt dagligen, men detta kan också följas upp systematiskt på gruppnivå. Inom detta projekt ligger fokus på uppföljning på gruppnivå.

Det finns tre huvudsakliga kategorier av läkemedelsuppföljning:

1. Uppföljning av hur läkemedel används

Här ligger fokus på att undersöka vilka grupper som använder ett läkemedel, hur länge de står på läkemedlet, hur det kombineras med andra behandlingar och var i en behandlingssekvens det inryms. För att genomföra den här typen av uppföljningsanalyser på gruppnivå räcker det oftast att ha tillgång till en datakälla med läkemedelsuppgifter.

2. Uppföljning av säkerhet

Vid godkännandet av nya läkemedel är oftast endast de vanligaste biverkningarna kända medan kunskapen om mer ovanliga biverkningar, eller biverkningar som uppträder efter lång tids behandling, är ofullständig. Epidemiologiska studier efter godkännandet är därför en viktig komponent i den fortlöpande nytta–risk–värderingen. De metodologiska utmaningarna är likartade som för uppföljning av behandlingseffekt med tillägget att de

utfall som studeras ofta är ovanliga och därför kan kräva mycket stora studiepopulationer. En mer strukturerad uppföljning kan ske inom ramen för så kallad Post Approval Safety Studies (PASS).

3. **Uppföljning av behandlingseffekt**

Behandlingsutfall hos patienter som fått exempelvis ett läkemedel i klinisk vardag studeras och analyseras. Även för den här typen av uppföljningsanalyser behövs tillgång till en datakälla med uppgifter om läkemedel, patentkaraktäristiska och behandlingsutfall, alternativt separata datakällor för läkemedel respektive behandlingsutfall som samkörts. Dessutom kan det behövas andra uppgifter som innehåller exempelvis demografisk och socioekonomisk information. Sådan uppföljning kan ske genom att använda etablerade kvalitetsregister eller via en Post Approval Efficacy Study (PAES).

Inom ramen för detta projekt är det i första hand den tredje kategorin som är av intresse, det vill säga uppföljning av behandlingseffekter.

Till skillnad från säkerhetsaspekter, finns inget system för uppföljning av behandlingseffekter av läkemedel. Sådan uppföljningar sker istället inom ramen för enskilda så kallade observationsstudier. Observationsstudier kan ha olika design, kohortstudier eller fall-kontrollstudier. Observationsstudier som utvärderar behandlingseffekter bygger ofta på retrospektivt uttag och statistisk analys av relevanta data från register eller vårdadministrativa databaser där läkemedelsanvändning och behandlingsutfall dokumenterats.

Kunskapen om nya behandlingsmetoder tilltar i takt med att de används. Utöver ren kunskapsgenerering finns en trend och ett antal konkreta projekt mot att ett ökat fokus riktas mot uppföljning av användningen. Nedanstående avsnitt belyser ett antal exempel som illustrerar den tilltagande efterfrågan på uppföljning av behandlingseffekter.

Ordnat införande av nya läkemedel

I januari 2015 sjösattes en landstingsgemensam samverkansmodell för nationellt ordnat införande av nya läkemedel. Syftet är att uppnå en jämlik, kostnadseffektiv och ändamålsenlig användning av nya läkemedel över hela landet. I förslaget som arbetat fram av SKL, inom ramen för den nationella läkemedelsstrategin, är uppföljning ett återkommande ledord. Möjligheter till strukturerad uppföljning av nya läkemedel har också studerats i ett separat delprojekt i den nationella läkemedelsstrategin (delprojekt 6.2). Den ursprungliga planen med det ordnade införandet var att det även ska omfatta andra interventioner än läkemedelsbehandling, även om organisationen i ett första skede i byggts för hantering av nya läkemedel.

I processen för ordnat införande introduceras läkemedel på en av tre ambitionsnivåer, enligt särskilda kriterier. För läkemedel där man har en hög ambitionsnivå tas ett införande- och uppföljningsprotokoll fram inför introduktionen. Det främsta syftet med uppföljningen är att klargöra om införandet följt NT-rådets rekommendationer och begränsningar, alternativt nämnden för

läkemedelsförmånens begränsningar i förmånsbeslutet. Man vill dessutom få en uppfattning om omfattningen av läkemedlets användning, om läkemedelsförbrukningen och om det använts till rätt patientgrupp och på rätt sätt. För vissa läkemedel med högt kostnadsläge, eller där osäkerheten i beslutsunderlaget bedöms vara stor, finns även en önskan att följa upp om behandlingseffekten kan antas vara i paritet med den som sågs i de kliniska prövningarna [1].

Myndigheternas behov av uppföljning av nya läkemedels effekt

Ett behov av att följa upp behandlingseffekter i klinisk vardag finns även hos TLV och LV. En allt högre andel av TLV:s förmånsbeslut är kopplade till någon typ av begränsning. Åren 2014–2015 innehöll cirka 50 procent av alla förmånsbeslut begränsningar och/eller villkor. Detta återspeglar en utveckling där nya läkemedel i allt större utsträckning är riktade mot små patientgrupper och är omgärdade av osäkerheter vad gäller behandlingseffekt och behandlingstid. Höga kostnader för många av dessa nya läkemedel gör att osäkerheter kan innebära att sjukvårdshuvudmännen prioriterar fel och drabbas av onödiga kostnader.

Eventuell uppföljning av subventionsbegränsningar har för TLV:s del i regel handlat om att följa upp att begränsningen efterlevs, det vill säga att behandling förskrivs till rätt patienter och inte till grupper där användning inte bedömts vara kostnadseffektiv. Men det finns en önskan om att också följa upp behandlingseffekter, speciellt i de fall då nya indikationer tillkommer. Idag utgör effektdata från RCT:er den centrala delen av beslutsunderlaget vid subventionsbeslut.

Den vision som har presenterats av TLV innebär en utvecklad värdebaserad prissättning med möjlighet till prisjusteringar i takt med att kunskaper om behandlingseffekterna i klinisk vardag blir kända. Detta skulle vara ett sätt att säkerställa en kostnadseffektiv läkemedelsanvändning över hela läkemedlets livscykel.

TLV:s utveckling av den värdebaserade prissättningen

TLV:s utveckling av den värdebaserade prissättningen medför nya krav på uppföljning. Som en följd av det utvecklade arbetssättet, där trepartsöverläggningar med landstingen och läkemedelsföretagen ingår, har ett antal så kallade riskdelningsavtal slutits. Dessa går i korthet ut på att läkemedelsföretagen och landstingen delar på risken kring osäkerheter för läkemedel som ansökt om subvention.

Vid ställningstagande av nya läkemedel råder ibland osäkerhet kring behandlingstid. När nya Hepatit C-läkemedel introducerade hade en sådan osäkerhet betydande konsekvenser för den uppskattade kostnadseffektiviteten. Hepatit C-läkemedlen valdes som pilotärendet för denna nya prissättningsmodell. Ett riskdelningsavtal slöts mellan landstingen och läkemedelsföretagen som gick ut på att företagen betalade tillbaka en andel av behandlingskosten för de

patienter vars behandlingstid översteg 12 veckor. Riskdelning avtalades också kring antalet behandlade patienter.

Möjligheten till uppföljning på individnivå är en grundförutsättning för denna typ av riskdelningsavtal. Hittills har all uppföljning av riskdelningsavtal baserats på data från Läkeemedelsregistret.

Utöver behandlingstid och patientantal har behandlingsresultat identifierats som en potentiell parameter som det kan råda osäkerhet kring vilket bör tas med som en parameter vid utformning av riskdelningsavtal. Av de sidoöverenskommelser som hittills träffats mellan landsting och läkeemedelsföretag har dock inget inkluderat faktiska behandlingsresultat som en uppföljningsparameter. Detta kan förklaras av en generell bristfällig tillgång på data över behandlingseffekter i klinisk vardag som av de metodmässiga utmaningarna.

Läkeemedelsverkets behov av uppföljningsdata

Efter att ett läkemedel godkänts för marknadsföring så följer Läkeemedelsverket varje läkemedels nytta–risk profil under hela dess livscykel. Arbetet bedrivs i nära samverkan med andra EU-länder även i enskilda vetenskapliga frågor. I många fall måste till exempel uppföljningsstudier av säkerhetsfrågor genomföras i flera länder för att ge ett tillräckligt säkert beslutsunderlag.

Det är särskilt viktigt att följa upp nya läkemedel som introduceras på marknaden, men epidemiologiska studier kan initieras av läkeemedelsmyndigheterna också för äldre läkemedel. Övervakningen efter godkännandet är betydelsefull eftersom en större och mer skiftande grupp av patienter kommer att få läkemedlet efter godkännandet och ovanliga händelser kan ofta inte upptäckas med tillräckligt precision i de studier som ligger till grund för godkännandet.

Effekt och säkerhet följs med ledning av en riskhanteringsplan som upprättas för varje nytt läkemedel. Där specificeras bland annat vilka studier som ska göras efter godkännandet för att ytterligare karaktärisera läkemedlets nytta–riskprofil och ändamålsenlighet. En viktig informationskälla för nya biverknings-signaler är också spontanrapporter av misstänkta biverkningar från hälso- och sjukvården och konsumenterna. Med statistiska metoder letar man efter mönster och strukturer i stora databaser där dessa spontanrapporter samlas. Ibland måste nya epidemiologiska studier initieras för att stödja utredningen av en sådan signal. Signaler om nya säkerhetsproblem kan också grunda sig på nya studier som publiceras i vetenskaplig litteratur och den metodologiska värderingen av dessa studier är då central.

Läkeemedelsföretagen ska med bestämda tidsintervall komma in med periodiska säkerhetsrapporter (PSUR) som granskas av läkeemedelsmyndigheterna. Förutom en värdering av nya data från den aktuella tidsperioden görs en uppdaterad vetenskaplig utvärdering av läkemedlets nytta–risk-balans. I dessa rapporter ingår dels resultat från nya kliniska studier, riktade uppföljningar och epidemiologiska studier som ålagts företaget, resultat av andra vetenskapliga studier och resultat av utredningar av signaler baserade på spontanrapportering och andra källor.

Uppföljningsstudier av effekt och säkerhet i klinisk praxis är alltså en viktig komponent i LäkeMedelsverkets kontinuerliga övervakning som fortgår under läkemedlets hela livscykel, det vill säga, så länge produkten finns på marknaden. Det finns således starka skäl för att fortsätta utveckla metodiken för denna typ av studier.

3 Begrepp och termer

Observationer och experiment

Forskningsstudier kan delas upp i experimentella och observationsstudier. Ett experimentellt studieupplägg innebär att forskaren intervenerar i verkligheten för att sedan studera konsekvenserna. Upplägget för observationsstudier innebär däremot att ingen aktiv styrning av åtgärder vidtas utan forskaren observerar verkligheten som den är. Vi kan säga att skillnaden mellan experimentella och observationsstudier är förekomsten av variabelmanipulation som görs på ett kontrollerat sätt i experimentella [2].

Interventionsstudier är ett exempel på en experimentell studie vars syfte är att kartlägga effekterna av en åtgärd. Under sådana studier har forskaren kontroll över hur gruppindelning och behandlingstilldelning sker. När grupptilldelningen sker på ett slumpmässigt sätt kan vi teoretiskt anta att deltagarnas kända och okända bakgrundsvariabler är jämnt fördelade mellan grupperna, vilket skapar jämförbara grupper. De studierna som jämför minst två behandlingsalternativ och i vilka deltagarna fördelas slumpmässigt mellan behandlingsalternativ kallas för RCT.

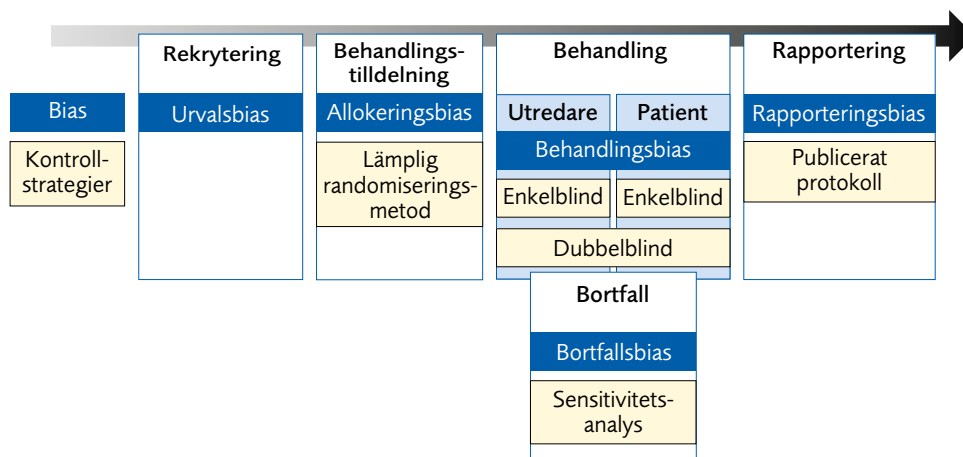
RCT

Målet med randomiserade kontrollerade kliniska studier är att kvantifiera en behandlingseffekt och/eller en säkerhet. Deltagarna indelas slumpmässigt till antingen kontroll- eller behandlingsgruppen och sannolikheten att hamna i den ena eller den andra gruppen är oberoende av deltagarnas bakgrundsvariabler. Därmed skapas jämförbara grupper där själva behandlingen är det enda som ska utgöra en skillnad mellan grupperna. Detta förutsätter dock att studien

är tillräckligt stor för att man ska kunna anta att viktigare bakgrundsvariabler inte råkat snedfördelas mellan grupperna. Målet med RCT:er är således att visa att de uppmätta skillnaderna i utfall mellan grupperna enbart beror på själva behandlingen och inte på några andra yttre faktorer.

Randomisering är ett sätt att försöka minimera de systematiska skillnaderna mellan jämförelsegrupperna i prognos eller mottaglighet för en behandling, det vill säga allokeringsbias. Det finns emellertid en rad andra systematiska fel (bias) som kan leda till fel slutsats. Dessa bias uppstår genom exempelvis procedurfel, effektbedömningsfel eller andra mänskliga fel och kan resultera i antingen under- eller överskattning av behandlingseffekten. I Figur 3.1 nedan visas de vanligt förekommande biasformerna under studiens gång. Fokus i denna rapport är urvalsbias, det vill säga systematiska skillnader som inträffar då en del patienter rekryteras till studien medan andra patienter utesluts. En alltför hård gallring av studiedeltagarna leder till att patienturvalet inte är representativt för patientpopulationen och resultaten blir inte lika allmängiltigt (dålig överförbarhet).

Figur 3.1
Potentiella biaskällor
i RCT och eventuella
kontrollåtgärder.



Trots närvaron av de olika formerna av bias är chanserna att minimera dem större i en RCT än för andra studieupplägg, vilket dock inte nödvändigtvis gäller urvalsbias och bortfallsbias. RCT:ers metodologiska fördelar gör dem till den studiedesign som föredras när regulatoriska krav utformas för bedömning av behandlingseffekter. De kan dock ha begränsningar när det gäller att skatta effekten av en behandling i en klinisk vardagspopulation som underlag för subventionsbeslut och framtagning av kliniska riktlinjer [3]. Det är dessutom inte alltid möjligt eller etiskt försvarbart att göra en RCT [4]. Exempel på detta kan vara en studie vars syfte är att bedöma långsiktiga risker av olika åtgärder eller studier som vill ta reda på hur alkoholkonsumtion påverkar fostrets utveckling. Ett alternativt tillvägagångssätt är att välja en annan studiemetodik, exempelvis observationsstudien.

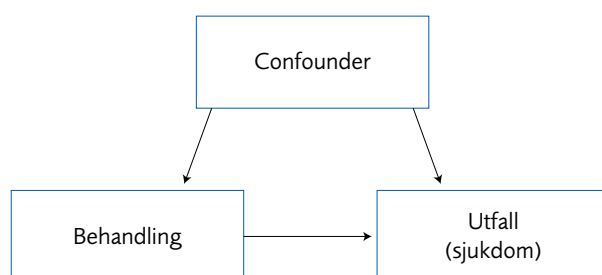
Observationsstudier

Utformning av icke-interventionsstudier kan se ut på olika sätt och omfattar bland annat studier som kartlägger behandlingspraxis eller kostnader, uppföljning av hur ett läkemedel används osv. Exempel på benämningar för dessa studier är observationsstudier, kartlägningsstudier, epidemiologiska studier, kvalitetssäkringsprojekt och uppföljningsstudier. Gemensamt för dessa studier är att patienter inkluderas om de fyller vissa urvalskriterier och därefter samlas data in, retrospektivt eller prospektivt genom exempelvis register eller journaler. Vidare gäller att behandlingen och uppföljning sker i enlighet med rådande klinisk praxis (LVFS 2011:19).

Observationsstudier kan planeras och utföras på olika sätt och metodologiskt brukar man dela in dem i exempelvis kohortstudier, retrospektiva studier (inklusive fall-kontrollstudier) och tvärsnittsstudier [5]. Variationen av studie-design gör att observationsstudier är olika känsliga för olika biasformer. Flera av de systematiska fel som förekommer i observationsstudier är desamma som för RCT:er, till exempel behandlingsbias, bedömningsbias, bortfallsbias och rapporteringsbias. Det är däremot risken för selektionsbias som utgör den största skillnaden mellan RCT och observationsstudier [6].

Selektionsbias¹ uppstår i observationsstudier när deltagarnas egenskaper är ojämnt fördelade mellan kontroll- och behandlingsgruppen. Att situationen uppstår beror på att valet av behandling görs av personalen på kliniken och inte slumpmässigt. Behandlingstilldelningen kan då komma att påverkas av läkarens personliga preferenser, patientens egenskaper, sjukdomsintensitet och patientens önskemål [7].

Selektionsbias skapar obalanser mellan jämförelsegrupper avseende prognostiska faktorer som är associerade med behandlingen och utfallet man vill mäta. Situationen kallas även för ”confounding” och innebär att den behandlade gruppen och kontrollgruppen visar skillnader i en eller flera egenskaper (confounders) som också är en riskfaktor för utfallet. En confounder är helt enkelt en faktor som stör skattningen av förhållandet mellan exponeringen (behandlingen) och utfallet (Figur 3.2).



Figur 3.2
Confounding uppstår när ett orsakssamband som man studerar påverkas av en eller flera länkade confounders (förväxlingsfaktorer, störfaktorer) som inte beaktats.

¹ Notera att termen ”selektionsbias” ofta används som synonym till ”urvalsbias”, men inte i denna rapport.

Kausal inferens

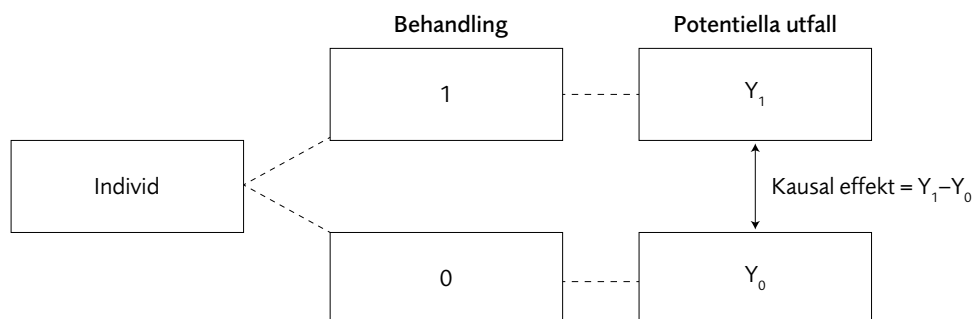
Trots skillnader i studieupplägg kan målet med RCT och observationsstudier vara detsamma till exempel när vi vill ta reda på effekten av en viss behandling (X) på ett eller flera utfall (Y). Närmare bestämt vill vi undersöka ett orsaks-samband (kausal samband) som innebär att en viss förändring i X har en viss effekt på Y allt annat likt. Det är viktigt att notera att kausalitet inte är detsamma som association. Kausalitet innebär att om man gör en intervention X har denna påverkan på Y. En association mellan två variabler innebär däremot inte nödvändigtvis att den ena variabeln orsakat den andra utan snarare en indikation att variablerna samvarierar med varandra.

Utmaningen med kausal inferens är att utifrån tillgänglig information kunna dra de bästa möjliga slutsatserna om det vi inte vet men vill ta reda på. De relevanta artiklarna som vi identifierat har alla sin grund i kausal inferens baserat på kontrafaktiska variabler. Dessa metoder klarlägger de antaganden som är nödvändiga för att man skall kunna genomföra en analys och kommer att diskuteras i Kapitel 4. Det är rimligt att se kausal inferens som ett ramverk inom vilket vi kan tillämpa flera specifika skattningsmetoder för behandlingseffekt. Det finns ett antal modeller som syftar till att hantera problem med kausal inferens och vi kommer att beskriva den dominerade modellen, nämligen "the potential outcome model" (även kallad the counterfactual model) [8–14].

The potential outcome model

Ett centralt begrepp inom modellen är potential outcome (möjligt utfall). Låt oss betrakta en studie där hälften av deltagarna tillhör behandlingsgruppen (behandling 1) och den andra hälften kontrollgruppen (behandling 0). Vi vill ta reda på effekten av behandling 1 jämfört med behandling 0 på ett utfall. Varje individ antas ha två potentiella utfall som betecknas med Y_1 och Y_0 . Det första, Y_1 , är det resultat vi kan mäta om individen får behandling 1 och Y_0 är utfallet vi kan mäta om individen får behandling 0. I modellen definieras den kausala effekten för en individ som skillnaden mellan dessa två potentiella utfall, (Figur 3.3).

Figur 3.3
Behandling 1 representerar den aktiva behandlingen medan behandling 0 är kontrollbehandling. Den kausala effekten av behandling 1 jämfört med kontrollbehandling 0 för individen är definierad som skillnaden mellan det potentiella utfallet Y_1 och det potentiella utfallet Y_0 .



I vanliga fall kan vi bara mäta det ena potentiella utfallet för en individ: Y_1 om individen fick den aktiva behandlingen eller Y_0 om individen fick kontrollbehandlingen (Tabell 3.1).

Behandling	Potentiella utfallen		Kausal effekt
	Y_1	Y_0	$Y_1 - Y_0$
1	Mätbar	?	$Y_1 - ?$
0	?	Mätbar	$? - Y_0$

Tabell 3.1
Potentiella utfallen och kausal effekt på individnivå.

Det gör att vi inte kan observera den kausala effekten på individnivå. Istället kan vi undersöka en grupp individer varav några har fått behandling 1 och några behandling 0. Syftet är att ta reda på väntevärdet av den genomsnittliga behandlingseffekten $E(\bar{Y}_1 - \bar{Y}_0)$.

Skattning av kausala effekter på gruppnivå kan försvåras om utfallen för en viss individ påverkas av behandlingar som andra individer i gruppen får. Därför måste ett stabilitetsantagande (stable unit-treatment value assumption, SUTVA) vara uppfyllt. Antagandet säger att det potentiella utfallet för en individ inte påverkas av behandlingar som tilldelas de andra individerna, det vill säga inga interferenser mellan individerna.

När vi nu tittar på en grupp individer istället för på den enskilda individen, märker vi att problemet med data som saknas kvarstår. Vi kan bara observera hälften av de potentiella utfallen (Tabell 3.2).

Individ	Behandling	Potentiella utfallen		Kausal effekt
		Y_1	Y_0	$Y_1 - Y_0$
1	1	Mätbar	?	$Y_1 - ?$
2	0	?	Mätbar	$? - Y_0$
3	1	Mätbar	?	$Y_1 - ?$
4	0	?	Mätbar	$? - Y_0$

Tabell 3.2
Potentiella utfallen och kausal effekt på gruppnivå.

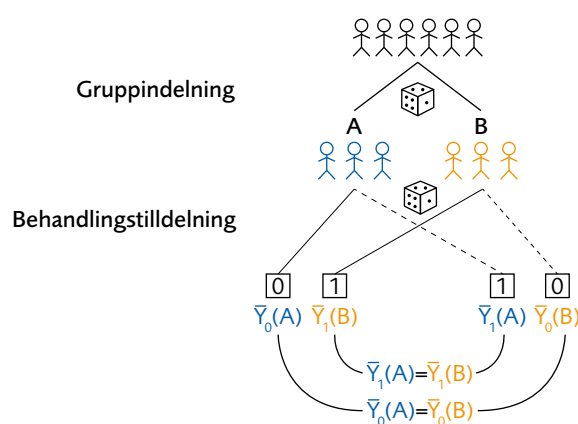
För att kunna skatta den kausala effekten behöver vi predicera de saknade potentiella utfallen. För att kunna fylla i de saknade värdena behöver dock flera antagande vara uppfyllda.

Innan vi går in på nästa antagande kommer vi att diskutera en viktig komponent i kausalanalys, nämligen, processen som avgör vilka individer som får vilka behandlingar (behandlingstilldelning). Det är den processen som avgör vilka potentiella utfall vi kommer att observera eller sakna.

Vid randomiserade studier indelas individerna slumpmässigt i två grupper, A och B, och därefter tilldelas behandlingar 0 och 1 (Figur 3.4) till var sin grupp. Randomisering och tillräcklig studiestorlek försäkrar att bakgrundvariabler som kan påverka utfallet är jämnt fördelade mellan grupperna. Utfallets genomsnittliga värde hos grupp A kommer att vara detsamma som det genomsnittliga utfallet hos grupp B om individerna i grupp A tilldelats samma behandling som gavs till individerna i grupp B. Vilken grupp som egentligen fick behandlingen A (eller B) är således irrelevant för utfallets genomsnittliga värde när grupptillhörigheten är randomiserad. Vi säger att randomisering leder till ”utbytbarhet” (exchangeability) och att grupperna A och B betraktas som utbytbara (exchangeable). Detta betyder att medelvärdet är detsamma i grupp A (om A fick behandling) och grupp B (om B fick behandling), det vill säga $\bar{Y}_1(A) = \bar{Y}_1(B)$. Samma resonemang gäller för utfall under kontrollbehandlingen, $\bar{Y}_0(A) = \bar{Y}_0(B)$.

Figur 3.4

Utbytbarhet. I RCT sker en slumpmässig gruppindelning som gör att alla individer har samma sannolikhet att hamna i grupp A eller B. Även behandlingen tilldelas slumpmässigt och vi får två potentiella utfall: Y_1 för behandlade individer och Y_0 för obehandlade individerna. Om studien är tillräcklig stor kommer bakgrundsvariablerna att jämfördelas mellan grupperna och vi säger att grupperna är utbytbara.



Den oberoende relationen mellan behandlingstilldelning och potentiella utfall under utbytbarhet innebär att vi inte får någon som helst information om de potentiella utfallen genom att känna till behandlingstilldelning. Därefter kan vi relativt enkelt skatta de kausala effekterna på gruppnivå genom att beräkna skillnaden mellan utfallets genomsnittliga värde för behandlingsgruppen $\bar{Y}_1(1)$, och kontrollgruppen $\bar{Y}_0(0)$.

I observationsstudier däremot kan vi inte längre anta att de potentiella utfallen och behandlingstilldelningen varit oberoende. Eftersom randomisering saknas kan det hända att vi får olika fördelningar av bakgrundsvariablerna i behandlingsgruppen och kontrollgruppen. Dessa bakgrundsvariabler kan påverka utfallet. Ett exempel kan vara ett läkemedel som tilldelas patienter med sämre prognos i större utsträckning än patienter med bättre prognos. Det leder till att bakgrundsvariabelprognos är snedfördelad mellan grupperna och kan påverka utfallet. En konsekvens blir att vi inte kan skatta kausala effekter genom att beräkna skillnaden i utfallets genomsnittliga värde mellan behandlings- och kontrollgruppen $\bar{Y}_1(1) - \bar{Y}_0(0)$.

För att kunna gå vidare med kausalanalys för icke-randomiserade studier behöver två antagande vara uppfyllda. Det första brukar kallas för "unconfoundedness" eller ignorerbarhet. Antagandet säger att, om vi känner till ett set av observerbara bakgrundsvariabler mätta före behandlingen, kan vi anta att behandlingstilldelning är oberoende av utfallen om setet innehåller alla bakgrundsvariabler som påverkar behandlingstilldelning och utfall, det vill säga alla confounders. För att bättre kunna förstå antagandet går vi tillbaka till exemplet ovan där benägenhet att bli behandlad beror på prognosvariabeln. Individerna i behandlingsgruppen har sämre prognos i genomsnitt än individerna i kontrollgruppen. Vi säger att grupperna inte är utbytbara eftersom fördelningen av prognosvariabeln varierar mellan behandlings- och kontrollgruppen. Istället skapar man två subgrupper varav den ena omfattar individer med dålig prognos och den andra utgörs av individer med bättre prognos. Vi kan nu tänka oss att våra data egentligen är en kombination av två separata randomiserade studier. En som är gjord på subgruppen "sämre prognos" och en som är gjord på subgruppen "bättre prognos". Nu kan vi anta att de behandlade och obehandlade individerna i varje subgrupp är villkorligt utbytbara. Med andra ord, om vi känner till utfallsprediktorerna som är snedfördelade mellan behandlings- och kontrollgruppen, kan vi skapa en villkorlig utbytbarhet.

Det andra antagandet kallas för överlapp eller positivet. Antagandet innebär att oavsett värdena på bakgrundsvariablerna har varje individ en positiv sannolikhet (det vill säga >0) att bli tilldelad behandlingen eller kontrollbehandlingen. Det enklaste sättet att förklara antagandet är genom att ge ett exempel när antagandet inte gäller. Antiretroviral behandling av HIV-infekterade patienter initieras när CD4-talet är under $200/\mu\text{L}$. Det innebär att det finns noll sannolikhet att patienten förblir obehandlad när cellnivån sjunker under gränsen. Situationen bryter mot antagandet eftersom för vissa värden av bakgrundsvariabler saknar vi behandlade och obehandlade individer som kan jämföras.

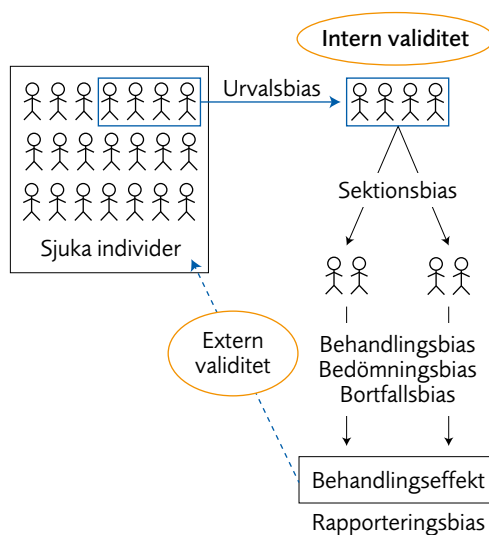
Validitet

Vilken typ av studie som bör väljas, RCT eller observationsstudie, bestäms av frågeställningen. Oavsett studieupplägg finns det två fundamentala frågor som bör ställas [15]:

1. Är studieresultaten korrekta för individerna som ingick i studien?
2. Är urvalet patienter i studien representativt för populationen av intresse?

Den första frågan handlar om intern validitet, det vill säga tillförlitligheten i själva undersökningen. Den avgörs av hur väl design, datainsamling och analyser genomförts och hotas bland annat av systematiska samt slumpmässiga fel. Intern validitet avser om själva behandlingen orsakat effekten eller om det är något annat som förklarar uppmätta resultat. Den andra frågan handlar om extern validitet eller generaliserbarhet. Den externa validiteten avser i vilken utsträckning resultaten har bredare giltighet, till exempel överförbarheten till andra individer än de som deltog i studien (Figur 3.5).

Figur 3.5
Validitet för en undersökning. Intern validitet avser tillförlitligheten hos en undersöknings resultat, medan extern validitet gäller i vilken grad undersökningens resultat har bredare giltighet.



Ett centralt antagande i RCT är att behandlingseffekten skattas från ett medelvärde i studiepopulationen eller definierade subpopulationer. Det finns inget krav på att studiepopulationen ska vara representativ för målpopulationen, däremot måste centrala egenskaper i målpopulationen vara representerade. Generellt gäller att RCT har bättre intern validitet än observationstudier eftersom studieupplägget minimerar uppkomsten systematiska fel. Den externa validiteten däremot är sämre för RCT (med undantag för så kallade pragmatiska RCT:er) och detta kan bero på en mängd faktorer, bland annat selektion av patienter, deras egenskaper samt skillnader mellan studieprotokoll och rådande klinisk praxis [16].

Godkännandet av ett läkemedel baseras på resultat från RCT. Deltagandet i dessa studier begränsas genom att införa specifika inklusions- och exklusionskriterier och därmed skapas en homogen grupp av patienter. En minskad heterogenitet i patientgruppen leder till lägre extern validitet men studiedesignen tillåter utredaren minimera confounding och därmed optimera möjligheten att skatta den kausala effekten av exponeringen. När behandlingen därefter tas i bruk, kommer den att förskrivas till patienter som kan skilja sig från de selekterade studiedeltagarna. Svar på behandlingen kan då vara annorlunda hos dessa patienter mot dem i klinisk vardag. Några av faktorerna som orsakar skillnader i behandlingsvar diskuterar i nästa avsnitt.

Efficacy och effectiveness

Efficacy och effectiveness är två relevanta engelska begrepp i detta sammanhang. Efficacy beskriver hur väl en medicinsk åtgärd fungerar när den undersöks i en klinisk prövning. Med hjälp av efficacy-studier vill utredaren veta om behandlingen fungerar under kontrollerade förutsättningar. Effectiveness däremot beskriver hur väl en medicinsk åtgärd fungerar när den är i allmänt bruk inom sjukvården. Problemet är att behandlingseffekten i klinisk vardag inte alltid är densamma som behandlingseffekten i RCT:n som utgjorde grunden för godkännandet [17]. Skillnaden i behandlingseffekt mellan patienter i RCT och patienter i klinisk praxis kallas för ”efficacy-effectiveness gap” [18].

Det finns flera orsaker till skillnader mellan efficacy och effectiveness [19]. En av orsakerna är variabilitet i behandlingsrespons som grundar sig inte enbart på biologiska faktorer utan även på patientens och förskrivarens beteende [17].

Biologiska faktorer kan påverka farmakodynamik (hur läkemedel påverkar kroppen), farmakokinetik (hur kroppen påverkar läkemedlet) och biverkningar. Genetiska skillnader mellan individer är ett exempel på en biologisk faktor som kan orsaka variabilitet i behandlingsrespons. Andra, icke-genetiska, biologiska faktorer som kan påverka behandlingseffekten kan delas i inre- och yttre faktorer. Med inre faktorer menas olika fysiologiska och patologiska tillstånd såsom ålder, kön, BMI, samsjuklighet eller sjukdomens svårighetsgrad. Yttre faktorer syftar till påverkan av miljön på individen, till exempel luftföroreningar eller andra pågående behandlingar inklusive läkemedel.

En annan orsak till skillnader mellan efficacy och effectiveness kan relateras till patientens och förskrivarens beteenden. Patientens beteende syftar till följsamhet (compliance), det vill säga i vilken utsträckning patienten använder läkemedlet. Med förskrivarens beteende menas exempelvis off-label förskrivning vilket innebär förskrivning av läkemedel utanför godkänd indikation vad gäller indikation, dosering, kontraindikationer eller administreringsätt.

4 Metod för översikten

Litteratursökning

Inom ramen för projektet genomfördes litteratursökningar i flera steg av en informationsspecialist. Inledningsvis med breda ansatser för att försöka ringa in ämnesområdet. Redan tidigt konstaterades att begreppsbildningen är heterogen i förhållande till den ursprungliga frågeställningen. Detta medförde svårigheter att skapa stringenta sökstrategier.

Litteratursökningar utfördes i databaserna PubMed via NLM, Scopus via Elsevier och i Google Scholar. Under processen med att identifiera potentiellt relevanta studier har kompletterande metoder som kontroll av referenslistor varit av största betydelse. Sökningen kompletterades med forskningskontakter. Till viss del har handsökning utförts av tidskrifter företrädesvis i "Journal of the Royal Statistical Society. Statistics in Society Series A". I arbetet med litteratursökningar har också projektets experter deltagit. Även statistiska termer från experterna har kontrollsökts och tolv studier, även de utvalda av experterna och projektledaren, har citeringssökts i Scopus och i Google Scholar. Dessa studiers referenslistor kontrollerades också. Sökstrategin redovisas i detalj, se Bilaga 1 på www.sbu.se/256.

Kvalitetsgranskning, sammanläggning och evidensgradering

Eftersom denna rapport inte avser att utvärdera effekterna av en insats eller tillförlitligheten av en diagnostisk metod, har inte de artiklar som befunnits vara relevanta för rapportens frågeställning kvalitetsgranskats på vanligt sätt utifrån olika risker för systematiska fel. Istället har de beskrivna strategiernas användbarhet bedömts utifrån rapportens syfte. I den mån effekter av interventioner ingår i de relevanta artiklarna dras inga slutsatser från utfallen utan resultaten används i första hand för att illustrera hur statistiska strategier använts. Någon formell sammanläggning av studiedata eller evidensgradering av resultat är följaktligen inte heller relevant.

5 Resultat

Litteratursökningarna identifierade nio relevanta artiklar som bedömdes användbara för att besvara uppdragets frågeställning [20–28]. Samtliga artiklar har sin grund i kausal inferens baserat på kontrafaktiska variabler. Metoderna som beskrivs i artiklarna klarlägger de antaganden som är nödvändiga för att man skall kunna genomföra en analys. Det är rimligt att se kausal inferens som ett ramverk inom vilket man kan tillämpa flera specifika skattningsmetoder för behandlingseffekten. Vi kan till exempel använda viktning med inversen av den skattade urvalssannolikheten eller en prediktionsmodell där man använder skattade responsytor. Däremot ger ingen av artiklarna information om vilken av dessa som är bäst i en given situation.

Frågeställningen som rapporten undersöker handlar i grunden om att betrakta populationen i RCT:n som ett urval från den population som finns i klinisk vardag men där patienter har olika sannolikhet att finnas med i RCT:n som således inte är ett vanligt enkelt slumpmässigt urval av populationen. De tillgängliga strategierna syftar till att korrigera resultat från RCT utifrån uppskattade patientspecifika urvalssannolikheter med vissa variationer utifrån vilka data som finns tillgängliga.

Behandlingseffekt kan definieras på olika sätt beroende på vad vi vill ta reda på. Ett mått vi kan utgå ifrån är genomsnittlig behandlingseffekt (average treatment effect, ATE) som mäter skillnaden i utfallens genomsnittliga värde mellan behandlings- och kontrollgruppen, $\bar{Y}_1 - \bar{Y}_0$. Om vi tänker oss en RCT som ett urval från den population vi vill generalisera till så kan vi definiera tre parametrar eller estimander vi kan skatta. Här beskriver vi dessa först med ord och sedan med en mer exakt matematisk definition.

PATE (Population average treatment effect (genomsnittlig populationseffekt)): Detta är effekten av att behandla alla i den bakomliggande populationen med den nya behandlingen jämfört med att inte behandla någon med det nya preparatet.

PATT (Population average treatment effect for the treated (genomsnittlig populationseffekt för de behandlade)): Detta är effekten av behandlingen bland de som i den bakomliggande populationen faktiskt blir behandlade och kan tolkas som effekten av behandlingen för de som blir behandlade i klinisk praxis.

PATC (Population average treatment effect for the controls (genomsnittlig populationseffekt för de obehandlade)): Definieras precis som ovan som effekten av behandlingen bland de i den bakomliggande populationen som fått referensbehandlingen och kan tolkas som effekten av att behandla de som i klinisk praxis ännu inte behandlas med den nya behandlingen.

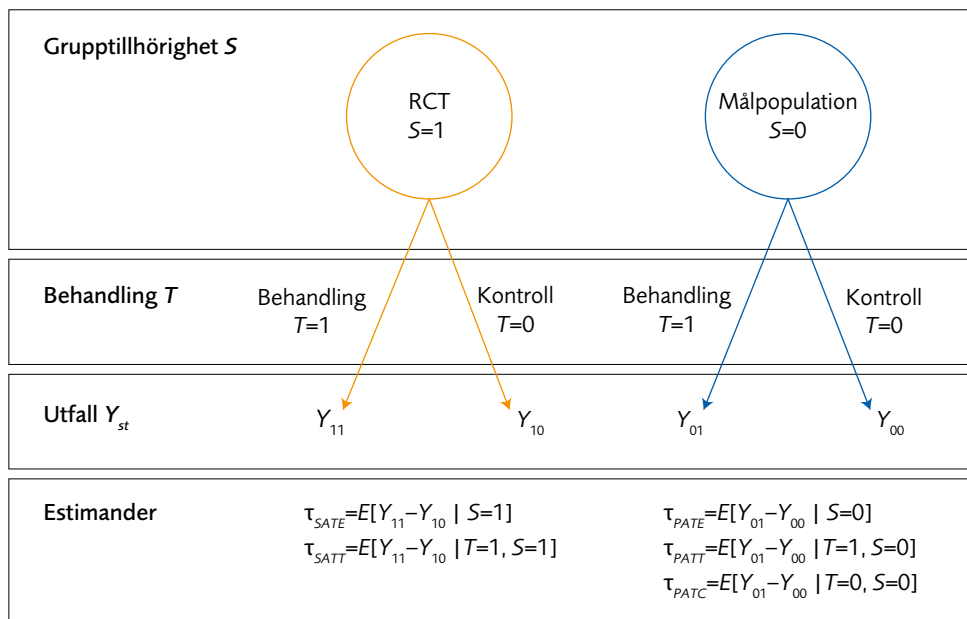
Den matematiska definitionen av estimanderna kräver några variabler. Om vi låter Y_{st} vara det potentiella utfallet för en patient där $s=1$ indikerar att patienten är med i urvalet som utgör RCT:n, $s=0$ indikerar att patienten inte är med och $t=1$ indikerar den nya behandlingen och $t=0$ den gamla (Figur 5.1). Matematiskt definieras PATE som skillnaden mellan det potentiella utfallet under den nya behandlingen Y_{01} om patienten inte selekteras till studien och det potentiella utfallet Y_{00} under kontrollbehandlingen om patienten inte är med i studien betingat med att patienten inte selekteras att delta i studien. Matematiskt skriver vi detta som:

$$\tau_{PATE} = E[Y_{01} - Y_{00} \mid S=0]$$

På liknande sätt kan vi definiera PATT och PATC genom dessutom betinga på att patienten får respektive inte får behandling, vilket vi skriver som:

$$\tau_{PATT} = E[Y_{01} - Y_{00} \mid T=1, S=0]$$

$$\tau_{PATC} = E[Y_{01} - Y_{00} \mid T=0, S=0]$$



Figur 5.1
Skattning av estimanderna görs med hjälp av tre variabler. S betecknar grupptillhörighet det vill säga RCT eller målpopulation. Behandling betecknas med variabeln T och kan antas ha två värden: 1 om individen fick behandling eller 0 om individen fick kontrollbehandling. För utfallsvariabeln Y specificerar vi grupptillhörighet och behandling med nedsänkta siffror.

Samtliga strategier för skattning av populationseffekter förutsätter att det finns en uppsättning variabler W som beskriver patienters sannolikhet att inkluderas i en RCT. Här betecknar $W = (W_1, \dots, W_k)$ en samling (eller vektor) som består av flera variabler såsom exempelvis ålder, kön och olika sjukdomsmarkörer. I de strategier som avser skattning av PATT och PATC kan man tillåta att olika uppsättningar variabler är relaterade till inklusion för patienter som i populationen är behandlande respektive obehandlade och vi betecknar dessa möjligen olika uppsättningar variabler med W^T respektive W^C .

Effekten vi vill skatta, det vill säga PATE, PATT eller PATC kan variera beroende på frågeställningen. De identifierade artiklarna presenterar strategier för att skatta de olika populationsspecifika effekterna (Tabell 5.1). I detta sammanhang syftar ”strategi” till en skattningsmetod som består av flera etablerade statistiska metoder som kan användas i kombination när populationsspecifika effekter skattas. De skattningsmetoder som presenteras i rapporten är de som beskrivits i de identifierade artiklarna. Det går dock inte att utesluta att det finns andra metoder som också kan användas för detta ändamål.

Tabell 5.1
Strategier för skattning
av populationsspecifika
behandlingseffekter.

	Populationsspecifika behandlingseffekter		
	Population Average treatment Effect PATE	Population Average treatment Effect for the Treated PATT	Population Average treatment Effect for the controlled PATC
Vilken effekt vill vi mäta?	Den förväntade effekten av att behandla alla i klinisk vardag med den nya behandlingen jämfört med att inte behandla någon med det nya preparatet	Effekten av behandlingen bland de i klinisk vardag som faktiskt blivit behandlade	Den förväntade effekten av den nya behandlingen för de i klinisk vardag som fått referensbehandlingen (eller ingen behandling)
Under vilka situationer?	Inför ett nytt policybeslut, riktlinjer, subvention av läkemedel	Uppföljning och utvärdering av befintlig behandling	Efter godkännandet men innan läkemedlet har börjat användas fullt ut i klinisk vardag
Exempel	Den förväntade effekten av att behandla alla med ett nytt cancerläkemedel jämfört med standardbehandling	Behandlingseffekten av nya läkemedel mot Hepatit C bland de som förskrivits det nya läkemedlet	Skattning av behandlingseffekten för ett nytt läkemedel mot hyperlipidemi hos de som ännu inte får det i klinisk vardag
Strategier	<ul style="list-style-type: none"> • Justering av RCT resultat till populationen utifrån urvalssannolikheter [21,22,25,26] • Hierarkiska Bayesianska regressionsmodellerna [23] • Skattning av generaliseringsbias med hjälp av inklusions- och exklusionskriterier [24] • Grafiska modeller [28] 	Justering av RCT resultat för jämförelse med behandlade patienter i populationen utifrån urvalssannolikheter [27]	Justering av resultat från kontrollgruppen i RCT:n för jämförelse med patienter i populationen som fått referensbehandlingen [20]

Grundantaganden

De strategier (som baseras på kontrafaktiska utfall) som tas upp i denna rapport bygger alla på ett antal antagande där flera är gemensamma eller likartade. Två helt gemensamma antaganden är SUTVA (Stable Unit Treatment Value Assumption) och konsistens. Antagande om konsistens handlar om att alla patienters potentiella utfall för de undersökta behandlingarna är de samma oavsett om man tillhör urvalet som inkluderas i RTC:n eller inte och kan formuleras matematiskt som:

$$Y_{i01} = Y_{i11}$$

$$Y_{i00} = Y_{i10}$$

Antagandet om SUTVA handlar om att varje individs potentiella utfall inte får bero på vilken behandling som tilldelas andra individer eller om andra individer är med i RCT:n eller inte. Matematiskt kan man skriva detta som:

$$Y_{ist}^{L_i} = Y_{ist}^{L_j} \forall i \neq j,$$

där L_i är behandling och urvalsindikatorerna för patient i .

Skattning av PATE

Skattning av PATE handlar om att utifrån resultat från en (eller flera) RCT samt observationell information, som beskriver målpopulationen, prediktera effekten av att behandla alla i målpopulationen med den nya behandlingen, jämfört med att inte behandla någon med den nya behandlingen.

Strategin bygger på att det finns tillgång till individuella data från både RCT:n och populationen som beskriver urvalet från populationen till RCT:n och att inklusionen i RCT:n inte beror på några andra variabler än de vi har data på. Dessutom måste kontrollbehandlingen i RCT:n vara grund för en relevant jämförelse i populationen. Vidare är det inte möjligt att generalisera utanför de inklusion/exklusionskriterier som används i RCT:n.

Ytterligare antaganden vid skattning av PATE

Skattning av PATE bygger, utöver grundantaganden som beskrivs ovan, på att det inte finns några variabler utöver som påverkar vilka patienter från målpopulationen som skulle kunna komma med i RCT:n. I termer av kontrafaktiska utfall betyder det att de kontrafaktiska utfallen under behandling eller kontroll för de som inkluderas i RCT:n och de som inte gör det, är oberoende av om de faktiskt kommer med i RCT:n, givet deras värde på W . I dagligt tal kan man säga att när vi väl känner patientens värden så är det bara slumpen som gör om denna kommer med i prövningen och matematisk skriver vi detta som:

$$(Y_{0t}, Y_{1t}) \perp S \mid W$$

Antagandet kallas ibland för ”unconfoundedness” vilket tillsammans med antagandet att det inte finns några patienter som har sannolikheten noll eller ett att komma med i RCT:n, det vill säga

$$0 < P(S = 1 \mid W) < 1$$

utgör antagandet om stark ignorerbarhet för urvalet in i RCT:n.

Skattningsmetoder

Utifrån antagandet ovan kan man använda flera metoder för att skatta PATE. Nedan beskrivs endast en av flera tillgängliga skattningsmetoder. Denna bygger på viktning och tas upp i flera artiklar [21,22,25,26]. Metoden går ut på att man för varje patient i RCT skattat sannolikheten att denne är med i prövningen, det vill säga:

$$e_i = P(S_i = 1 \mid W_i)$$

Man kan betrakta dessa sannolikheter på samma sätt som de sannolikheter man använder i propensity score (PS) metoder. I vanliga fall kan PS-metoder användas när vi vill jämföra en behandlingsgrupp och en kontrollgrupp som skiljer sig med avseende på fördelning av bakgrundvariablerna. Om vi känner till individens bakrundsvariabler kan vi skatta ett PS som avspeglar individens sannolikhet att få en behandling. Individer med samma PS-värde antas då ha samma fördelning av bakgrundvariablerna. Vi kan sedan välja två individer med likadant PS-värde, en från behandlingsgruppen och en från kontrollgruppen, och jämföra deras utfall. Till skillnad från denna standard PS-användning, modellerar PS i artiklarna inkluderbarhet i RCT vs. målpopulation, det vill säga sannolikheten att ingå i en RCT givet ett set av uppmätta bakgrundsviabler.

Rent praktiskt kan man skatta med en logistisk regression men man bör då noga överväga vilken modell man använder så att modellen innehåller relevanta interaktionstermer och möjligen högre gradseffekter. Som alternativ kan exempelvis Random Forest eller Gradient Boosting användas där man inte behöver specificera en modell explicit.

De skattade sannolikheterna används i sin tur för att definiera vikter vilka används till att justera analysen av RCT:n så att resultatet nu utgör en skattning av PATE. Syftet med vikterna är att få RCT-observationerna att likna observationerna i målpopulationen. Ett exempel kan vara en RCT där 20 procent av deltagarna är kvinnor medan andelen kvinnor i målpopulationen är 50 procent. Genom viktning kan vi ge mer vikt till kvinnorna i RCT:n så att resultatet bättre speglar fördelningen i målpopulationen.

De flesta vanligen använda statistiska programvarorna erbjuder möjligheter att använda vikter i analysprocedurerna, men man bör vara noggrann då standardfelen inte alltid skattas korrekt utan att man även gör andra anpassningar.

Vi kan summera metoden med följande steg:

1. Anpassa en modell som predicerar deltagandet i RCT som funktion av bakgrundsvariablerna, (till exempel logistisk regression). Här får vi fram våra PS.
2. Skapa vikter för individer i RCT:n med hjälp av de framräknade PS (steg 1).
3. Uppskatta behandlingseffekt med hjälp av RCT-individerna genom att göra en viktad regression av utfallen som en funktion av behandlingsstatus och bakgrundsvariabler, med vikterna från steg 2.

Skattning av PATT

PATT är effekten av behandlingen bland de som i målpopulationen faktiskt blir behandlade och kan tolkas som effekten av behandlingen såsom den används i klinisk praxis. Skattning av PATT förutsätter tillgång till information om exponering för populationen vi avser generalisera till samt att den behandling vi vill utvärdera används i den aktuella populationen.

Ytterligare antaganden vid skattning av PATT

Skattning av PATT bygger utöver grundantaganden på likande antaganden om stark ignorerbarhet som skattning av PATE med den skillnaden att vi tillåter att inklusion i RCT:n kan styras av olika variabler för patienter som i populationen är exponerade respektive oexponerade. Antaganden om stark ignorerbarhet måste därför delas upp på behandlade respektive obehandlade (kontroll) vilket för de behandlade skrivs som:

$$(Y_{01}, Y_{11}) \perp S \mid (W^T, T = 1),$$

$$0 < P(S = 1 \mid W^T, T = 1) < 1$$

Och för kontrollgrupp:

$$(Y_{00}, Y_{10}) \perp S \mid (W^C, T = 0),$$

$$0 < P(S = 1 \mid W^C, T = 0) < 1$$

Om alla antaganden är uppfyllda kan vi konstruera en skattning av behandlingseffekter i populationen. Effekten för de behandlade blir då:

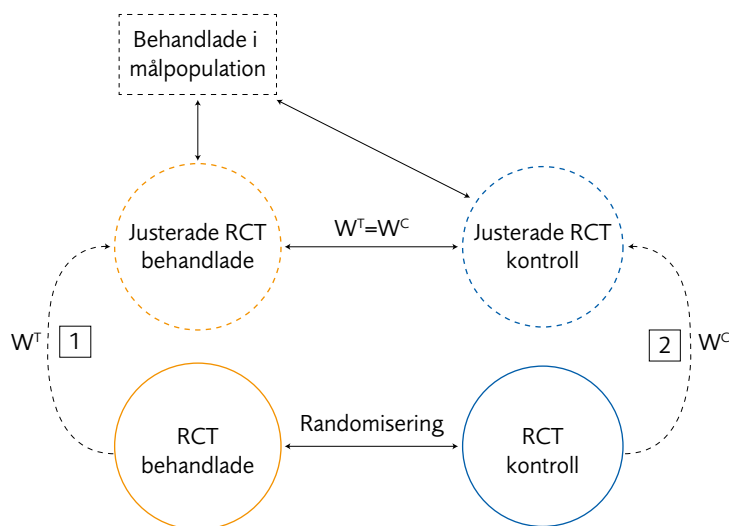
$$PATT = E_{01} \{E\{Y \mid W^T, S=1, T=1\}\} - E_{01} \{E\{Y \mid W^C, S=1, T=0\}\}$$

vilket betyder att PATT kan skattas utifrån utfallsdata för RCT:n och information om hur urvalet till endast RCT:n beror av observerade variabler. Det båda termerna i högerledet är väntevärden över fördelningen för respektive och representerar utfallen för behandlade och kontrollgrupp i RCT justerat för selektion in i RCT:n.

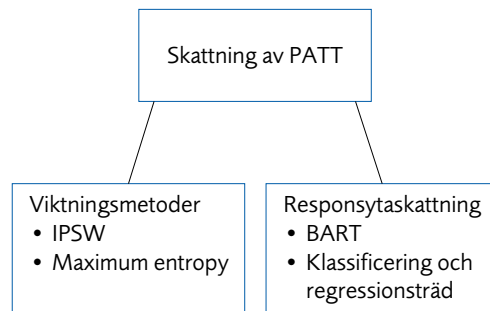
Skattningsmetoder

En strategi för skattning av PATT beskrivs av Hartman och medarbetare [27]. Bakgrundsvariabler för behandlade individer i RCT, W^T , kan med hjälp av viktningmetoder justeras för att matcha bakgrundsvariablerna hos behandlade individer i målpopulationen (Figur 5.2, steg 1). Med hjälp av den justerade behandlingsgruppen från RCT:n (streckade röd cirkel i Figur 5.2) kan vi återskapa det potentiella utfallet Y_{s1} för behandlade individer i målpopulationen. Nästa steg blir att justera bakgrundsvariablerna hos RCT:ns kontrollgrupp till bakgrundsvariablerna hos behandlade individer i målpopulationen (Figur 5.2, steg 2). På samma sätt kan vi använda den justerade kontrollgruppen från RCT:n (streckad blå cirkel i Figur 5.2) för att återskapa det potentiella utfallet Y_{s0} . Den justerade kontrollgruppen från RCT:n ersätter kontrollindivider i målpopulation som har samma bakgrundsvariabler som behandlingsgruppen i målpopulation. De obehandlade (kontroll) och behandlade individerna antas inte vara utbytbara och därför används den justerade kontrollgruppen från RCT:n. Kontroll- och behandlingsgruppen i RCT:n är utbytbara och justering av dessa grupper, som är baserad på samma bakgrundsvariabler (det vill säga $W^T = W^C$), skapar utbytbara grupper. Om utbytbarhet råder kan vi använda utfallen från de justerade grupperna för att skatta PATT.

Figur 5.2
Justerung av RCT-individer (behandlade och kontroll) till behandlade individer i målpopulationen. Dubbelpilar representerar utbytbarhet av potentiella utfall, och streckade pilar representerar justering av bakgrundvariablernas fördelning. Bilden är tagen från Hartman och medarbetare [21].



Utifrån dessa antaganden kan man skatta PATT med hjälp av justerade RCT-data och det finns flera sätt att rent praktiskt göra detta (Figur 5.3).



Figur 5.3
Skattningsstrategier av PATT kan delas i två klasser: viktningmetoder och responsytaskattning.

IPSW=Inverse propensity score weighting; **BART**=Bayesian additive regression tree

Den första klassen av strategier består av viktningmetoder, där RCT-data vägs med hjälp av information om målpopulationen. Den andra klassen av strategier går ut på att skatta en responsyta med hjälp av RCT-data för att sedan extrapolera denna till behandlade individer i målpopulationen. Responsytan är egentligen det resultat vi genererar vid regressionsanalysen när vi undersöker relationen mellan responsvariabeln (utfallet) och de oberoende variabler (bakgrundsvariablerna).

Hartman och medarbetare beskriver en estimeringsstrategi som bygger på viktningmetoder. I det första steget skapas matchande par i RCT där varje par består av en individ från behandlingsgruppen och en matchande individ från kontrollgruppen [27]. Matchningsmetoden heter Genetic Matchning (GenMatch) och detaljer kring metoden kan hittas i studien av Diamond och Sekhon [29]. Vid nästa steg viktas de matchande paren från RCT:n till de behandlade individerna i målpopulationen. Viktningmetoderna som testades var IPSW och Maximum entropy.

I samma artikel redovisas även en alternativ estimeringsstrategi som bygger på responsytan [27]. Genom att studera relationen mellan utfall och bakgrundsvariabler i RCT:n kan man skatta en responsyta och prediktera respons under behandling samt utan behandling bland de behandlade i målpopulationen. Författarna föreslår BART (Bayesian adaptive regression trees) för detta ändamål, men det finns inget som hindrar att man använder andra enklare modeller.

Placebotest

I de fall där vi har tillgång till utfallsdata från individerna i målpopulationen kan ett placebotest konstrueras. Ett placebotest syftar till att empiriskt undersöka delar av de antaganden som ligger till grund för metodiken. Testet bygger på en jämförelse mellan utfallet bland de behandlade i målpopulationen och den justerade skattningen av utfallet av behandlingsgruppen i RCT:n. Om placebotestet misslyckas kan det bero på att individernas icke-observerbara bakgrundsvariabler eller att standardbehandlingen skiljer sig mellan RCT:n och målpopulationen [20,27]. För mer detaljer om detta test, se Hartman och medarbetare [27].

Skattning av PATC

Den genomsnittliga populationseffekten för de obehandlade, är den förväntade effekten av en ny behandling bland de i klinisk vardag som fått standardbehandling. Detta kan tolkas som effekten av att behandla de i klinisk vardag som ännu inte behandlas med den nya behandlingen. Situationen kan exempelvis uppstå vid införande av ett nytt läkemedel där endast de som har fått behandlingen är RCT-deltagarna. I sådana situationer kan vi definiera målpopulationen som patienter i klinisk vardag som kan bli aktuella för den nya behandlingen men som i nuläget får standardbehandling. Vi kan då skatta PATC genom att jämföra viktade utfall för kontrollgruppen i RCT:n med patienter som får standardbehandling i klinisk praxis. Ett exempel på denna strategi beskrivs av Stuart och medarbetare [20].

Övriga strategier

Hierarkiska Bayesianska regressionsmodellerna

Ett exempel på en strategi där man drar slutsatser om patientkategorier som inte ingår i kliniska studier är de Hierarkiska Bayesianska regressionsmodellerna som i sin tur omfattar underordnade modeller för data från studier med olika typer av studiedesign [23]. En typ av data är RCT:er där information endast finns på aggregerad nivå. En annan typ av data är prospektiva kohorter där man har tillgång till information på individnivå (till exempel kvalitetsregister). Ett krav är att samma utfallsvariabler finns i samtliga studier inklusive de prospektiva kohorterna.

Uppskattning av generaliserbarhetsbias

En annan strategi som inte kräver RCT-individdata beskrivs av Pressler och Kaizar [24]. I artikeln undersöks hur ett RCT-baserat estimat för behandlingseffekt påverkas av inklusion- och exklusionskriterier.

Utgångspunkten för studien är de skillnader i behandlingseffekt mellan RCT och målpopulation som uppstår på grund av andelen exkluderade patienter i RCT:n. Behandlingseffektens storlek skiljer sig således hos patienter som uppfyller inklusionskriterierna och patienter som inte uppfyller dem. Ett sätt att studera heterogeniteten i behandlingseffekt är att undersöka det bias som uppkommer när vi generaliserar RCT-resultat till en bredare population. Författarna kallar detta bias för ”generalisability bias”, eller generaliserbarhetsbias.

I denna metod används RCT:ns inklusion- och exklusionskriterier för att skapa två grupper i målpopulationen. Den ena gruppen består av individer som uppfyller RCT:ns inklusionskriterier (grupp I). Den andra gruppen består av individer som inte uppfyller inklusionskriterierna, det vill säga exkluderade individer (grupp E). Med hjälp av observationsdata skattas behandlingseffekt separat för varje grupp (PATE (I) och PATE (E)). Genom att sedan räkna

skillnaden mellan PATE (I) och PATE (E) och därefter vikta den mot andelen exkluderade i målpopulationen kan vi skatta generaliserbarhetsbias.

Utöver kännedom om inklusion- och exklusionskriterier, kräver metoden tillgång till observationsdata om de behandlingar som ska jämföras. Metoden är därför inte användbar när behandlingseffekt av en ny behandling ska studeras i målpopulationen.

Grafiska modeller

I denna rapport har vi använt potentiella utfall som ett ramverk för att resonera kring kausalitet. Ett alternativ är att använda grafiska modeller [30]. Även inom grafiska modeller kan vi använda samma statistiska metoder, exempelvis viktning. Båda modellerna har för- och nackdelar och möjliggör olika antaganden. Den praktiska användningen av Pearls ramverk är inte lika väl beskriven för vår frågeställning och därför kanske något av ett andrahandsval i dagsläget [28].

6 Datakällor

Vilka datakällor från klinisk vardag finns att tillgå?

Genom systemet med personnummer, datoriserade journaler och förekomsten av lagstadgade nationella register har vi i Sverige i teorin goda förutsättningar att följa enskilda patienters väg från förskrivning av ett läkemedel till behandlingsutfall. Eftersom olika register och datakällor kan samköras med varandra finns heller inget krav på att en och samma datakälla ska innehålla all information som behövs för en uppföljning.

Nedan följer en kort redovisning av de primära källorna till data från klinisk vardag som finns att tillgå. Viktiga begränsningar i dessa datakällor avseende möjligheter till uppföljning av läkemedelsanvändning och behandlingsutfall lyfts fram. Det kan också finnas legala begränsningar i hur individbaserade uppgifter från dessa datakällor får användas. En genomgång av legala aspekter har inte genomförts inom ramen för denna rapport.

Läkemedelsregistret

För att kunna följa upp behandlingseffekter av läkemedel är det till att börja med helt centralt att de individer som fått ett visst läkemedel kan identifieras. Den främsta källan till data på läkemedelsanvändning är Läkemedelsregistret som är ett hälsodataregister som förvaltas av Socialstyrelsen och innehåller uppgifter om samtliga läkemedel som förskrivs och hämtats ut på svenska öppenvårdsapotek.

I Läkemedelsregistret finns bland annat uppgifter om produktnamn, förpackningsstorlek och förskrivningsdatum för den produkt som expedierats. Uppgiftsinhämtning sker via eHälsomyndigheten, som tar emot data från apoteken och därefter skickar uppgifterna vidare till Socialstyrelsen. Läkemedelsregistret uppdateras månadsvis med nya uppgifter från apoteken.

Sedan juli 2005 innehåller Läkemedelsregistret också uppgifter om personnummer för den patient som ordinerats läkemedlet, något som öppnat upp för möjlighet att samköra med andra register och därmed påtagligt utökat registrets nytta.

Läkemedelsregistret har ett stort värde i uppföljningssammanhang men på två viktiga områden är det otillräckligt:

- För det första innehåller registret inga uppgifter om läkemedel som ordineras inom slutenvården eller på öppenvårdskliniker där patienten inte själv hämtat ut läkemedlet på recept. Läkemedelsregistret ger således ingen möjlighet att identifiera patienter som behandlas med så kallade klinisläkemedel/rekvisitionsläkemedel.
- För det andra saknar Läkemedelsregistret uppgifter om vilket användningsområde, eller så kallad indikation, som en produkt förskrivits inom. Många nya läkemedel, inte minst inom onkologi, får successivt godkännanden för fler indikationer. Uppföljning av dessa läkemedel skulle behöva vara indikationsspecifik men någon sådan möjlighet ger i dagsläget inte Läkemedelsregistret.
- Man kan genom att jämföra data i läkemedelsregistret och förskrivningsdata få ett mått på patienternas följsamhet gentemot ordinationen i form av en skattning av hur mycket läkemedel som hämtats från apotek i relation till hur mycket som skrivits ut. Om patienten sedan undlåter att läkemedlet är en fråga som inte går att komma åt med hjälp av register.

I slutrapporten till SKL:s projekt Individdata om rekvisitionsläkemedel som drivits inom Nationella läkemedelsstrategin område 6.4 ges ett antal konkreta förslag till vidare arbete för att möjliggöra insamling av individdata om rekvisitionsläkemedel till Läkemedelsregistret [31].

Övriga hälsodataregister

Bland Socialstyrelsens övriga hälsodataregister är Patientregistret och Cancerregistret de som har störst relevans i uppföljningssammanhang. Även Dödsorsaksregistret kan innehålla användbara data.

Patientregistret innehåller uppgifter om diagnoser, kvalificerade åtgärds-koder och operationskoder för vårdtillfällen i slutenvård samt i den öppna specialistvården. Dessa kan vara till nytta när det gäller att identifiera subgrupper av patienter som behandlats med ett visst läkemedel, men har begränsad nytta som utfallsmått i uppföljningssammanhang eftersom fokus då ligger på medicinska resultat. För läkemedel som utvärderats i kliniska prövningar med hårda

utfallsmått (exempelvis kardiovaskulära händelser) är Patientregistret däremot en värdefull källa till uppföljningsdata eftersom utfallet som observeras i registret då kan ställas i direkt relation till vad som observerats i de kliniska studierna. Tyvärr täcker Patientregistret i dagsläget inte primärvården.

Cancerregistret skulle kunna bli ett värdefullt register för uppföljning av behandlingseffekter med tanke på att många nya läkemedel, där uppföljning är som mest relevant, är onkologiska preparat. Uppgiftsskyldighet gäller för nyupptäckta tumörer och tumörliknande tillstånd som upptäcks vid kliniska undersökningar och laborietester. Cancerregistret innehåller dock inga uppgifter om läkemedelsbehandling. Dessa skulle kunna hämtas in genom samkörning med en annan datakälla men den möjligheten begränsas av att Läkemedelsregistret inte innehåller information om läkemedel som administreras i slutenvården. Ett betydande hinder för uppföljning av läkemedel är att upprepade mätpunkter för samma patient saknas i Cancerregistret, så att behandlingsutfall och sjukdomsförlopp inte kan följas över tid.

Nationella kvalitetsregister

Enligt patientdatalagen definieras ett kvalitetsregister som en automatiserad och strukturerad samling av personuppgifter som inrättats särskilt för ändamålet för att systematiskt och fortlöpande utveckla och säkra vårdens kvalitet. Registren ska enligt samma lag möjliggöra jämförelser inom hälso- och sjukvården på nationell och regional nivå. Till skillnad från Socialstyrelsens hälsodataregister har patienten möjlighet att begära att data från dem raderas ur kvalitetsregistren.

Genom en överenskommelse mellan SKL och staten har en särskild satsning på de nationella kvalitetsregistren gjorts under åren 2012–2016. Myndigheten för vård- och omsorgsanalys har fått i uppdrag att utvärdera satsningen som omfattar cirka en och en halv miljard kronor över fem år.

För närvarande finns 96 nationella kvalitetsregister som varierar stort i fråga om täckningsgrad, det vill säga hur stor andel av den avsedda patientpopulationen som återfinns i registret. År 2015 hade 72 procent av registren en täckningsgrad på över 80 procent och 41 procent en täckningsgrad på över 95 procent. Den högsta certifieringsnivån (nivå 1) som bland annat ställer krav på validering av kvaliteten på data och hög täckningsgrad utifrån aktuell patientgrupp uppnåddes av 13 register, däribland

- Nationella Diabetesregistret (NDR)
- Swedeheart (Nationellt register för hjärtintensivvård, kranskärlsröntgen, PCI, hjärtkirurgi och sekundärprevention)
- Svensk Reumatologisk Kvalitetsregister (SRQ) [32]
- Infektionsregistret InfCare (HIV)

Många av kvalitetsregistren kräver fortfarande manuell registrering av uppgifter men det finns pågående projekt för automatiserad datainsamling som bedrivs under Nationella programmet för datainsamling (NPD_i), vars övergripande syfte är att förenkla datainsamling till kvalitetsregister och begränsa dubbelregistrering.

När det gäller uppföljning av behandlingseffekter ser förutsättningarna väldigt olika ut i de olika registren. De flesta kvalitetsregister fokuserar på en viss diagnos/åtgärd och inte alla omfattar läkemedelsinformation. Eftersom uppföljning av behandlingseffekter är särskilt angelägen för nya läkemedel som är föremål för nationellt ordnat införande, är det en viktig fråga om de nationella kvalitetsregistren kan tillhandahålla de data som krävs för att kunna följa upp dessa läkemedel, som ofta är kostsamma och omgärdade av osäkerheter i fråga om kostnadseffektivitet. I många fall finns formella krav på att dessa läkemedel ska följas upp inom ramen för riskdelningsavtal som slutits i samband med TLV:s subventionsbeslut. Det kan också synas rimligt att användningen av nya, mycket kostsamma behandlingar förenas med krav på uppföljning av effekt och säkerhet, även om det för närvarande kan medföra ett extra arbetsmoment för forskrivarna. Det är dock angeläget att insamling av data för uppföljning i största möjliga utsträckning kan underlättas med hjälp av digitala lösningar som inte kräver dubbelregistrering av data i både patientjournal och kvalitetsregister. I vissa register är det möjligt att införa nya uppföljningsparametrar för en speciell behandlingskohort, medan detta är svårare i andra register.

Vid en genomgång av de uppföljningsprotokoll som tagits fram för läkemedel som är föremål för ordnat införande på nivå 1, syns en uttrycklig ambition att följa upp behandlingseffekter i klinisk vardag och i flera fall föreslås något av kvalitetsregistren som källa till data. Tabell 6.1 redovisar en sammanställning av de läkemedel i nivå 1 av den nationella processen samt de kvalitetsregister som föreslagits ligga till grund för uppföljning av dessa läkemedel [33]. Den sammantagna bilden är att kvalitetsregistren inte räcker till för att tillgodose de krav på data som krävs för att följa upp behandlingseffekter av dessa läkemedel. Flera av uppföljningsprotokollen nämner kvalitetsregistret "Nya läkemedel i cancervården" som källa till uppföljningsdata. Eftersom många nya cancerläkemedel kommer att introduceras under de kommande åren är detta register viktigt för möjligheten att följa upp dessa. Kvalitetsregistret Nya läkemedel i cancervården är i full drift sedan 1 november 2015 men har i dagsläget en otillräcklig täckningsgrad.

Läkemedel	Indikation	Kvalitetsregister som föreslås i uppföljningsprotokollet
Evolocumab (Repatha)	Primär hyperkolesterolemi	<ul style="list-style-type: none"> Kardiogenetikregistret, som för närvarande är under uppbyggnad
Guanfacin (Intuniv)	ADHD hos barn och ungdomar 6–17 år	<ul style="list-style-type: none"> På grund av begränsade resurser har NT-rådet beslutat att introduktionen av Intuniv inte kommer att följas upp på nationell nivå. BUSA är ett nationellt kvalitetsregister för ADHD. Det har för närvarande för låg täckningsgrad för att användas för uppföljning av introduktionen av Intuniv
Sakubitril-valsartan (Entresto)	Kronisk symptomatisk hjärtsvikt hos vuxna med nedsatt ejektionsfraktion	<ul style="list-style-type: none"> RiksSvikt. Data ska rapporteras till detta kvalitetsregister i samband med behandlingsstart. Fokus är att registrera NYHA-klass (svårighetsgrad för hjärtsvikt) samt LVEF (vänsterkammarens ejektionsfraktion, det vill säga hjärtmuskelfunktion) enligt dokument/mallar
Nivolumab (Opdivo)	Icke småcellig lungcancer av icke-skivepiteltyp	<ul style="list-style-type: none"> Kvalitetsregistret "Nya läkemedel i cancervården" På sikt nationellt kvalitetsregistret för lungcancer, som är under utveckling för att kunna följa läkemedelsbehandlingar
Nivolumab (Opdivo)	Icke-småcellig lungcancer av skivepiteltyp	<ul style="list-style-type: none"> Kvalitetsregistret "Nya läkemedel i cancervården" På sikt nationellt kvalitetsregistret för lungcancer, som är under utveckling för att kunna följa läkemedelsbehandlingar
Nivolumab (Opdivo) och pembrolizumab (Keytruda)	Malignt melanom	<ul style="list-style-type: none"> Kvalitetsregistret "Nya läkemedel i cancervården"
Abirateron (Zytiga) och enzalutamid (Xtandi)	Prostatacancer	<ul style="list-style-type: none"> Inget kvalitetsregister angivet
Olaparib (Lynparza)	Ovarialcellscancer	<ul style="list-style-type: none"> Kvalitetsregistret "Nya läkemedel i cancervården" Kvalitetsregistret för gynekologisk cancer
Sofosbuvir (Sovaldi), Simeprevir (Olysio), Daclatasvir (Daklinza), Ledipasvir/sofosbuvir (Harvoni), Ombitasvir/paritaprevir/ritonavir (Viekirax) och Dasabuvir (Exviera)	Hepatit C	<ul style="list-style-type: none"> InfCare Hepatit

Tabell 6.1
Sammanställning av datakällor som föreslås ligga till grund för uppföljning av läkemedel med nivå 1 i det nationella ordnade införandet*.

* Baserat på en genomgång av de införande- och uppföljningsprotokoll publicerade på <http://www.janusinfo.se/Nationellt-inforande-av-nya-lakemedel/> den 6 september 2016.

Vårddatabaser

Ett antal landsting använder sig av elektroniska vårddatabaser som samlar data över all vård som produceras i landstinget eller konsumeras av landstingets invånare. Exempel på elektroniska vårddatabaser är VAL i Stockholms Läns Landsting och VEGA i Västra Götalandsregionen, Region Skånes Vårddatabas och Vårddatalagret i Östergötland.

Vårdgivare som har slutit avtal med dessa landsting är skyldiga att leverera elektronisk information om vårdhändelser till respektive landsting för inläsning i databaserna. Informationen används sedan för att följa upp behov, volymer och effektivitet i vården och som ett underlag för utbetalda ersättningar till vårdgivarna.

VEGA och vårddatalagret saknar uppgifter om läkemedel. Vad gäller VAL-databasen innehåller den individbaserade uppgifter om läkemedel som hämtats på recept men inte läkemedel som administrerats på sjukhus (rekvisitionsläkemedel). VAL lider således av samma begränsning som Läkemedelsregistret. Aggregerade uppgifter om läkemedel som rekvirerats till en vårdinrättning finns emellertid i VAL-databasen [34].

När det gäller behandlingsutfall innehåller vårddatabaserna inte några medicinska resultat utan endast uppgifter om besöksdiagnoser och utförda åtgärder. Detta gör att det idag är svårt att använda dem för att utvärdera behandlingseffekter.

Journalssystem

I ett fullt utbyggt journalssystem bör uppgifter om såväl läkemedelsanvändning som behandlingsutfall registreras vilket möjliggör uppföljning av behandlingseffekter om registrering skett på ett standardiserat sätt. Eftersom data på öppenvårdsläkemedel finns i Läkemedelsregistret är det framför allt möjligheterna att extrahera data på rekvisitionsläkemedel som står i fokus när det gäller läkemedelsuppgifter i journalssystem.

Resultaten från det så kallade INNOVIS-projektet och Stockholms läns landstings (SLL) arbete inom projektet Individdata om rekvisitionsläkemedel [31] visar att det är möjligt att från journalsystemet TakeCare extrahera data om rekvisitionsläkemedel inom slutenvård eller sjukhusdriven öppenvård och koppla samman dessa data med medicinska journaldata. Undantaget är vissa cancerläkemedel och andra läkemedel med komplexa beredningar där de läkemedelsmoduler som finns i journalsystemen inte är anpassade, utan specifika ordinationssystem krävs [34]. Ett sådant ordinationssystem (CytoDos) har nyligen implementerats på Karolinska Universitetssjukhuset. Även om möjligheten ännu inte utvärderats torde det vara möjligt att inhämta data om läkemedelsanvändning från denna typ av ordinationssystem. Inom SLL har uppgifter från apotekets beredningscentraler också använts som källa till data på läkemedel med komplexa beredningar, framför allt cytostatika [35].

Erfarenheterna av de extraktioner av individdata på rekvisitionsläkemedel som genomförts visar att det går att få fram data med hög täckningsgrad. Men dokumentationen var inte enhetlig och det krävdes manuell bearbetning för att harmonisera doseringar och strukturera data. En automatisering av dessa processer bör var möjlig och är alldeles nödvändig om uppföljning i stor skala ska bli verklighet.

Även om en rad olika journalsystem används på nationell nivå domineras marknaden av ett fåtal aktörer. För sjukhusen och psykiatrien har fyra dominerande system 92,4 procent av marknaden uttryckt i antalet användare. Inom primärvården har fyra leverantörer 100 procent av marknaden [36]. Därmed kan antalet journalsystem anses vara hanterbart. Men många menar att en strategisk utveckling mot en nationell strukturerad vårdokumentation krävs för att fullt ut kunna genomföra de studier som behövs för att utvärdera nya läkemedel.

Prövningsdatabaser

En förutsättning för att de strategier som presenteras i denna rapport ska vara användbara är att de företag som äger data från de pivotala kliniska prövningarna ställer data från aktuella prövningsdatabaser till förfogande. Under arbetet med denna rapport har ett stort intresse uttryckts från representanter från ett antal företag, att samverka omkring arbetet med att värdera behandlingars effektivitet i klinisk praxis. Detta kräver emellertid att de som i praktiken ska genomföra databearbetningen och analysen får tillgång till avidentifierade individdata. För att resultaten av sådana analyser ska bli trovärdiga, krävs en öppenhet om hur data genererats. Vissa företag kan dock komma att ha invändningar på att rådata från deras kliniska prövningar blir transparenta. Det är därför viktigt att möjligheten att sekretessbelägga sådana data klagas. Det torde vara lättare att hantera dessa frågor för myndigheter som har legal möjlighet att skydda data som är känslig ur affärssynpunkt.

Om istället företagen, inom ramen för ett myndighetsbeslut, åläggs att ta fram data av den typ som beskrivs i denna rapport, kan känsliga patientdata komma att behöva hanteras av företagen. Detta förekommer redan idag och har oftast lösts genom avidentifiering av registerdata. Alternativt kan ett oberoende organ (exempelvis en forskargrupp inom ett universitet) utföra arbetet på uppdrag av företaget.

Det är inte säkert att företagens prövningsdataregister innehåller all data om viktiga faktorer av betydelse, för att kunna använda de föreslagna strategierna. Detta kommer säkerligen att variera från fall till fall. Förhoppningsvis kommer företagen bli mer angelägna om att registrera sådan data om myndigheter och finansärer i olika länder börjar efterfråga denna typ av uppföljningar.

7 Ekonomiska aspekter

I ett samhällsekonomiskt perspektiv är det viktigt att beslut om introduktion av nya behandlingsmetoder såsom nya läkemedel utvärderas utifrån förväntad effekt och patientnytta samt resursanspråk och kostnader. Kunskapen om den förväntade effekten och därmed patientnyttan, men också om storleken på förväntade resursanspråk och kostnader utvecklas efter hand under ett läkemedels livscykel. De metoder som belyses i denna rapport väntas få betydelse för processerna både för att bedöma ett läkemedels kostnadseffektivitet vid introduktionen och då nya data blir kända, vid en strukturerad uppföljning av läkemedlets effekt och kostnadseffektivitet i klinisk praxis. Det är inte metoden för att utvärdera kostnadseffektivitet i sig som påverkas, men däremot kan slutsatser om kostnadseffektivitet påverkas av såväl bedömningen om hur säkert och relevant dataunderlaget som används i analysen är, som i de fall då kunskapen om effektstorleken ändras. För att få en så tillförlitlig skattning av kostnadseffektiviteten som möjligt är det viktigt att en strukturerad uppföljning också sammanställer uppgifter om faktisk resursanvändning och kostnader vid användning av det nya läkemedlet.

I Sverige har det hittills funnits två processer för introduktion av nya läkemedel. Vid introduktionen av ett nytt läkemedel som avses förskrivas i öppenvården granskas läkemedlets kostnadseffektivitet av TLV. Om TLV gör bedömningen att det nya läkemedlet är kostnadseffektivt inkluderas det i läkemedelsförmånen. TLV gör bedömningen utifrån underlag från företaget och kan dessutom begära in kompletterande underlag. Läkemedel som används inom slutenvården, ibland kallade rekvisitionsläkemedel, upphandlades tidigare av landstingen utan att genomgå en formell process för bedömning av kostnadseffektivitet. På senare år har detta förfarande ändrats och nu görs en strukturerad bedömning på nationell nivå av kostnadseffektiviteten även för utvalda nya läkemedel avsedda

för slutenvården. I denna process samverkar landsting och regioner via SKL och dess råd för nya terapier, NT-rådet, med TLV för att ta fram underlag om kostnadseffektivitet och rekommendationer om nya läkemedel. Även om processerna ser olika ut är det samma metoder och principer som används för att analysera kostnadseffektiviteten [37].

Modellanalyser av kostnadseffektivitet

Bedömning av ett nytt läkemedels kostnadseffektivitet görs vanligen i modellanalyser som bygger på den skattade effekten av läkemedlet från den eller de studier som legat till grund för läkemedlets regulatoriska godkännande. Effekten i de kliniska prövningarna är vanligen ett kliniskt relevant utfallsmått, men detta mått är samtidigt ett intermediärt mått i förhållande till patientnära utfall såsom livskvalitet och överlevnad. I modellanalyser används därför vanligen kompletterande datakällor såsom epidemiologiska studier av överlevnad och relevanta sjukdomshändelser som kopplas till tillståndet som behandlas, beskrivande studier av resursanvändning och kostnader för patientgrupper och behandlingar samt studier av livskvalitet kopplat till delar av händelser och tillstånd i ett sjukdomsförlopp. Det är nödvändigt att använda dessa kompletterande källor så snart läkemedel och andra behandlingar kan förväntas påverka sjukdomsutfall, patientnytta och kostnader under längre perioder än de studieperioder som används i de randomiserade kontrollerade prövningar som ligger till grund för det regulatoriska godkännandet. Eftersom kliniska prövningar vanligen har korta uppföljningstider (från några månader till två år) blir det i normalfallet att en studie av kostnadseffektivitet är en modellanalys som också inkluderar kompletterande datakällor.

För läkemedel som används vid kroniska sjukdomar och tillstånd är det särskilt angeläget att med simulerande modellanalyser pröva olika scenarier där data från kliniska prövningar extrapoleras för att undersöka hur robust skattningen av kostnadseffektivitet är under olika antaganden. I tabellen nedan illustreras tre exempel på sjukdomar och tillstånd tillsammans med effektmått som är vanliga i kliniska prövningar samt vilka ytterligare uppgifter som hämtas från andra källor till den hälsoekonomiska modellanalysen.

Sjukdom/tillstånd	Vanligt effektmått för behandlingsutfall i klinisk prövning	Uppgifter som kompletteras från andra källor i analys av kostnadseffektivitet
Bröstcancer	Progressionsfri överlevnad	Total överlevnad Livskvalitet Resursanvändning, kostnader
Diabetes	Primärt utfallsmått: Blodglukoskontroll mätt med HbA _{1c} Sekundära utfallsmått: Viktpåverkan, förekomst av hypoglykemi	Total överlevnad Diabeteskomplikationer såsom mikro- och makrovaskulära händelser Livskvalitet Resursanvändning, kostnader
Förmaksflimmer	Ischemisk stroke Hjärnblödning	Total överlevnad Större gastrointestinal blödning, hjärtinfarkt Livskvalitet Resursanvändning, kostnader

Tabell 7.1
Data till hälsoekonomiska modellanalyser av kostnadseffektivitet: Tre exempel på sjukdomar/tillstånd och effektdata från klinisk prövning samt kompletterande uppgifter från andra datakällor.

Som framgår av Tabell 7.1 är effektdata från en RCT ett av flera underlag som ingår i analysen av kostnadseffektivitet. I princip är det tekniskt enkelt att ändra parametrar i hälsoekonomiska modeller och i så motto är en ny skattning eller omvärdering av effektstorleken från den ursprungliga RCT:n inte ett undantag. Det som är centralt är därför säkerheten och relevansen i det parametervärde som används i modellanalysen. Hälsoekonomiska modeller byggs vanligen för att användaren ska kunna variera modellens centrala parametrar för att pröva hur robusta slutsatserna om kostnadseffektivitet är. Effekten av läkemedelsbehandling är naturligtvis ett av de centrala måtten som inte bara kan, utan också bör varieras i känslighetsanalyser. Det betyder att ny kunskap om effektstorlek anpassad till svensk klinisk praxis enligt strategierna som beskrivs i denna rapport, i princip kan vara gjord också i de inledande modellanalyserna om sådana justerade behandlingseffekter tas fram. Konsekvensen blir då att det som tidigare var en mer hypotetisk känslighetsanalys byggd på bedömningar av hur effektstorleken kan förväntas bli i en svensk patientpopulation nu med mer stringent metodik skulle kunna bedömas vara tillräckligt relevant och välgrundat för att utgöra underlag för en ny huvudanalys, och den tidigare huvudanalysen baserad på den kliniska prövningens ojusterade resultat istället inkluderas bland känslighets- och scenarieanalyserna.

Det kan emellertid finnas skäl att inte enbart justera effektstorleken i den hälsoekonomiska modellen, mot bakgrund av en ny förväntad effekt för svensk klinisk praxis utan även överväga andra delar av modellen såsom patientsammansättning och konsekvenser av denna. Detta diskuteras kort i följande avsnitt.

Nya effektdata från svensk klinisk praxis och modellanalyser

Om effektstorleken i svensk klinisk praxis skiljer sig från den i den kliniska prövningen på grund av att patientsammansättningen skiljer sig, finns det anledning att se över modellstrukturen och i vilken utsträckning värdena på övriga modellparametrar kan anses rimliga. Om den kliniska prövningen exempelvis hade förhållandevis fler män än vad som observerats i svensk klinisk praxis så bör ändringen i modellen omfatta även denna fördelningsförändring eftersom den kan förväntas påverka behandlingskohortens förväntade överlevnad eftersom livslängden skiljer något mellan män och kvinnor. Samma sak kan gälla andra delar av modellen och de antaganden som görs. Vid introduktion av nya läkemedel gör TLV och andra myndigheter som granskar hälsoekonomiska modeller av kostnadseffektivitet redan idag en granskning av modellens relevans för svenska patienter och gör justeringar i samråd med företagen eller egna analyser. Det som skulle kunna bli nytt med den presenterade strategierna är att den kliniska prövningens resultat för effektdata i större utsträckning viktas om efter svensk klinisk praxis.

Emellertid kan det vara fler faktorer som skiljer mellan populationen från den kliniska prövningen och populationen i svensk klinisk praxis. Om patientpopulationerna avviker så mycket att de i vissa dimensioner inte i stort överlappar så bör analysen av kostnadseffektivitet ta hänsyn till skillnaderna. Ett sätt är att redovisa två parallella bedömningar en av kostnadseffektivitet för de patientgrupper där egenskaperna från den kliniska prövningen och klinisk praxis överlappar, en annan av kostnadseffektivitet för den subgruppen som inte matchar populationen i den kliniska prövningen. I det förra fallet kan denna rapports strategier tillämpas, men i den senare blir bedömningen av kostnads effektivitet mer osäker och ett större behov av scenarie- och känslighetsanalyser föreligger. Osäkerheten beror på att modellen behöver vila på bedömningar av kliniska experter för att täcka in relevanta patientgrupper som inte inkluderats i den kliniska prövningen. Den sammantagna bedömningen beror då på hur robust den skattade inkrementella kostnadseffektkvoten för skillnaderna mellan patientgrupperna är. Om den grupp som inte motsvaras av de kliniska prövningarna är liten, blir påverkan på de övergripande resultaten liten. Det går inte på förhand att ange hur stor en eventuell avvikelse blir utan varje fall behöver analyseras för sig utifrån beslutsproblemets specifika förutsättningar.

Strategierna i denna rapport för att kombinera kunskap om effekt från randomiserade kliniska prövningar och uppföljningsdata i svensk klinisk praxis är relevanta både vid inledande skeden och vid uppföljning när faktiska uppföljningsdata från svensk klinisk praxis föreligger. I det senare skedet kan effektskattningar antingen baseras på mätningar av behandlade patienter som jämförs med kontrollgruppen i de kliniska prövningarna eller om det finns relevanta jämförelsegrupper i svensk sjukvård, exempelvis som historiska kontroller eller om behandlingen införs stegvis och i olika takt i olika regioner.

Vid en strukturerad uppföljning av ett nytt läkemedel där också effekten mäts, kan dessa mått också användas i nya modellanalyser av kostnadseffektiviteten. I de fall där läkemedlet används väsentligen för samma patientgrupper som i den pivotala kliniska prövningen, kan jämförelser mellan beräknad kostnadseffektivitet i olika scenarier med alternativa effektstorlekar göras. I de fall då den beräknade kostnadseffektkvoten påverkas påtagligt kan det vara utgångspunkt för att revidera priset på läkemedlet. Det går inte på förhand att säga om priset i så fall skulle behöva sänkas (på grund av att effekten över överskattad) eller höjas (om uppföljningsdata visar större effekter än den tidigare kliniska prövningen). Lika viktigt som att mäta och analysera effektdata i klinisk praxis är det att också följa upp faktisk resursanvändning och kostnader i klinisk praxis. Det är väl känt att protokollstyrd resursanvändning från kliniska prövningar är ett dåligt underlag för att beräkna behandlingskostnader. Det är möjligt att en välstrukturerad uppföljning av faktisk resursanvändning och kostnader kan vara minst lika betydelsefull för att göra en korrekt bedömning av vilka värden ett nytt läkemedel får.

8 Etiska aspekter

Behovet av prioriteringar

Utveckling av metoder för att bättre kunna mäta effekten av nya läkemedel i klinisk vardag motiveras av behovet av prioriteringar i sjukvården och samhället i stort. Svenska folket, genom sina demokratiskt valda ombud, avgör i vilken grad samhällets resurser ska användas till olika gemensamma åtaganden varav hälso- och sjukvården är ett. De resurser som tas in täcker till en del, men inte fullt ut, de behov av hälso- och sjukvård som finns. Den demografiska utvecklingen (fler äldre innebär mer behov av sjukvård) och den medicinska utvecklingen (fler behandlingsmetoder och teknologier som tillåter behandling av fler tillstånd, i mer avancerade stadier och i högre åldrar) gör att klyftan mellan utbud och efterfrågan ökar. Detta innebär att behovet av att prioritera tillgängliga resurser blir allt tydligare. Att öka sjukvårdens totala resurser (på bekostnad av någon annan sektor) eliminerar inte behovet av prioritering, men ändrar nivån på samhällets betalningsvilja och gör att fler insatser i sjukvården kommer att bedömas vara kostnadseffektiva.

Riksdagen har beslutat vilka riktlinjer för prioritering av resurser i hälso- och sjukvården som gäller (proposition 1996/97:60). Dessa riktlinjer bygger på tre principer.

- Människovärdesprincipen, vilken hindrar särbehandling utifrån individers egenskaper och samhällsställning,
- Behovs- och solidaritetsprincipen som ger företräde till de med de största behoven, samt
- Kostnadseffektivitetsprincipen, som slår fast att kostnaden för en insats måste stå i rimlig proportion till dess nytta.

Dessa tre principer bildar en enhet och kan inte användas var för sig. De metoder som presenteras i denna rapport berör dock främst användningen av kostnadseffektivitetsprincipen, även om behoven också är en funktion av sjukdomssvårighetsgrad och möjligheten att behandla. De övriga principerna kommer dock endast undantagsvis diskuteras vidare.

Problem med att mäta nytta

För att en behandlingsmetod ska kunna prioriteras krävs ett korrekt mått på dess nytta. Skillnaden i nytta mellan en ny metod (den behöver inte vara ny, men som föreslås att introduceras i hälso- och sjukvården) och den mest relevanta referensmetoden (jämförelsealternativet) sätts i relation till skillnaden i kostnad för användning av respektive metod. Det är denna inkrementella kostnads/nyttokvot som visar hur kostnadseffektiv en ny behandlingsmetod är. Idag baseras flertalet beräkningar av kostnadseffektivitet på resultat som erhållits från RCT:er. Eftersom patienter som ingår i RCT-prövningar som regel är starkt selekterade utifrån ålder och övrig sjuklighet så motsvarar de sällan den patientgrupp som senare kommer att behandlas. Utfallet av en behandling i klinisk vardag kan därför skilja sig från det som uppmätts i en RCT [38,39]. Det gäller dock även utfallet i kontrollgruppen i RCT:n jämfört med en obehandlad grupp (eller grupp som enbart fått tidigare standardbehandling) i klinisk vardag. Hur effekten (skillnaden i utfall mellan behandlings- och kontrollgrupp) skiljer sig mellan RCT och klinisk vardag är därför oklart. Osäkerheten gäller också om det finns skillnader i nytta mellan subgrupper av individer och därmed skillnader i kostnadseffektivitet mellan subgrupper. När prioriteringar baseras på osäkra underlag finns det en stor risk för att beslut i praktiken innebär suboptimering och orättvisa prioriteringar. Det finns därför ett moraliskt imperativ i att försöka minska osäkerheten så gott det går med uppbyggande av rimliga resurser för detta arbete. De metoder som beskrivs i denna rapport kan ge en bättre uppskattning av hur stora sådana skillnader är och därmed ge ett bättre underlag för att beräkna en ny/alternativ behandlingsmetods kostnadseffektivitet. Detta är dock endast en liten del av den information om en behandlingsmetods effekter och biverkningar i klinisk vardag som en god uppföljning kan generera och som succesivt minskar osäkerheten om hur den ska användas.

Att använda data från RCT:er för att förbättra kvaliteten på uppföljningen av behandling i klinisk vardag innebär också att det bidrag och den risk som RCT-deltagarna valt att ta ges ytterligare värde.

Hantering av olika kunskapsunderlag

Om effekten i klinisk vardag är annorlunda än vad som uppmätts i RCT:er riskerar en beräkning av kostnadseffektivitet utifrån effektmätning i RCT att felskatta kostnadseffektiviteten jämfört med den man har vid användning av behandlingsmetoden i klinisk vardag. Detta skulle i vissa fall kunna innebära att kostnadseffektiviteten, när den beräknats utifrån RCT-data, är inom den nivå som är rimlig medan en sämre effekt av behandlingen i klinisk vardag gör att kostnadseffektiviteten i verkligheten är orimligt låg. Att införa eller fortsätta att använda behandlingsmetoden innebär i regel att annan angelägen vård undanträngs. Detta är resursslöseri.

Detta resonemang bygger i grunden på ett utilitaristiskt grupptänkande där mesta möjliga nytta till största antal individer är målet. Behovs- och solidaritetsprincipen, som säger att de med störst behov (de svårast sjuka) ska ha företräde, styr dock tillämpningen av kostnadseffektivitetsprincipen bort från en rent utilitaristisk användning. Detta är också en tydligt uttalad intention i propositionen och riksdagsbeslutet där man poängterar att en liten nyttovinst hos många individer inte får tränga undan möjligheten att erbjuda en större nyttovinst till färre individer med svårare tillstånd.

Användningen av de beskrivna metoderna innebär således inget etiskt dilemma i sig. Det är snarare så att dagens praxis som baseras på användning av RCT-data för att beräkna kostnadseffektivitet utan att uppdatera beräkningarna med hjälp av effektdata från klinisk vardag, riskerar att medföra en felanvändning av tillgängliga resurser. Metoderna har således potential att leda till en mer effektiv och ändamålsenlig användning av samhällets resurser vilket kan betraktas som positivt ur etisk synvinkel. Införandet av de nya metoderna aktualiserar dock vissa andra etiska frågor som diskuteras i det följande.

Även om de presenterade metoderna innebär att man i möjligaste mån minskar de problemen med selektionsbias som är sammanfogad till observationella data kvarstår dock alltid en osäkerhet om behandlingsgruppens jämförbarhet med kontrollgruppen. Denna osäkerhet måste också vägas in när beslut tas på grundval av beräkningar av nyttan med dessa metoder.

Rättviseproblem i ett övergångsskede

Om användning av de beskrivna strategierna innebär en snabbare övergång från en tidigare statisk beräkningsmetod för kostnadseffektivitet till en mer dynamisk kostnadseffektivitet skulle det under ett övergångsskede innebära vissa rättviseproblem mellan olika patienter och olika företag. De behandlingsmetoder som introducerats i tidigare skeden kommer att ha genomgått sedvanlig hantering och bedömning av kostnadseffektivitet medan det för senare utvecklade behandlingsmetoder även kommer att genomföras uppdaterade beräkningar av kostnadseffektivitet med hjälp av data från klinisk vardag. Olika krav kommer därmed att ställas på nya respektive gamla behandlingsmetoder. Detta problem kan hanteras genom att även äldre behandlingsmetoder testas. Men det kommer inte med säkerhet finnas data tillgänglig för motsvarande analys av effekt i klinisk vardag för dessa äldre behandlingsmetoder, se avsnitt Rättviseproblem mellan olika bransch, i detta kapitel.

Konsekvenser av ett ändrat kunskapsläge

I framtiden kan således behandlingsmetoder som introducerats i större utsträckning komma att omvärderas hälsoekonomiskt och ur prioriteringssynpunkt när data från klinisk vardag blir tillgänglig och analyseras med de strategier som presenteras i denna rapport. För de behandlingsmetoder som då inte är kostnadseffektiva (när kostnaden inte längre står i rimlig proportion till nyttan) kan priset behöva justeras ner för att uppnå kostnadseffektivitet med dessa nya data. Eftersom prisnivån i ett land påverkar prissättningen även i andra, ibland kommersiellt betydligt viktigare, länder kommer detta inte alltid att vara möjligt. Vissa behandlingsmetoder kan då behöva utrangeras (man upphör att använda dem).

Det kan också handla om utrangering av vissa metoder i speciella subgrupper med så låg grad av nytta jämfört med andra grupper, att det inte längre är försvarbart ur prioriteringssynpunkt att erbjuda dem behandlingen.

Att avsluta pågående behandling

Ett etiskt problem som behöver hanteras är hur man väljer att göra med patienter som startat behandling innan en omvärdering. I de fall behandlingen är begränsad i tiden kan det anses rimligt att låta dessa patienter avsluta påbörjad behandling. I de fall behandlingen är långvarig eller till och med livslång inställer sig frågan om det kan anses motiverat att fortsätta i åratal med en behandling som inte längre anses kostnadseffektiv.

För vissa långvariga behandlingar kan det vara möjligt att få en rimlig uppfattning om effekten på enskilda individer i klinisk praxis, exempelvis vid behandling av en manifest metastaserande tumörsjukdom eller sjukdomar som hypertoni eller diabetes. Avsaknad av en kontroll gör dock att även sådana individuella bedömningar av nyttan är osäkra. För andra behandlingssituationer är en värdering av den enskilda patientens nytta av pågående behandling omöjlig, exempelvis vid adjuvant behandling av tumörsjukdomar och all profylaktisk behandling där validerade surrogatutfall saknas.

Att avsluta en pågående behandling med en positiv effekt (i den mån detta kan avgöras på individuell nivå) eller förväntad positiv effekt, för att den inte är kostnadseffektiv är svårt för alla inblandade. Att inte avsluta en påbörjad behandling innebär dock en tydlig orättvisa mot tillkommande patienter som inte får tillgång till samma behandling trots lika stort behov och mot andra patienter som nekas tillgång till andra, mer kostnadseffektiva behandlingsmetoder (den så kallade alternativkostnaden).

Detta principiella problem är ingalunda nytt utan har länge funnits i samband med kliniska prövningar av behandling av kroniska sjukdomar där patienter startar en behandling innan man vet om den är effektiv eller kostnadseffektiv. I de fall den aktuella prövningen visar att behandlingen har en klinisk effekt uppkommer frågan om hur fortsatt behandling ska tillhandahållas när prövningen avslutats. Det slås fast i Helsingforsdeklarationen (§34) att detta ska arrangeras i förväg av sponsor (företaget), forskare och värdlandet samt att patienten ska informeras om detta. Det är dock ofta oklart vem som har finansieringsansvaret för fortsatt, klinisk behandling efter en avslutad klinisk prövning. På samma sätt kan det vara rimligt att patienter som påbörjar en behandling med ett nytt läkemedel redan initialt informeras om att beslutet att använda detta läkemedel ännu vilar på osäker grund och kan komma att omprövas och att det kan komma att innebära att patientens behandling kommer att avslutas.

De etiska frågorna om att avbryta en behandling i dessa olika situationer, behöver belysas ytterligare. De är omfattande, mångfacetterade och komplicerade och kan därför inte kartläggas eller utredas till fullo inom ramen för denna rapport. Det finns dock ett stort behov av att detta arbete drivs vidare av därför lämpade organ.

Rättviseproblem mellan olika branscher

De behandlingsmetoder som idag kan undersökas vad gäller effekt i klinisk vardag är i första hand läkemedel, som är mer dokumenterade avseende nyttan än andra metoder i hälso- och sjukvården. Läkemedel kommer därför att bli föremål för en kontinuerlig granskning till skillnad från andra behandlingsmetoder i hälso- och sjukvården, vilka kan komma att både introduceras och fortsätta användas utan förnyade undersökningar av effekt i klinisk vardag. Detta skulle kunna innebära en olikartad hantering av metoder från olika företag inom life

science-området. Det är därför angeläget ur såväl prioriteringssynpunkt som rättvisesynpunkt att olika behandlingsmetoder på sikt bedöms enligt samma kriterier och har samma villkor för införande och utrangering. Detta skulle innebära krav på samma omfattning och kvalitet på data för övriga behandlingsmetoder som gäller för läkemedel.

Eftersom marknaden för behandlingsmetoder inom hälso- och sjukvård är global är det nödvändigt att frågor om villkor för olika aktörer även drivs internationellt. Man kan dock notera att FDA i USA redan idag ställer betydligt högre krav på dokumentation av medicintekniska produkter än vad som är fallet inom EU. Det borde därför finnas förutsättningar för en mer likartad bedömning av olika behandlingsmetoder. De svenska kvalitets- och hälsodataregistren bör således innehålla den typ av data som krävs för att kunna jämföra effekt i klinisk vardag med data från RCT-prövningar. När det gäller icke-kommersiella behandlingsmetoder, sådana som utvecklats inom akademien utan involvering av någon kommersiell aktör, kan nationella myndigheter och internationella vårdprogramgrupper spela en roll för att kräva ökad dokumentation innan sådana behandlingsmetoder introduceras. Sådan dokumenterad forskning stärker också möjligheten att erbjuda patienter mer ändamålsenliga behandlingar och stärker också patientens autonomi och möjlighet till delaktighet i besluten.

9 Diskussion

Varför behövs denna rapport?

Det är angeläget att utveckla och använda strategier för att få en bättre uppfattning om nyttan av de behandlingsmetoder som används eller är på väg att introduceras i klinisk praxis. Att ha system för att säkerställa att nytta med en behandling är större än risken är visserligen av stor vikt för patientsäkerheten och svensk sjukvårds trovärdighet. Det torde dock också vara viktigt för medborgarna och deras valda ombud att myndigheternas regelverk garanterar att de behandlingar som ges är ändamålsenliga och kostnadseffektiva. Likaså är det viktigt för läkemedelsföretagen att regelverket stödjer en konsistent handläggning och att priset för produkten står i rimlig relation till patientnyttan. Med nytta avser vi i detta sammanhang den effekt (utfallet vid behandling minus det utfall som skulle ha erhållits vid behandling med jämförelsealternativet) som uppmätts med hänsyn taget till den livskvalitetsnivå som förknippas med tillståndet, med respektive utan behandlingen (QALY).

Olika angreppssätt

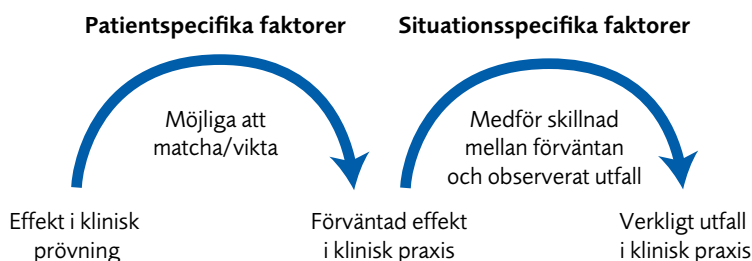
Man kan tänka sig olika angreppssätt och strategier för att erhålla en tillräckligt tillförlitlig uppskattning av patientnyttan i klinisk praxis som underlag för olika former av beslut. En strategi är att genomföra en RCT med patienter i klinisk praxis. Till exempel har registerrandomiserade studier föreslagits och genomförts, framförallt för medicintekniska produkter. Men det råder en oklarhet om regelverket för kliniska läkemedelsprövningar och dess kvalitetskrav på sponsorer (monitorering av överföring av klinisk data till prövningsregister) i

praktiken möjliggör denna prövningstyp även för nya läkemedel. Många nya behandlingar har, i de prövningar som ligger till grund för marknadsföringsgodkännande av LV, visats medföra betydande effekter i form av förlängd överlevnad mm. Det vore i detta läge svårt och till och med oetiskt att genomföra en ny RCT, denna gång i klinisk praxis. Det är bland annat i sådana situationer, då man har en kohort av behandlade patienter utan en kontrollgrupp att jämföra med, som man kan ha nytta av de strategier som beskrivs i denna rapport.

Det gemensamma för de strategier som diskuteras i denna rapport är att man med hjälp av information om kända faktorer av betydelse för både urvalmekanism (eller selektionen till RCT) och som kan modifiera utfallet av en behandling, erhåller en uppfattning om vilket utfall man kan förvänta sig vid behandling av patienter i klinisk praxis med såväl standardbehandling som en ny, experimentell behandling. Rent tekniskt kan dessa strategier baseras på matchning av patienter från kliniska prövningar gentemot patienter i klinisk praxis och viktning av patienter i RCT utifrån frekvensen av dessa kända faktorer i klinisk praxisgruppen (Kapitel Resultat). Man kan således erhålla en rimlig uppskattning av hur kända patientspecifika faktorer påverkar vilken effekt man kan förvänta sig av en behandling i en kohort av patienter i klinisk praxis. Patientspecifika faktorer har att göra med patientens och sjukdomens egenskaper (exempelvis ålder, samsjuklighet, stadium och undergrupp) och påverkar både utfallet och sannolikheten för att en patient ska inkluderas i en klinisk prövning. Det förväntade utfallet i en behandlad kohort kan sedan jämföras med det verkliga utfallet i klinisk praxis. Skillnader i situationsspecifika faktorer (sådant som skiljer kliniska prövningssituationen från situationen i klinisk praxis) kan man dock inte kompensera för och dessa kan därför påverka utfallet och visa sig i form av en skillnad mellan det förväntade och det observerade utfallet i klinisk praxis (Figur 9.1). Sådana situationsspecifika faktorer kan exempelvis vara att ordinationsföljsamheten är bättre när patienterna kontrolleras tätare, att doseringen är högre när ett strikt protokoll följs, att patienter som behandlas på en universitetsklinik med specialister och forskningsköterskor får en behandling som skiljer sig från vad som erbjuds i den övriga sjukvården. De presenterade strategierna kan hantera skillnader i patientspecifika faktorer men inte skillnader som beror på situationsspecifika faktorer.

Figur 9.1

Skillnader i utfall mellan en grupp patienter i klinisk prövning och en annan grupp i klinisk praxis beror på patientspecifika faktorer och situationsspecifika faktorer. Dessa går att använda för att beräkna ett förväntat utfall i en patientgrupp i klinisk praxis. Situationsspecifika faktorer kan medföra att det förväntade och det verkliga utfallet i klinisk praxis skiljer sig åt.



Strategiernas antaganden, problem och begränsningar

Det finns dock en rad antaganden, problem och begränsningar som ökar osäkerheten i beräkningen av den förväntade effekten av en behandling i klinisk praxis.

- Ett antagande är att man känner till alla faktorer av betydelse för selektionen till RCT och utfallen. Om detta antagande inte stämmer finns en risk för över- eller underskattning av utfallet i klinisk praxis.
- Om antagandet stämmer återstår frågan om den nödvändiga informationen om de speciella och betydelsefulla faktorerna finns tillgänglig, både för RCT-patienterna och för patienterna i den kliniska praxiskohorten. Det finns anledning att kartlägga datatillgången i såväl de datafiler för RCT:er som företagen förfogar över som för de kvalitetsregister och hälsodataregister som sannolikt är den viktigaste datakällan i klinisk praxis. Detta är dock ett omfattande arbete som inte ryms inom ramen för denna rapport.
- De strategier som tas upp i rapporten kan inte användas för att generalisera resultat utanför den population som undersöks i den RCT som används, vilket i princip innebär att generaliserbarheten begränsas av de inklusions- och exklusionskriterier som används i RCT:n. För övriga patienter i klinisk praxis som inte ingår i prövningen kan man inte dra några slutsatser. Det kan till exempel gälla patienter som exkluderats från den kliniska prövningen på grund av en absolut gräns för kronologisk ålder eller sådan grad av samsjuklighet att det utgjort ett exklusionskriterium. Således kan dessa strategier bara ge information om det förväntade utfallet och effekten av en behandling för en delmängd av alla patienter som behandlas i klinisk praxis. Med andra ord, även om man ser att det verkliga utfallet i en kohort i klinisk praxis motsvarar det förväntade utfallet samt att detta innebär att effekten är god i relation till den som setts i klinisk prövning, kan man inte utgå ifrån att detta gäller även för patienter utanför den typ av patienter som ingått i RCT:n.

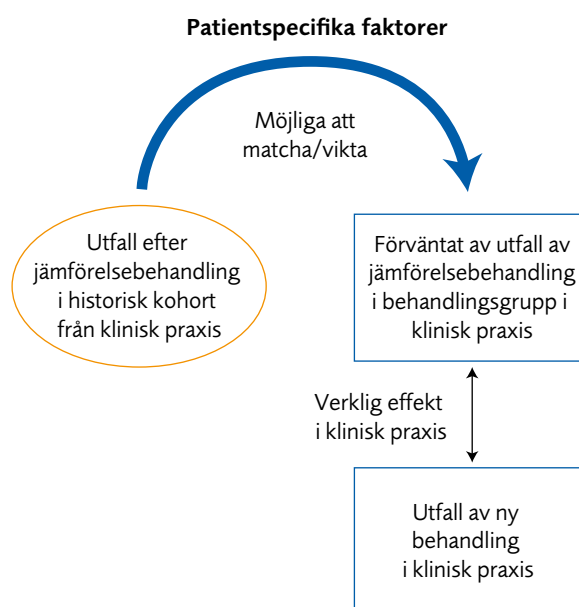
Alternativa datakällor

De presenterade strategierna är i första hand avsedda att justera data från kliniska prövningar för att erhålla en uppskattning av effekten i en kohort i klinisk praxis. Man kan även tänka sig att använda andra datakällor i avsikt att få en uppfattning om effekten (och patientnyttan) av en ny behandling i klinisk praxis i uppföljningssituationen. Historiska kontroller från samma patientgrupp, exempelvis från samma kvalitetsregister, vilka erhållit adekvat jämförelsebehandling är ett alternativ (Figur 9.2). Dessa patienter kommer sannolikt att skilja sig avseende patientspecifika faktorer som påverkar utfallet och man kan då använda strategierna för att matcha/vikta denna grupp så att man får en bättre uppfattning om det förväntade utfallet om den aktuella kohorten (vilka erhållit den nya behandlingen) istället skulle ha erhållit den

adekvata jämförelsebehandlingen. På detta sätt erhåller man en matchad kontrollgrupp med vars hjälp man kan beräkna effekten av den nya behandlingen i klinisk praxis. Denna beräkning lider inte heller i samma utsträckning av det ovan beskrivna problemet med skillnader i situationsspecifika faktorer eftersom dessa patienter också är behandlade i klinisk praxis. Det finns dock alltid en risk att faktorer av betydelse för behandlingsutfallet har ändrats över tid vilket försämrar möjligheten att använda historiska kontroller. Sådana faktorer kan exempelvis vara skillnader i understödjande behandling, selektion på basen av ändrad diagnostik och skillnader i behandling orsakade av förändringar i ekonomiska styrsystem. Möjligheten att använda historiska kontroller från kvalitetsregister bygger också på att man prospektivt insamlar alla data om de faktorer som behövs för att göra en rimlig matchning/viktning av de historiska patienterna.

Figur 9.2

Vid användning av historiska kontroller kan de beskrivna strategierna användas för att förbättra jämförbarheten gentemot aktuell behandlingskohort avseende utfallet efter behandling med jämförelsebehandlingen. Man kan på detta sätt skapa en kontrollgrupp som i princip inte skiljer sig vad gäller de situationsspecifika faktorerna (man kan dock inte utesluta att även dessa kan ändras under tiden mellan den historiska kohortens och aktuell kohorts behandling).



Motsvarande historiska data kan också extraheras direkt från elektroniska journaler. Detta bygger på att nödvändig data finns registrerad på ett sätt som tillåter extraktion av dessa med rimlig insats. Detta alternativ begränsas av att journalsystemen skiljer sig inom och mellan landstingen.

Alternativa användningsområden och frågeställningar

Det förväntade utfallet i klinisk vardag med en ny behandling respektive jämförelsebehandlingen kan användas för att ge en uppfattning om den teoretiska nyttan i klinisk praxis redan innan en kohort i klinisk praxis är genomförd, givet att man har kännedom om relevanta bakgrundsvariabler. En begränsning är att man då inte kan beräkna betydelsen av situationsspecifika faktorer. Det krävs också att man har tillgång till information om faktorer av betydelse för utfallet från den tilltänkta behandlingskohorten. Fördelen är att denna beräkning kan göras mycket tidigt, redan innan regulatoriska beslut eller beslut om förmån respektive användning av den nya behandlingen börjat gälla.

De strategier som beskrivs i denna rapport kan vara ett alternativt sätt att matcha/vikta patientkohorter i observationsstudier för att förbättra jämförbarheten avseende utfall mellan kohorter. I den mån man har tillgång till tillräckligt bra individdata bör det också vara möjligt att göra en sammanläggning av data från RCT:er och observationsstudier. Det är dock viktigt att komma ihåg att randomisering, i bästa fall, innebär att man får en jämn fördelning mellan behandlingsarmarna av alla kända och okända faktorer av betydelse för utfallet. Medan en väl genomförd matchning/viktning mellan grupper i en observationsstudie bara innebär att man har kontroll över vissa kända och uppmätta prognosfaktorer.

Resursåtgång

Den största delen av arbetet med att ta fram data om det finns nytta med att använda en ny behandlingsmetod består i allmänhet av att samla och bygga upp den informationsmängd som behövs. Det slutgiltiga statistiska analysarbetet tar sannolikt inte lång tid. Därför kan det vara rimligt att testa olika statistiska strategier på samma datasätt för att se om man erhåller samma resultat med olika strategier. Val av statistisk metod kan annars göras utifrån vilken som är mest lämpad beroende på sammansättningen av data eller vana och förtroendet hos den som gör analysen.

Det är förmodligen inte rimligt att följa upp och bestämma nyttan i klinisk praxis för alla nya metoder som införs i vården, av resursskäl. Redan idag bestäms också på vilken ambitionsnivå som nya läkemedel ska följas upp efter introduktionen. För merparten av de nya behandlingar som introduceras i vården såsom medicinteknik och kirurgiska metoder saknas dock data som möjliggör en uppföljning av nyttan i klinisk praxis. Detta medför att man för dessa insatser tvingas göra prioriteringar utan lika tillförlitliga underlag som för läkemedel. Detta medför rättviseproblem som diskuterats i Kapitel 8 Etiska aspekter.

Fortsatt arbete

Denna rapport visar att det finns strategier som skulle kunna användas för att möjliggöra en bättre uppskattning av vilken nytta en behandlingsmetod ger i klinisk praxis. Det återstår dock att testa om de presenterade strategierna är genomförbara i praktiken eller om det finns praktiska, etiska eller legala hinder som omöjliggör dem. Detta steg ingår i samma aktivitet (1.2) inom den Nationella Läkemedelsstrategin som denna rapport. Ett hinder för en bredare användning av de beskrivna strategierna är dock ifall tillgången till nödvändiga data, såväl patientspecifika faktorer som utfallsdata, i praktiken inte går att erhålla ur befintliga datakällor. På sikt handlar detta naturligtvis om vilken ambition profession och beslutsfattare har när det gäller att möjliggöra en detaljerad uppföljning av nyttan av behandlingar som introducerats i hälso- och sjukvården.

10 Projektgrupp, externa granskare, råd och nämnd

Projektgrupp

Sakkunniga

XAVIER DE LUNA
professor, Handelshögskolan
vid Umeå universitet

STEFAN FRANZÉN
fil dr, Registercentrum
Västra Götaland

KATARINA STEEN CARLSSON
fil dr, Institutet för Hälso- och
Sjukvårdsekonomi samt Institutionen
för kliniska vetenskaper, Lunds
universitet

SBU

NAAMA KENAN MODÉN
projektledare

STEN ANTTILA
biträdande projektledare

AGNETA BROLUND
informationsspecialist

JAN LILIEMARK
avdelningschef

ANNA RINGBORG
hälsoekonom

ANNA ATTERGREN GRANATH
projektadministratör

Externa granskare

SBU anlitar externa granskare av sina rapporter. Dessa har kommit med värdefulla kommentarer, som i hög grad bidragit till att förbättra rapporten. I slutversionen av rapporten är det möjligt att SBU inte kunnat tillgodose alla ändrings- eller tilläggsförslag från de externa granskarna, bland annat därför att de inte alltid varit samstämmiga. De externa granskarna står därför inte nödvändigtvis bakom samtliga slutsatser eller andra texter i rapporten.

Externa granskare har varit:

MATS ELIASSON

professor, överläkare, Institutionen för folkhälsa och klinisk medicin, Umeå universitet

JUNI PALMGREN

professor, Institutionen för medicinsk epidemiologi och biostatistik, Karolinska institutet

LARS-ÅKE LEVIN

professor, Institutionen för medicin och hälsa, Linköpings universitet

Bindningar och jäv

Sakkunniga och granskare har i enlighet med SBU:s krav inlämnat deklARATION rörande bindningar och jäv. Dessa dokument finns tillgängliga på SBU:s kansli. SBU har bedömt att de förhållanden som redovisas där är förenliga med kraven på saklighet och opartiskhet.

SBU:s vetenskapliga råd – Brage

SBU:s vetenskapliga råd har granskat det vetenskapliga underlaget i rapporten.

LARS HANSSON

ordförande, professor, vårdvetenskap, Lunds universitet

ARNE GERDNER

professor, socialt arbete, Hälsohögskolan i Jönköping

CHRISTEL BAHTSEVANI

leg sjuksköterska, med dr, vårdvetenskap, Malmö Högskola

LENNART ISELIUS

docent, Hälso- och sjukvårdsdirektör, Landstinget Västmanland

PER CARLSSON

professor, hälsoekonomi, Linköpings universitet

MUSSIE MSGHINA

docent, överläkare, psykiatri, Karolinska Universitetssjukhuset

BJÖRN-ERIK ERLANDSSON

professor, medicinteknik, KTH, Stockholm

LARS SANDMAN

professor, vårdetik, Högskolan i Borås

BRITT-MARIE STÅLNACKE
professor/överläkare, rehabiliterings-
medicin, Umeå Universitet

SVANTE TWETMAN
professor, tandvård, Halmstad samt
Köpenhamns Universitet

SBU:s nämnd

SBU:s nämnd har fattat beslut om slutsatserna i rapporten.

NINA REHNQVIST
ordförande, professor,
Karolinska Institutet

STEFAN LINDGREN
professor, ordförande,
Svenska läkaresällskapet

SUSANNA AXELSSON
generaldirektör, SBU

STIG NYMAN
ordförande, Handikappförbunden

HEIKI ERKERS
förbundsordförande,
Akademikerförbundet SSR

SVEN OHLMAN
med dr, Socialstyrelsen

EVA FRANZÉN
forsknings- och utvecklingschef,
Statens Institutionsstyrelse

SINEVA RIBEIRO
förbundsordförande, Vårdförbundet

ÅSA HIMMELSKÖLD
sektionschef,
Sveriges Kommuner och Landsting

AGNETA VON SCHOTING
ordförande, Nationella samverkans-
gruppen för kunskapsstyrning i social-
tjänsten (NSK-S)

JAN-INGVAR JÖNSSON
huvudsekreterare för ämnesrådet
för medicin, Vetenskapsrådet

HEIDI STENSMYREN
ordförande Sveriges läkarförbund

BJÖRN KLINGE
professor, Odontologiska
fakulteten, Malmö högskola,
och Karolinska Institutet

ANDERS SYLVAN
landstingsdirektör, Västerbottens
Läns Landsting

LARS-TORSTEN LARSSON
docent, avdelningschef,
Socialstyrelsen

HÅKAN SÖRMAN
verkställande direktör,
Sveriges Kommuner och Landsting

KARIN TENGVOLD
professor emerita,
Linköpings universitet

11 Extern samverkan och referensgrupp

Denna rapport bygger i delar på informationsinhämtning från olika aktörer i området. Förutom referensgruppen med representanter från LV, TLV, Socialstyrelse, SKL, SLL och Läkemedelsindustriföreningen (LIF), Kapitel 10, har projektgruppen diskuterat rapportens frågeställning, möjligheter och begränsningar med kvalitetsregisterägare, akademiska forskare företag som arbetar med liknande frågor samt enskilda läkemedelsföretag vilka äger kliniska studiedata. Följande grupper/personer har träffats:

ANN-MARIE SVENSSON
biträdande registerhållare,
Registercentrum Västra Götaland

SVEN-ÅKE LÖÖV
Stockholms läns landsting

RAGNAR LINDER
MATS ROSENLUND
JULIA ROCKBERG
QuintilesIMS (fd IMS Health,
fd Pygargus)

JOAKIM RAMSBERG
Myndigheten för vård-
och omsorgsanalys

SOPHIE LANGENSKIÖLD
universitetslektor hälsoekonomi
vid Institutionen för folkhälso- och
vårdvetenskap, Uppsala universitet

PER JOHANSSON
Institutet för arbetsmarknads- och
utbildningspolitisk utvärdering
(IFAU)

AMY LEVAL
Nordic Outcomes Research Manager,
Janssen-Cilag

OLA GHATNEKAR
Market Access Manager, Astellas

LIF:s arbetsgrupp Real World Data:

ANINA FRASCHKE
Bayer AB

STINA SALOMONSSON
MSD

MARIA JAKOBSSON
Pfizer

ANDERS BERGLUND
Biogen

SARA LARSEN
Novonordisk

GUNNEL TORSTENSSON
Abbvie

AMY LEVAL, ANNA RINGBORG
GSK

OLA GHATNEKAR
Astellas

Referensgruppen har bestått av:

DOUGLAS LUNDIN
Tandvårds- & läkemedels-
förmånsverket (TLV)

TOMAS CARS
Stockholms läns landsting (SLL)

ROLF GEDEBORG
Läkemedelsverket (LV)

KAROLINA ANTONOV
Läkemedelsindustriföreningen (LIF)

MARCUS GRY
Socialstyrelsen

SUSANNA EKLUND
Sveriges Kommuner och Landsting
(SKL)

12 Ordförklaringar och förkortningar

Allokeringsbias	Systematiska skillnader som beror på hur patienterna har blivit indelade i interventions- och kontrollgrupp. Allokeringsbias kan exempelvis inträffa när patienter med sämre prognos får den aktiva behandling och patienter med bättre prognos hamnar i kontrollgruppen. En lämplig randomiseringsmetod som är dold för utredaren, t ex. en datorgenererad slumpvalsserie som hanteras av en oberoende funktion, kan förhindra allokeringsbias.
ATE	Average Treatment Effect. Den genomsnittliga behandlingseffekten.
Bakgrundsvariabel	Egenskaper hos deltagare i undersökning, fastställda vid undersökningens början. Exempel: ålder, kön, kroppsvikt, blodtryck, blodanalysvärden, tidigare eller aktuella sjukdomar.
Bedömningsbias	Resultatfel som kan uppstå när en undersökare som bedömer behandlingseffekt med mjuka kriterier (patientens muntliga symtombeskrivning, observation av funktionsstörning, till exempel gångsvårigheter) känner till vilken behandling patienten fått. Undersökaren kan då tendera att göra förmånligare bedömningar hos de patienter som fått den behandling hen tror vara den bästa. Denna typ av fel är ett viktigt skäl för blindning.
Behandlingsbias	Orsakas av systematiska skillnader som är relaterade till hur deltagare som tillhör kontroll- respektive behandlingsgruppen har behandlats i studien, utöver det som har planerats i förväg. Felet kan uppstå när behandlare eller deltagare agerar annorlunda vid vetskap om erhållen behandling. Behandlingsbias kan minskas genom att blinda studien, dvs. behandlare och deltagarna är ovetande om gruppindelningen, och genom att ha ett strukturerat protokoll som styr och kontrollerar viktig övrig behandling och undersökningar.
Behandlingseffekt	Den observerade relationen mellan en behandling och ett utfall
Bias (systematiska fel)	Ett resultatfel som uppstått genom procedurfel, effektbedömningsfel eller annat mänskligt fel under en undersökning; även fel som görs i bedömningen eller hanteringen av resultaten.
Bortfallsbias	Systematiska skillnader mellan jämförelsegrupper i antal som hoppar av en utvärdering

Confounding	Vilseledning när data från en studie tolkas. Uppstår när ett orsaks-samband som man studerar påverkas av en eller flera länkade stör-faktorer (confounders) som inte beaktats. Felaktig slutsats orsakad av confounding kan undvikas genom att studien konstrueras med hänsyn till confounders eller genom att den statistiska resultat-bearbetningen innefattar korrektion för confounders.
Compliance	Följsamhet, en patients (eller frisk försökspersons) efterlevande av medicinska föreskrifter, t ex. beträffande intag av läkemedel. Termen används både vid kliniska prövningar och i det ordinära sjukvårdsarbetet.
Effectiveness	Hur väl en medicinsk åtgärd (t ex. behandling med visst läkemedel) fungerar när den är i allmänt bruk inom sjukvården.
Efficacy	Hur väl en medicinsk åtgärd (t ex. behandling med visst läkemedel) fungerar när den undersöks i en klinisk prövning
Exklusionskriterier	Omständigheter som förhindrar att en person föreslås delta i en undersökning, eller att en viss studie inte tas med i en systematisk granskning. Dessa kriterier ska vara angivna i undersökningens protokoll.
Inklusionskriterier	De betingelser som ska föreligga för att en person ska kunna föreslås att delta i en undersökning. Kriterierna, som ska vara angivna i projektplanen, kan gälla viss sjukdomsdiagnos, åldersgrupp mm. Deltagande i undersökningen kräver dessutom att inget av exklusionskriterierna föreligger.
Kausal inferens	Orsak-verkansamband. Ett tidssamband, dvs. att A händer före B, behöver inte betyda att A orsakar B, dvs. att kausalsamband föreligger. Likaså gäller att en statistisk korrelation inte bevisar kausalsamband. Ett flertal kriterier och metoder kan behöva tillämpas för att ge starkt stöd för att kausalsamband föreligger eller saknas.
LV	Läkemedelsverket
Matchning	Åtgärder för att åstadkomma maximal jämförbarhet i fall-kontrollundersökning. För varje "fall" utses en eller flera "matchade kontroller"; som i flera avseenden ska ha stor likhet med "fallen". I andra jämförande undersökningar utförs i regel inte matchning; jämförbarheten gäller där mellan grupper, och den blir oftast tillfredsställande som följd av randomisering.
NLS	Nationella läkemedelsstrategin togs fram av regeringen och Sveriges Kommuner och Landsting, tillsammans med en bred uppslutning av aktörer på läkemedelsområdet. Syftet med strategin är att få en nationell kraftsamling kring prioriterade frågor inom läkemedelsområdet. Strategin innefattar tre målområden i syfte att nå en bättre läkemedelsanvändning till nytta för patient och samhälle.
NT-rådet	Nya Terapier rådet har landstingsmandat att avge rekommendationer om förhållningssätt till nya läkemedelsterapier.
Observationsstudie	Studie där ingen aktiv åtgärd vidtas. Exempel: ekologisk undersökning, tvärsnittundersökning, kohortundersökning och fall-kontrollundersökning.
Off-label förskrivning	Uttrycket betecknar förskrivning och användning av läkemedel utanför de begränsningar som godkänts av läkemedelsmyndigheten i Europa eller Sverige.
Ordnat införande	Nationellt ordnat införande av nya läkemedel innebär samverkan mellan landstingen, läkemedelsföretagen och myndigheter som Läkemedelsverket och Tandvårds- och läkemedelsförmånsverket. Avsikten är på så goda grunder som möjligt introducera nya läkemedel i vården.

PATC	Population average treatment effect for the controls (genomsnittlig populationseffekt för de obehandlade). Effekten av behandlingen bland de i den bakomliggande populationen som fått referens-behandlingen och kan tolkas som effekten av att behandla de som i klinisk praxis ännu inte behandlas med den nya behandlingen.
PATE	Population average treatment effect (genomsnittlig populations-effekt). Detta är effekten av att behandla alla i den bakomliggande populationen med den nya behandlingen jämfört med att inte behandla någon med det nya preparatet
PATT	Population average treatment effect for the treated (genomsnittlig populationseffekt för de behandlade). Detta är effekten av behandlingen bland de som i den bakomliggande populationen faktiskt blir behandlade och kan tolkas som effekten av behandlingen för de som blir behandlade i klinisk praxis.
Potentiellt utfall	Det möjliga utfallet som observeras när en individ får behandling eller inte får behandling. Varje individ kan ha två potentiella utfall men bara den ena observeras beroende på behandlingsstatus.
Propensity score	Beräknad sannolikhet för deltagande i RCT. Varje individ får en Propensity Score som är mellan 0 och 1, beräknas utifrån de variabler som påverkar deltagande i RCT och utfall av interventionen.
Rapporteringsbias	Snedvridning av publicerade resultat av studier orsakad av att undersökare, ibland också tidskriftsredaktörer, föredrar att publicera undersökningar som visat positivt resultat, t ex. visat att en behandling har effekt. Studier som inte visat någon effekt blir till stor del okända, och bilden av behandlingens värde blir omotiverat gynnsam.
RCT	Randomiserade kontrollerade prövning. Undersökning som avser att fastställa för- och nackdelar med en metod för diagnostik, profylax eller behandling. En prövning bör vara jämförande: på den ena deltagargruppen tillämpas den nya metoden, på den andra en etablerad metod, placebo eller ingen åtgärd. Den sistnämnda gruppen kallas kontrollgrupp. Vanligt är att deltagargrupperna kommer till stånd genom randomisering och att blindning tillämpas.
Regression	En form av statistisk analys av samvariation där det finns en oberoende och en beroende variabel, till skillnad från vad som är fallet vid korrelation.
Responsyta	Resultat som genereras vid regressionsanalys när vi undersöker relationen mellan responsvariabeln (utfallet) och de oberoende variabler (bakgrundsvariablerna)
Selektionsbias	Resultatstörning som orsakas av deltagare inte fördelats slumpmässigt mellan grupperna i en studie. Till exempel den undersökare som svarar för intagningen av deltagare fördelar vissa av dessa till den grupp som han/hon tror vara mest gynnsam för vederbörande. Exempel: i en behandlingsstudie vid högt blodtryck kan undersökaren placera dem med särskilt höga blodtryck i den grupp som får den behandling han/hon tror vara mest effektiv. Förutsättningen är att han/hon känner till fördelningsssekvensen genom att undersökningen inte är blindad eller att han/hon vid en studie som planerats som blindad lyckats skaffa sig kännedom om randomiseringskoden. Selektionsbias medför att de grupper vars behandlingsresultat ska jämföras inte är jämförbara.
SKL	Sveriges Kommuner och Landsting
SUTVA	Stable unit treatment value assumption. Antagandet säger att det potentiella utfallet för en individ inte påverkas av behandlingar som tilldelas de andra individer, dvs. inga interferenser mellan individerna. Vidare säger SUTVA att det enbart finns en version av varje behandling, t ex. styrkan av en viss aktiv substans eller administreringsväg för ett läkemedel. Detta för att motverka olika möjliga utfall som orsakas av olika behandlingsversioner.
TLV	Tandvårds- och läkemedelsförmånsverket

Unconfoundedness	Antagandet kräver att vi känner till/kan mäta alla variabler som påverkar behandlingstilldelning och behandlingsspecifika utfall. Därefter kan vi betinga på dessa variabler och anta att utfallet är oberoende av behandlingstilldelning.
Utbytbarhet	Exchangeability. Grupperna är jämförbara med avseende på bakgrundsvariablerna.
Validitet	Egenskaper hos en undersökning. Intern validitet avser tillförlitligheten hos en undersöknings resultat, medan extern validitet gäller i vilken grad undersökningens resultat har bredare giltighet, till exempel. kan förmodas gälla alla personer med en viss sjukdom.
Viktning	Vägning av data används för att korrigera för den bristande representativiteten hos urvalet.

13 Referenser

1. Sveriges Kommuner och Landsting. Ordnat införande i samverkan, slutrapport från den nationella läkemedelstrategin, delprojekt 6:1 Sveriges Kommuner och Landsting 2014. ISBN 978-91-7585-059-7.
2. Sundell K. Att göra effektutvärderingar. Stockholm, Gothia Förlag, 2012.
3. Anglemyer A, Horvath HT, Bero L. Healthcare outcomes assessed with observational study designs compared with those assessed in randomized Reviews 2014, Issue 4. Art. No.: MR000034. DOI: 10.1002/14651858.MR000034.pub2.
4. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-8.
5. SBU. Utvärdering av metoder i hälso- och sjukvården: En handbok. 2 uppl. Stockholm: Statens beredning för medicinsk utvärdering (SBU); 2014.
6. Reeves BC, Deeks JJ, Higgins JP, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JP, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration. UK: John Wiley & Sons; 2011.
7. Deeks JJ, Dinnes J, D'Amico R, Sowden AJ, Sakarovitch C, Song F, et al. *Evalu-Health Technol Assess* 2003;7:1-173.
8. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health* 2000;21:121-45.
9. Hernan MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health* 2004;58:265-71.
10. Imbens GW, Rubin DB. *Causality: The Basic Framework*. Causal Inference for Statistics, Social, and Biomedical Sciences. New York: Cambridge University Press; 2015.
11. Holland PW. Statistics and Causal Inference. *J Am Stat Assoc* 1986;81:945-60.
12. Morgan SL, Winship C. *Counterfactuals and causal inference: Methods and principles for social research*. New York: Cambridge University Press; 2015.

13. Pearl JG, Glymour M, Jewell NP. Causal Inference in Statistics: A Primer. UK: John Wiley & Sons; 2016.
14. Hernán MA, Robins JM. Causal Inference. USA: Chapman & Hall/CRC, forthcoming; 2016.
15. Fletcher RH, Fletcher SW, Fletcher GS. Clinical Epidemiology: The Essentials. USA: Lippincott Williams & Wilkin; 2014.
16. Rothwell PM. External validity of randomised controlled trials: "to whom do the results of this trial apply?". *Lancet* 2005;365:82-93.
17. Eichler HG, Abadie E, Breckenridge A, Flamion B, Gustafsson LL, Leufkens H, et al. Bridging the efficacy-effectiveness gap: a regulator's perspective on addressing variability of drug response. *Nat Rev Drug Discov* 2011;10:495-506.
18. Srikanthan A, Amir E. Efficacy-effectiveness gap as an obstacle to translating clinical trials to clinical practice. *Eur J Cancer* 2015;51:905-6.
19. Nordon C, Karcher H, Groenwold RH, Ankarfeldt MZ, Pichler F, Chevrou-Severac H, et al. The "Efficacy-Effectiveness Gap": Historical Background and Current Conceptualization. *Value Health* 2016;19:75-81.
20. Stuart EA, Cole SR, Bradshaw CP, Leaf PJ. The use of propensity scores to assess the generalizability of results from randomized trials. *Journal of the Royal Statistical Society: Series A* 2011;174:369-86.
21. Stuart EA, Bradshaw CP, Leaf PJ. Assessing the generalizability of randomized trial results to target populations. *Prev Sci* 2015;16:475-85.
22. Stuart EA, Rhodes A. Generalizing Treatment Effect Estimates From Sample to Population: A Case Study in the Difficulties of Finding Sufficient Data. *Eval Rev* 2016.
23. Verde PE, Ohmann C, Morbach S, Icks A. Bayesian evidence synthesis for exploring generalizability of treatment effects: a case study of combining randomized and non-randomized results in diabetes. *Stat Med* 2016;35:1654-75.
24. Pressler TR, Kaizar EE. The use of propensity scores and observational data to estimate randomized controlled trial generalizability bias. *Stat Med* 2013;32:3552-68.
25. Kern HL, Stuart EA, Hill J, Green DP. Assessing Methods for Generalizing Experimental Impact Estimates to Target Populations. *J Res Educ Eff* 2016;9:103-27.
26. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *Am J Epidemiol* 2010;172:107-15.
27. Hartman E, Grieve R, Ramsahai R, Sekhon JS. From sample average treatment effect to population average treatment effect on the treated: combining experimental with observational studies to estimate population treatment effects. *Journal of the Royal Statistical Society: Series A* 2015;178:757-78.
28. Pearl J. Generalizing experimental findings. *Journal of Causal Inference* 2015;3:259-66.
29. Diamond A, Sekhon JS. Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. *Rev Econ Stat* 2012;95:932-45.
30. Pearl J. Causality: Models, Reasoning and Inference 2nd ed, New York, Cambridge University Press, 2009.
31. Sverige Kommuner och Landsting. Individdata om rekvisitionsläkemedel. Uppföljning, dokumentation och informatik. Nationella Läkemedsstrategin, Område 6.4 Slutrapport 2014. Sverige Kommuner och Landsting 2014.
32. Nationella Kvalitetsregister. Nulägesrapport våren 2016. Satsningen på nationella kvalitetsregister, hämtat från <http://bit.ly/2dbSxIU> den 161012. Nationella Kvalitetsregister 2016.
33. Stockhoms läns landsting. Nationellt ordnat införande av nya läkemedel, hämtat från <http://www.janusinfo.se/Nationellt-inforande-av-nya-lakemedel/> den 161012. Janusinfo, Stockhoms läns landsting.

34. Läkemedelsverket. Stegvis godkännande och införande av nya läkemedel NLS projekt 6.6. Uppsala; Läkemedelsverket; 2014.
35. Lilja B, Miranda-Téllez J, Ljunggren G, Loov S, Wettermark B, Lissmats A, Henriksson R. A Study on Cancer Patients in the Region of Stockholm by Linking Data from Multiple Sources. *Adv Pharmacoepidemiol Drug Saf* 2015;4:187.
36. Jerlvall L, Pehrsson T. eHälsa i lands-
tingen, Inventering på uppdrag av
SLIT-gruppen; 2016.
37. Husereau D, Drummond M, Petrou S,
Carswell C, Moher D, Greenberg D, et al.
Consolidated Health Economic Evalu-
ation Reporting Standards (CHEERS) –
explanation and elaboration: a report of
the ISPOR Health Economic Evaluation
Publication Guidelines Good Reporting
Practices Task Force. *Value Health*
2013;16:231-50.
38. Jernberg T, Janzon M. Sanningen
om den kliniska verkligheten finns i
registren. Registerbaserad randomiserad
klinisk prövning kan ge representativ
bild. *Läkartidningen* 2015;112:1590-1.
39. Shafrin J, Brookmeyer R, Peneva D,
Park J, Zhang J, Figlin RA, et al. The
value of surrogate endpoints for pred-
icting real-world survival across five
cancer types. *Curr Med Res Opin*
2016;32:731-9.